THE NEW SECOND CIRCUIT LOCAL RULES: ANATOMY AND COMMENTARY

Jodi S. Balsam*

ABSTRACT

The New Second Circuit Local Rules provides a general account of the origins, accretion, and renewal of local rules in the federal appellate courts, and specific commentary on the wholesale revision of the Second Circuit's local rules, adopted in 2010. The Second Circuit local rules had not been holistically reappraised in over 100 years when, in 2008, the Court engaged me to spearhead a comprehensive review and rewrite. Among other things, the project undertook to comply with appellate local rulemaking strictures imposed by 1995 amendments to the Federal Rules of Appellate Procedure and by Judicial Conference mandates that originated in a burst of local rule study and superintendence during the 1990s. Notably, the Second Circuit is alone among the thirteen federal appellate courts to conduct such a comprehensive overhaul of its local rules in accordance with the new strictures. This ambitious project has met with great success and provides a role model for other courts considering rules reform.

The article has two objectives: (1) to explain the context of the Second Circuit local rules revision project by providing a history of local rulemaking in the federal appellate courts, including

---

* Acting Assistant Professor of Lawyering, New York University School of Law. I also served as Counsel to the Clerk of the Second Circuit Court of Appeals, and functioned as the reporter and drafter for the Second Circuit’s local rules revision project that culminated in the adoption of the 2010 Local Rules and Internal Operating Procedures. I thank Chief Judge Dennis Jacobs of the Second Circuit Court of Appeal and members of the NYU Lawyering Faculty Scholarship Colloquium for comments on drafts, as well as Zachary Levin for research assistance. The views expressed in this article are solely my own.
efforts at the national level to reform the circuits’ fragmented collections of local rules, and the Second Circuit’s response to those efforts; and (2) to supply a commentary on the new and revised Second Circuit local rules that details how the new rulemaking parameters were applied, and the practical effect of the resulting reforms. It is hoped that an account of how the Second Circuit local rules revision was accomplished may animate and facilitate other federal courts’ efforts at local rules reform, and assist the practicing bar to understand the proper purposes of local rulemaking generally and to navigate the Second Circuit’s new rules.
Table of Contents

INTRODUCTION .................................................................472
I.ANATOMY OF THE NEW SECOND CIRCUIT LOCAL RULES ....475
   A. Local Appellate Rulemaking........................................475
      1. History of Local Rulemaking Power........................475
      2. The Standing Committee’s Local Rules Project ...480
      3. 1995 Amendments to FRAP 47 ..............................486
      4. The Restyled FRAP and the Guidelines for
         Drafting and Editing Court Rules .........................490
   B. History of Second Circuit Local Rulemaking ..........494
   C. Methodology of the Second Circuit Local Rules
      Revision Project ....................................................500
II.COMMENTARY ON THE NEW SECOND CIRCUIT LOCAL RULES ..503
   A. Scope and Organization .........................................504
   B. Docketing the Appeal and Preparing the Record ....504
   C. Electronic Case Filing ...........................................509
   D. Motion Practice ...................................................511
   E. Briefing Schedules and Requirements for Briefs and
      Appendices ..........................................................516
   F. Criminal Appeals ..................................................521
   G. Habeas Corpus .....................................................524
   H. Civil Appeals Management Plan and Related Rules ...525
   I. Oral Argument ......................................................527
   J. Post-Disposition Matters .........................................529
   K. Attorney Regulation ..............................................531
   L. Local Rules Relating to the Organization of the
      Court ....................................................................533
   M. Rescinded Rules ...................................................535

CONCLUSION .......................................................................536
INTRODUCTION

Effective January 1, 2010, the Court of Appeals for the Second Circuit adopted new and revised Local Rules (LRs) and Internal Operating Procedures (IOPs) (collectively, the “local rules”). The new and revised rules are the culmination of a wholesale review of Second Circuit Local Rules taking into account rulemaking strictures that the Federal Rules of Appellate Procedure (FRAP) introduced in 1995 and contemporaneous guidance from the Judicial Conference of the United States (the “Judicial Conference”). Since this 1990s burst of local rulemaking study and superintendence, the Second Circuit is alone among the thirteen federal circuits to conduct such a comprehensive review and rewrite in accordance with the new parameters. The review also coincided with the Second Circuit’s implementation of a new case management system and electronic case filing, presenting an opportunity to issue a comprehensive set of rules that reflect and advance the prevailing methods of processing appeals and administering court business.

The Second Circuit’s local rules were perhaps most in need of an overhaul, given that they were first promulgated in 1892 and had not since been comprehensively audited. The review and revision process exposed an array of defects in this body of rules. The wording of some local rules had not changed for over one hundred years. Other local rules had been overtaken by national rules or federal statutes, and consequently were either impermissibly inconsistent or obsolete. Many rules existed in

1 See 2D CIR. LR 1–47; 2D CIR. IOP A–I. All circuit rules cited in this article are current to December 1, 2010, based on the text of the local rules available on that circuit’s official website.
2 FED. R. APP. P. 47(a) & Note to 1995 Amendments [hereinafter, all citations to the Federal Rules of Appellate Procedure will be abbreviated FRAP].
The New Second Circuit Local Rules

desuetude.\footnote{See infra Parts I.B, II.M.} Filling in the gaps were a patchwork of standing orders, instruction booklets, and website notices.\footnote{See, e.g., COMM. ON FED. COURTS, THE ASS’N OF THE BAR OF THE CITY OF N.Y., APPEALS TO THE SECOND CIRCUIT 55–60 (9th ed. 2007) (deriving instructions for filing and serving briefs from the Civil Appeals Management Plan, the \textit{CAMP Guidelines}, the \textit{Second Circuit Handbook}, the Revised Plan to Expedite the Processing of Criminal Appeals, and forms available on the court’s website).} The need to consult and reconcile multiple sources of guidance and instruction undoubtedly increased the expense and complexity of appellate practice in the Second Circuit, and impaired the court’s efficiency in processing those appeals. In short, the Second Circuit local rules were no longer adequate to convey court policies or to impart necessary guidance to attorneys practicing before the court.\footnote{“Local rules of a court are rules on rules. They had better be guides to navigation or they become submerged rocks and hidden shoals.” Interview with Chief Judge Dennis Jacobs, U.S. Court of Appeals for the Second Circuit, in N.Y., N.Y. (July 6, 2010).}

How the court came to this pass is a cautionary tale that is best understood in the context of the history of local rulemaking generally and the Second Circuit’s experience in particular. This article aims to illustrate this history and how it has been redeemed through an ambitious and auspicious project to rebuild a creaky apparatus of local appellate legislation. Over the course of eighteen months from Summer 2008 through the new rules’ effective date on January 1, 2010, the Second Circuit met the challenges of totally revising its local court rules through the efforts and cooperation of its judges, court personnel, and an advisory committee of private attorneys.\footnote{See infra Part I.C.} It is hoped that an account of how it was accomplished may animate and facilitate other courts’ efforts at local rules reform, and assist the practicing bar to understand the proper purposes of local rulemaking generally and to navigate the Second Circuit’s new rules.\footnote{See, e.g., JEANNE JOHNSON BOWDEN, FED. JUDICIAL CTR., A PRACTICAL GUIDE TO REVISION OF LOCAL COURT RULES (1988), available at http://www.fjc.gov/public/pdf.nsf/lookup/pgrlcr.pdf/$file/pgrlcr.pdf (describing the rules revision process undertaking by the Northern District of Georgia to serve}
Part I of this article presents an anatomy of federal courts local rulemaking in general and in the Second Circuit in particular. Section A of this anatomy charts the origins and history of local rulemaking in the federal circuit courts, describing the sources of this judicial power, and how its exercise led to fragmentation and disunity of appellate practice rules across the thirteen federal circuits. This section also reviews various reform efforts to unify and nationalize appellate practice, including statutory imperatives, survey projects, FRAP amendments to clarify local rulemaking authority, and Judicial Conference standardization of stylistic conventions. Section B of Part I traces the history of local appellate rulemaking in the Second Circuit, and the circuit’s response to federal reform efforts, which has included periods of both inattention as well as innovative responses to the challenges of judicial administration. This section further describes the Second Circuit’s deviation over time from rulemaking standards imposed by FRAP and the Judicial Conference. Section C of the anatomy describes the methodology and parameters the Second Circuit employed in conducting its review and renovation of the local rules.

Part II of the article presents commentary on the new and revised Second Circuit local rules. This part is essentially a catalogue of the practical effects of the alteration or introduction of particular individual rules on Second Circuit appellate practice and court operation, organized to correspond, categorically and chronologically, to how an appellate practitioner might interact with the rules during the course of a case. This section addresses how the governing parameters for revision were applied to particular rules, and how Second Circuit appellate practice differs in the new regime.

I conclude with practical suggestions for how the Second Circuit might maintain the freshness and vitality of its rules going forward and how other circuit courts might similarly take up this enterprise. Local appellate rulemaking may be necessary to accommodate the interpersonal dynamics among a court’s members, its specialized docket management needs, or local
practitioners’ expectations and relationships with the bench. However, variation among local appellate rules should be minimized and mitigated through regular review and compliance efforts that honor FRAP strictures and Judicial Conference guidance.

I. ANATOMY OF THE NEW SECOND CIRCUIT LOCAL RULES

A. Local Appellate Rulemaking

In 1968, the Federal Rules of Appellate Procedure became effective, unifying federal appellate procedure for the first time. FRAP 47, however, reserved the appellate courts’ local rulemaking power. The history of this power and its exercise in the United States courts provide necessary context to a discussion of the Second Circuit’s local rules revision project.

1. History of Local Rulemaking Power

It has long been accepted that courts have inherent power to prescribe local rules of practice and procedure that they deem necessary to conduct their business, so long as such rules are within the scope of the courts’ jurisdiction and authority. In the

---

   Each court of appeals by action of a majority of the circuit judges in regular active service may from time to time make and amend rules governing its practice not inconsistent with these rules. In all cases not provided for by rule, the courts of appeals may regulate their practice in any manner not inconsistent with these rules.
11 McDonald v. Pless, 238 U.S. 264, 266 (1915) (“In the very nature of things the courts of each jurisdiction must each be in a position to adopt and enforce their own self-preserving rules.”); United States v. McSherry, 226 F.3d 153, 155–56 (2d Cir. 2000) (“[F]ederal courts may, within limits, formulate procedural rules not specifically required by the Constitution or the Congress.” (quoting United States v. Johnson, 221 F.3d 83, 96 (2d Cir. 2000))).See also Stephen B. Burbank, The Rules Enabling Act of 1934, 130 U. PA. L. REV. 1015,
Judiciary Act of 1789, the first Congress endorsed local rulemaking power, announcing that “all the said courts of the United States shall have power . . . to make and establish all necessary rules for the orderly conducting business in the said courts, provided such rules are not repugnant to the laws of the United States.” Subsequent judiciary legislation has repeatedly acknowledged the power of courts to make rules necessary for the conduct of their business. When it established the Circuit Courts of Appeal in 1891, Congress again expressly acknowledged the power of each appellate court to specify local rules pertaining to its operation.

Local rulemaking power survived even the 1934 passage of the

12 Judiciary Act of 1789, ch. 20, § 17, 1 Stat. 73, 83.

13 See, e.g., Process Act of 1792, ch. 36, § 2, 1 Stat. 275, 276 (regulating practice in equity and admiralty courts, “subject however to such alterations and additions as the said courts respectively shall in their discretion deem expedient”); Process Act of 1828, ch. 68, § 1, 4 Stat. 278, 278–81 (regulating procedure in common law suits in federal court, “subject, however, to such alterations and additions, as the said courts of the United States respectively shall, in their discretion, deem expedient”); Act to Further the Administration of Justice (“Conformity Act”) of 1872, ch. 255, § 5, 17 Stat. 196, 197 (authorizing the district courts to “regulate their own practice as may be necessary or convenient for the advancement of justice and the prevention of delays in proceedings”). See generally Stephen N. Subrin, Federal Rules, Local Rules, and State Rules: Uniformity, Divergence, and Emerging Procedural Patterns, 137 U. Pa. L. Rev. 1999, 2012 & n.65 (1989) (noting that the Conformity Act did not nullify district court local rulemaking power).

14 Circuit Courts of Appeals Act of 1891, ch. 517, § 2, 26 Stat. 826, 827 (Each of said circuit courts of appeals “shall have power to establish all rules and regulations for the conduct of the business of the court within its jurisdiction as conferred by law.”), amended by Act of Mar. 3, 1911, ch. 231, § 122, 36 Stat. 1087, 1132.
The New Second Circuit Local Rules

Rules Enabling Act (REA), authorizing the Supreme Court to prescribe procedural rules for the lower courts subject to Congressional approval.\(^\text{15}\) The REA was the outcome of decades of efforts to reform the disunity, confusion, and complexity of federal rules of practice and procedure.\(^\text{16}\) The operative provision, Section 2072(a), currently provides: “The Supreme Court shall have the power to prescribe general rules of practice and procedure and rules of evidence for cases in the United States district courts (including proceedings before magistrates thereof) and courts of appeals.”\(^\text{17}\) The Supreme Court exercises its authority to create and amend federal court rules in cooperation with the Judicial Conference and its Committee on Rules of Practice and Procedure, commonly referred to as the “Standing Committee.”\(^\text{18}\)

Although the 1934 REA was intended to centralize and unify federal practice and procedure, it did not abrogate or qualify courts’ authority to promulgate local rules, leaving intact the various provisions of the Judicial Code empowering specified


\(^{16}\) See generally Burbank, supra note 11.


\(^{18}\) The Judicial Conference sits at the top of a three-level rulemaking hierarchy. For appellate rules, the first level is the Advisory Committee on Appellate Rules, which drafts proposed rule changes based on suggestions from interested individuals such as judges, clerks of court, lawyers, professors, and government agencies. The Advisory Committee recommends these rule changes to the second level, the Committee on Rules of Practice and Procedure (the “Standing Committee”). If the Standing Committee approves the rule changes, it then sends the rules to the public for a six-month comment period. After reviewing and synthesizing the public comments, the Advisory Committee meets again to refine the rule and submit it to the Standing Committee for re-approval. If the Standing Committee again approves the rule, it transmits them to the third level—the Judicial Conference. Only after the Judicial Conference approves the rules are they then submitted to the Supreme Court for final approval. If the Supreme Court approves the rules, Congress has a period of time to act to stop a rule adoption or amendment. See generally James C. Duff, *A Summary For the Bench And Bar: The Federal Rules of Practice and Procedure*, ADMIN. OFFICE OF THE U.S. COURTS (Oct. 2010), http://www.uscourts.gov/RulesAndPolicies/FederalRulemaking/RulemakingProcess/SummaryBenchBar.aspx.
courts, other than the Supreme Court, to make rules. REA amendments in 1948 consolidated the provisions authorizing local rulemaking in what is now codified at 28 U.S.C. § 2071(a). The current version of Section 2071(a) authorizes local rulemaking as follows: “The Supreme Court and all courts established by Act of Congress may from time to time prescribe rules for the conduct of their business. Such rules shall be consistent with Acts of Congress and rules of practice and procedure prescribed under section 2072 of this title.”

Procedural reform in the wake of the 1934 REA focused chiefly on adoption of Federal Rules of Civil Procedure governing practice in the district courts. Despite the call for national uniformity, the original version of the Federal Rules allowed for district court local rulemaking. Appellate court local rulemaking did not make it onto the agenda of the Standing Committee for thirty more years, after still more complaints about idiosyncratic variations in practice. Yet, when the Federal Rules of Appellate
The New Second Circuit Local Rules

Procedure took effect on July 1, 1968,\(^{25}\) they also expressly reserved the appellate courts’ local rulemaking power.\(^{26}\) Both the REA and FRAP 47 limited local rulemaking power to those rules relating to a court’s “practice” or “business” that are “consistent” with all other federal statutes.\(^{27}\)

The circuit courts of appeal, in the habit of defining their local rules of practice since their creation in 1891, were not significantly constrained by FRAP, especially given its explicit endorsement of local rulemaking. Local rules continued to multiply in both the district and circuit courts,\(^{28}\) and practitioners complained about divergences of practice and profusion of rules.\(^{29}\) In response, Congress amended the REA in 1988 to require public notice and an opportunity to comment before any federal court adopted local rules, and to empower the Judicial Conference to abrogate any appellate local rule it finds to be inconsistent with federal law.\(^{30}\)

---


\(^{27}\) Rules Enabling Act, 28 U.S.C.A. § 2071(a) (West 2010); FRAP 47(a)(1).

\(^{28}\) See generally supra note 24 (discussing appellate local rulemaking).


\(^{30}\) Rules Enabling Act, Pub. L. No. 100-702, § 403(a), 102 Stat. 4642, 4650–51 (1988) (codified as amended at 28 U.S.C. § 2071(b) (2010)). While the federal courts have generally complied with the notice and comment requirement, the Judicial Conference has yet to exercise its veto power over inconsistent local rules. See Sisk, supra note 29, at 51–52. The 1988 amendments also required each court, other than the Supreme Court, to appoint an advisory committee to study the court’s rules of practice and internal
Meanwhile, as noted in the House Report accompanying the 1988 legislation, the Standing Committee was making progress on a special project to study the local rules problem.  

2. The Standing Committee’s Local Rules Project

Congress had been complaining about the proliferation and inconsistency of local court rules since 1983, leading the Judicial Conference to commission a study of those rules later that year. Under the auspices of the Standing Committee, an initiative known as the Local Rules Project collected and analyzed all the local rules of the federal courts. The Project began with the district courts, operating procedures and make recommendations concerning them. § 403(a), 102 Stat. at 4648–49 (codified as amended at 28 U.S.C. § 2077(b) (2010)). See generally Paul D. Carrington, Learning from the Rule 26 Brouhaha: Our Courts Need Real Friends, 156 F.R.D. 295, 300–01 (1994) (discussing the 1988 amendments to the Rules Enabling Act).


33 Minutes of Meeting of the Committee on Rules of Practice and Procedure, at 5–6 (Jan. 23, 1986) (describing generally the plan for the study of local rules). The Local Rules Project ultimately also studied local rules addressing admiralty and criminal cases, and the Advisory Committee on Bankruptcy Rules surveyed bankruptcy local rules. See Minutes of Meeting of the Advisory Committee on Federal Rules of Appellate Procedure, at 18 (Oct. 23, 1990); Minutes of Meeting of the Committee on Rules of Practice and
and in 1988 issued a comprehensive report on district court local rules. The next phase included a survey of appellate court local rules, but by that time, the business of the Local Rules Project had taken on a statutory imperative. The 1988 REA amendments specifically obligated the Judicial Conference to review local appellate rules to identify and reform those in conflict with national rules. The Judicial Conference delegated the responsibility to review local appellate rules to the Standing Committee’s Advisory Committee on Appellate Rules, and deferred compliance with the statutory directive to abrogate inconsistent rules until completion of the Local Rules Project’s report.

In January 1991, the Local Rules Project issued its report on appellate local rules (the “Appellate Local Rules Report” or “Report”). It contains two sections, each of which organizes local rules according to one of four categories: (1) rules that constitute permissible local variation; (2) rules that repeat existing law; (3) rules that are inconsistent with existing law; and (4) rules that may

34 See DANIEL R. COQUILLETTE & MARY P. SQUIERS, REPORT OF THE LOCAL RULES PROJECT: LOCAL RULES ON CIVIL PRACTICE (1988). This report identified over 5,000 local rules and included a treatise listing every rule that either conflicted with or duplicated national rules. Professor Squiers worked with district courts to reform their practices and most voluntarily deleted or modified their questionable rules. The Local Rules Project also developed a uniform numbering system for district court rules, a set of model local rules, and a manual of administrative rules and forms. See Minutes of Meeting of the Committee on Rules of Practice and Procedure, at 12–13 (Jan. 19–20, 1989). The Civil Justice Reform Act of 1990, 28 U.S.C. §§ 471–82, reversed much of the Project’s work because it encouraged the adoption of new local court rules. This led to renewed complaints about “balkanization” of federal district court practice and the proliferation of local rules. Accordingly, the Standing Committee initiated a second local rules project, using a similar methodology as the first study. See Minutes of Meeting of the Committee on Rules of Practice and Procedure, at 12 (Jan. 10–11, 2002).


37 See APPELLATE LOCAL RULES REPORT, supra note 3.
be topics for FRAP amendment. The Report’s first section is arranged according to FRAP, considering all circuit rules that correspond to a national appellate rule, and analyzing them according to the four categories above. The second section is arranged by circuit, listing every local rule according to that appellate court’s numbering system, and again assigning each to one of the four categories. The resulting compendium exposed layers of clutter and confusion.

As of the date of the Appellate Local Rules Report, the thirteen appellate circuits had promulgated over 1,300 local rules, imposing a major burden on an appellate practitioner with a national practice. Furthermore, the Local Rules Project deemed 33 percent of those rules to be repetitive of national rules and 15 percent to be inconsistent with national rules. The Second Circuit was one of the more egregious offenders, with the Report identifying 53 percent of its local rules to be either repetitive or inconsistent.

Several categories of rules were especially susceptible to circuit court deviation. For example, nine circuits had local rules that contradicted aspects of FRAP 21 concerning extraordinary writs; twelve circuits’ local rules contradicted FRAP 28 and 31 concerning the requirements for formatting, serving, and filing of briefs; and seven circuits’ local rules contradicted FRAP 34 concerning oral argument. This sort of end-run around the

---

38 Minutes of Meeting of the Advisory Committee on Appellate Procedure, at 18 (Oct. 23, 1990).
39 See APPELLATE LOCAL RULES REPORT, supra note 3.
40 See id. The Appellate Local Rules Report calculated this number by counting each rule or part of a rule that its methodology required to be separately addressed.
41 See Memorandum from the Advisory Comm. on Appellate Rules to Chief Judges of the Circuits 2 (Apr. 19, 1991) (on file with author) (providing percentage breakdown of local rules according to the project’s categories).
42 See APPELLATE LOCAL RULES REPORT, supra note 3, at Appendix for Court of Appeals for the Second Circuit.
43 Id. at 31–32 (describing how nine courts have local rules that are inconsistent with portions of FRAP 21); Id. at 47–48, 55–56 (describing how twelve courts have local rules that are inconsistent with various subsections of FRAP 28 and 31); Id. at 62–63 (describing how seven courts have local rules that are inconsistent with portions of FRAP 34).
national rulemaking process was a principal reason Congress authorized the Judicial Conference to flush out inconsistent local rules.44 Also troubling to the Advisory Committee was that many circuits differed in technical requirements, such as formatting of briefs and requirements for the number of copies of various documents.45 These variations could lay traps for the unwary and increase the cost of appellate justice.46 The Report also found fault with local appellate rules for failing to use numbering systems that correlate with FRAP, making it more difficult for litigants to figure out whether a circuit court locally regulates a particular procedural issue.47

The Appellate Local Rules Report was distributed to the circuit courts to serve as a starting point for review of each circuit’s local rules. The Advisory Committee asked the circuit courts to do three things: (1) for local rules identified as inconsistent with FRAP, take steps to eliminate the conflict; (2) for inconsistencies noted in the Report that were not clear to a circuit, obtain clarification from the Local Rules Project director; and (3) for those portions of the Report that a circuit believed to be incorrect, communicate with the project director.48 The circuit courts were to report back to the Advisory Committee and twelve circuits did so.49 However, their responses were as idiosyncratic as their local rules, quibbling with the Local Rules Project methodology, and with the desirability of the project’s core objectives to improve the level of uniformity and decrease the amount of repetition.50 Many of the circuits, including

50 See id. at 5–7. For example, the Eighth Circuit responded that, because it
the Second Circuit, opined that the project used too expansive a definition for inconsistency, venturing that where the national rules are silent, a local rule on the subject is not necessarily inconsistent or inappropriate. In the words of the Second Circuit’s response, “the fact that a national rule sets a base line requirement need not be taken as implying that no greater or more stringent requirement may be imposed by a court of appeals.” 51 Not surprisingly, reform at the circuit level in response to the Report was limited and spasmodic, and virtually nonexistent in the Second Circuit. 52

At the national level, the Advisory Committee mined the Report and circuit reactions to it for topics that might be appropriate for FRAP amendment, in order to reduce the inefficiencies from unnecessary variation and to cull best practices of particular circuit courts into a uniform rule. 53 High priority items on this list included curtailing the authority of local circuit
The New Second Circuit Local Rules

clerks to return or refuse to file documents that do not comply with national or local rules, clarifying procedures for prehearing conferences, and unifying standards for granting a stay of mandate. 54

The dialogue with the circuits also led the Advisory Committee to exhort the Judicial Conference to adopt specific requirements for local rules, including three items that later became the basis for amendment of FRAP 47: (1) a uniform numbering system under which the local rules would be keyed to the national rules, (2) the removal of language in local rules that repeats national rules, and (3) stricter observation of the distinction between local rules and internal operating procedures. 55 The Advisory Committee also recommended that the Judicial Conference establish a process to review new local rules before their implementation. 56 The Judicial Conference has not acted on this recommendation, nor has it ever exercised its authority, conferred by the 1988 REA amendments, to veto a local rule. 57


3. 1995 Amendments to FRAP 47

The criticism implicit in the 1991 Appellate Local Rules Report, along with continuing pressure from the practicing bar, led to the amendment of FRAP 47 in 1995 to clarify the prerogatives and boundaries of local rulemaking. The amendment imposed three new strictures on local rulemaking:

- “A generally applicable direction to parties or lawyers regarding practice before a court must be in a local rule rather than an internal operating procedure or standing order.”


communicated with the circuits in an effort to persuade them to simplify and unify briefing requirements, with little success. See generally 16A CHARLES ALAN WRIGHT & ARTHUR R. MILLER, FEDERAL PRACTICE AND PROCEDURE § 3945 (4th ed. 2009).

58 See Sisk, supra note 29, at 4–5, 5 n.18.
The New Second Circuit Local Rules

- “A local rule . . . must conform to any uniform numbering system prescribed by the Judicial Conference of the United States.”

The amended FRAP 47 recognizes that even when a local rule is not facially inconsistent with FRAP, local rules require the practitioner to master both FRAP and the local rule and then to determine how the two sets of rules interact, thus imposing transaction costs that often outweigh whatever benefit might derive from the local rule. Thus, FRAP 47’s three new strictures were

---

60 FRAP 47(a)(1). The full text of Rule 47 reads:

(a) Local Rules.

(1) Each court of appeals acting by a majority of its judges in regular active service may, after giving appropriate public notice and opportunity for comment, make and amend rules governing its practice. A generally applicable direction to parties or lawyers regarding practice before a court must be in a local rule rather than an internal operating procedure or standing order. A local rule must be consistent with—but not duplicative of—Acts of Congress and rules adopted under 28 U.S.C. § 2072 and must conform to any uniform numbering system prescribed by the Judicial Conference of the United States. Each circuit clerk must send the Administrative Office of the United States Courts a copy of each local rule and internal operating procedure when it is promulgated or amended.

(2) A local rule imposing a requirement of form must not be enforced in a manner that causes a party to lose rights because of a nonwillful failure to comply with the requirement.

(b) Procedure When There Is No Controlling Law. A court of appeals may regulate practice in a particular case in any manner consistent with federal law, these rules, and local rules of the circuit. No sanction or other disadvantage may be imposed for noncompliance with any requirement not in federal law, federal rules, or the local circuit rules unless the alleged violator has been furnished in the particular case with actual notice of the requirement.


61 See Sisk, supra note 29, at 26 (opining that appellate local rules impose a disproportionate expense burden on litigants compared to district court local rules because an appeal proceeds on a faster schedule than a trial so “each hour
calculated to improve local rules’ transparency, clarity and accessibility.62

First, by requiring that a generally applicable direction be set out in a local rule, rather than in an internal operating procedure or standing order, the revision seeks to make it easier for practitioners to identify those local directives that govern practice before a court of appeals. Litigants should not be ambushed by provisional or hard-to-find procedural requirements.63 And circuits should not be permitted to impose procedural burdens in internal operating procedures that, unlike local rules, have not been adequately identified and subject to public notice and opportunity for comment. This provision of FRAP 47 also implicitly delimits the appropriate content of internal operating procedures; they “should not contain directives to lawyers or parties; they should deal only with how the court internally conducts its business.”64 Adherence to this distinction between local rules and internal operating procedures also prevents cluttering the former with administrative minutiae that might obscure the import of instructions to practitioners.65

Second, the prohibition against duplicating the language of the national rules strives for clarity as to which practices are truly local. A rule that contains both the national and local requirements “observes the local variation.”66 Extirpating all reiterative language of attorney time added by the need to discover, understand, and comply with idiosyncratic local rules meaningfully inflates the expense of an appeal”).


63 See id. at 92 (“Placing a practice oriented provision in the internal operating procedures may cause a practitioner, especially one from another circuit, to overlook the provision.”).

64 ADVISORY COMM. ON APPELLATE RULES, REPORT ON LOCAL RULES PROJECT TO THE COMMITTEE ON RULES OF PRACTICE AND PROCEDURE OF THE JUDICIAL CONFERENCE OF THE UNITED STATES 9 (Jan. 8, 1992).

65 See id. at 10.

66 Id.; see also FRAP 47 & Note to 1995 Amendments (“[L]ocal rules should not repeat national rules and Acts of Congress.”).
flags the local variation. Local rules that repeat national rules are also problematic because minor variations, poor paraphrasing, or selective duplication can introduce confusion. The interpretative problems multiply when there is a change in one rule but not the other. Accordingly, the amended FRAP 47 restricts the content of local rules to only those directives that depart from or supplement national rules.

Third, the 1995 amendment mandates conformity to any local rules numbering system prescribed by the Judicial Conference. This amendment reflects the concern that “[l]ack of uniform numbering might create unnecessary traps for counsel and litigants,” and the desire to “make it easier for an increasingly national bar and for litigants to locate a local rule that applies to a particular procedural issue.” In 1991, the Judicial Conference had issued a recommendation to all circuit chief judges that they adopt a numbering system for local rules that corresponds with FRAP, and that recommendation was later adopted as a formal prescription. Accordingly, if a court of appeals promulgates, for example, a local rule governing motions, the court must number the local rule to correlate to FRAP 27, which sets out the national requirements for motions. Using the same number for the local rule and the federal rule covering the same topic improves notice of the existence of the local rule and accessibility to it. It was also hoped that linking the number of a local rule to the corresponding national rule would dispel the inclination to repeat language from the national rules in the local rules.

Remarkably, although the 1995 amendments to FRAP 47 are absolutely binding on the circuit courts, compliance has been

68 Id. (“[T]he restriction prevents the interpretation difficulties that arise when there are minor variations in the wording of a national and a local rule.”).
69 FRAP 47 & Note to 1995 Amendments.
71 See June 1992 ADVISORY COMMITTEE REPORT, supra note 62, at 92.
72 Id.
virtually nonexistent.\textsuperscript{73} Localism continues to advance, diversifying practice in such basic areas as appellate motions, briefing, and oral argument.\textsuperscript{74} According to one commentator:

The appeals courts have expressly ignored the instructions of the High Court and lawmakers by prescribing even more local measures, many of which conflict with or reiterate the Federal Rules of Appellate Procedure or Acts of Congress. The Judicial Conference, however, has never undertaken the rigorous scrutiny of these mechanisms that the Supreme Court and legislators envisioned.\textsuperscript{75}

4. The Restyled FRAP and the Guidelines for Drafting and Editing Court Rules

While the 1995 amendments to FRAP 47 were wending their way through the rulemaking process, the Standing Committee, moving on a parallel track, established a Style Subcommittee to “clarify, simplify, and eliminate inconsistencies in proposed rules amendments.”\textsuperscript{76} The ultimate objective of the Style Subcommittee was to unify the stylistic approaches of each of the advisory committees on appellate, bankruptcy, civil, criminal, and evidence rules, whose disparate modes of rules-drafting had led to “unnecessary ambiguity and the loss of simplicity.”\textsuperscript{77} Respected legal-writing guru Bryan Garner led the style project and developed uniform drafting guidelines detailing a common set of

\textsuperscript{73} See Tobias, \textit{supra} note 57, at 567 (finding “very little evidence that the appellate courts had undertaken efforts to discharge the obligations which the 1995 revision of FRAP 47 or [Congress] imposed.”).
\textsuperscript{74} See Sisk, \textit{supra} note 29, at 7–24.
\textsuperscript{75} See Tobias, \textit{supra} note 60, at 153–54; see also Tobias, \textit{supra} note 57, at 570–72 (offering possible explanations as to why the appeals courts and Judicial Conference have never fulfilled their duties under FRAP 47, including deference to local expertise, notions of professional courtesy, competing demands of increasingly large and complex caseloads, and lack of resources).
\textsuperscript{77} \textit{Id.}
The New Second Circuit Local Rules

style preferences. Initially the Style Subcommittee applied these guidelines only to rules amendments, but eventually was charged with restyling the entire sets of civil and appellate rules.

The first set of rules to be tackled was FRAP. In 1996, the Style Subcommittee offered a comprehensive restyling in accordance with the uniform drafting guidelines. The restyled rules sought to eliminate ambiguities and inconsistencies in FRAP, and generally make the rules more readable by breaking up long narrative passages with section dividers and headings. The changes were intended to be non-substantive, and quickly cleared through the Supreme Court and Congress to become effective on December 1, 1998. As a side-by-side comparison of the redraft with the then-existing rules demonstrates, simply manipulating the format achieved a much clearer presentation. The revision breaks down rules into constituent parts, using progressively indented subparagraphs with headings and substituting vertical for horizontal lists. These basic formatting changes graphically render the structure of the rules and make them easier to read and understand even where wording is essentially unchanged.

Possibly the most significant change is the redraft’s elimination of the use of “shall,” an inherently ambiguous term that can variously mean “must,” “may,” or something else, depending on context. Because “shall” is no longer generally used in spoken or clearly written English, its use exacerbated the potential for

79 Robert E. Keeton, Preface to GUIDELINES, supra note 78, at iii.
80 Proposed Amendments to FRAP, supra note 76.
81 Id. at 123 (introductory note by Judge James K. Logan, Chair of the Advisory Committee on Appellate Rules).
82 Id.
84 Proposed Amendments to FRAP, supra note 76, at 129–273.
85 See, e.g., id. at 161–65 (Rule 10).
confusion. The restyled rules replace “shall” with “must,” “may,” or “should,” depending on which one the context and established interpretation make correct in each rule. The restyled rules also eliminate other ambiguous terms. For example, changing “receives” to “docketed” in FRAP 4(c) eliminates uncertainty as to the deadline for filing a cross-appeal: a court may “receive” a paper in the mail that is not processed for a day or two, making the date of receipt uncertain, while “docketing” is an easily identifiable event.

The restyled rules reduce the use of inconsistent terms that say the same thing in different ways. Because different words are presumed to have different meanings, such inconsistencies confuse the careful reader. For example, FRAP 1, 2, 5, 8, 13, 18, and 24 previously used the word “application” interchangeably with “motion” and “petition.” The restyled rules achieve consistent expression without affecting meaning by eliminating “application” except when it is a term of art, for example, in the context of FRAP 15 applications for enforcement of an agency decision and FRAP 22 applications for a writ of habeas corpus. In the same vein, “considered” replaces “deemed” in FRAP 3(a), 13(b), and 22(b), and “believed” in FRAP 10(b), to achieve consistent expression without changing meaning.

The restyled rules also replace redundant instructions with cross-references to the relevant rule. For example, document-formatting requirements in FRAP 5(c), 5.1(c), and 21(d) are replaced with a cross-reference to FRAP 32(a)(1) describing the required form for all paper submissions.

Anticipating adoption of the restyled FRAP, the Style Subcommittee published Bryan Garner’s *Guidelines for Drafting*
The New Second Circuit Local Rules (the “Guidelines”).92 The booklet served the dual purpose of explaining drafting and editing choices reflected in the revised FRAP, as well as establishing a standard going forward for any court engaged in a rules drafting and editing project.93 Over time, the Guidelines has become the “accepted style for federal rules” and recommended for use in drafting substantive statutes, practice codes, and local rules.94

The Guidelines first sets out “Basic Principles” summarizing its goals: clarity, readability and brevity.95 The second chapter of the Guidelines lists general conventions to effectuate the basic principles, including drafting rules in the present tense, in the active voice, and in the singular number unless the sense is undeniably plural.96 The general conventions also address syntax, instructing the drafter to place conditions, exceptions, and modifiers at the beginning or end of a sentence, and to avoid interruptive and prepositional phrases.97 The Guidelines recommends minimizing “of-phrases” by replacing them with possessives and adjectives.98 The Guidelines also prescribes short sentences of no more than 25–30 words, and specific punctuation to enhance readability.99

The Guidelines’ third chapter establishes organizational principles that put the broadly applicable before the narrowly applicable, the general before the specific, more important items before less important, rules before exceptions, and contemplated

---

92 See GUIDELINES, supra note 78.
93 See Alicemarie H. Stotler, Foreword to GUIDELINES, id. (statement of the Chair of the Standing Committee).
94 See George C. Pratt, Introduction to GUIDELINES, supra note 78, at vi–vii (written by a retired judge and former chair of Style Subcommittee); see also Joseph Kimble, How to Mangle Court Rules and Jury Instructions, 8 SCRIBES J. LEGAL WRITING 39, 42–43 (2002) (recommending GUIDELINES for drafting court rules).
95 GUIDELINES, supra note 78, at 1.
96 See id. at 3–4.
97 Id. at 5–12.
98 Id. at 11–12. For example, GUIDELINES rewrites “the clerk of the court of appeals” as “the circuit clerk”; “statute of the United States” as “a federal statute”; and “failure of an appellant” as “an appellant’s failure.”
99 Id. at 13–15.
events in chronological order.\textsuperscript{100} Along with requirements for structural divisions and enumerations, this chapter of the Guidelines focuses on reformatting rules for clearer presentation.\textsuperscript{101} The fourth chapter of the Guidelines considers particular “Words and Phrases,” with specific pointers for achieving brevity, using the active voice, and avoiding legalese and jargon.\textsuperscript{102} The Guidelines also dictates rules for using words of authority, providing a glossary that disfavors the use of “shall,” replacing it with “must,” “may,” or “should.”\textsuperscript{103}

\textbf{B. History of Second Circuit Local Rulemaking}

Despite the 1995 amendment of FRAP 47 and the 1996 publication of the Guidelines, few circuit courts made efforts to comply with the new substantive prescriptions and style suggestions for local rules.\textsuperscript{104} There is no record that the Second Circuit responded specifically to these developments prior to the comprehensive review that led to the 2010 local rules revision. Notably, the Second Circuit did not conform its local rules to FRAP 47(a)(1)’s new strictures on local rulemaking.\textsuperscript{105} The 1991 Appellate Local Rules Report continued to be a relevant source of criticism in this regard—in particular as to repetitive and inconsistent local rules.\textsuperscript{106}

How the Second Circuit arrived at this juncture is largely a product of its history as one of the first intermediate federal courts,
The New Second Circuit Local Rules

created in 1789 and known simply as “circuit courts.” For the first one hundred years of the republic, circuit courts were primarily trial courts of original jurisdiction with limited appellate jurisdiction. There were no separate circuit judges; Supreme Court justices and district judges presided over the circuit courts.

Three circuit courts were initially established—eastern, middle and southern. Connecticut, New York, and later Vermont were part of a larger group of states comprising the eastern circuit. As the country expanded, more circuits were added, and existing ones were reorganized and numbered, resulting in grouping those three states alone as the “Second Circuit.” The circuit courts’ appellate workload remained sparse; mercantile, patent, and admiralty trials occupied most of the Second Circuit’s docket. To the extent the Second Circuit engaged in local rulemaking, its efforts were principally focused on trial administration.

The burdens of travel and the Supreme Court’s own docket

108 See Judiciary Act of 1789, ch. 20, § 21, 1 Stat. 73, 83 (giving appellate jurisdiction to circuit courts in “causes of admiralty or maritime jurisdiction, where the matter in dispute exceeds the sum or value of three hundred dollars”); id. § 22, 1 Stat. at 84 (permitting a “writ of error” from the district courts in civil actions exceeding “the sum or value of fifty dollars”).
109 See id. § 4, 1 Stat. at 74–75. Subsequent judiciary legislation required that the justices only attend one term of circuit court in each year. See Act of June 17, 1844, ch. 96, § 2, 5 Stat. 676.
110 See § 4, 1 Stat. at 74–75. The eastern circuit originally comprised New Hampshire, Massachusetts, Connecticut and New York, and later Rhode Island and Vermont. See MORRIS, supra note 107, at 10.
112 See MORRIS, supra note 107, at 44–48; see also SCHICK, supra note 29, at 40–41.
113 See, e.g., O. HALSTED, RULES AND ORDERS OF THE UNITED STATES SUPREME COURT, CIRCUIT COURT FOR THE SECOND CIRCUIT AND THE DISTRICT COURT FOR THE SOUTHERN DISTRICT OF NEW YORK (1829) (setting forth procedures governing federal trials); see also Aqua Stoli Shipping Ltd. v. Gardner Smith Pty Ltd., 460 F.3d 434, 439–40 (2d Cir. 2006) (noting the early adoption of local admiralty rules by district courts within the old Second Circuit).
increasingly kept the justices from participating in circuit court proceedings. It was common for a single district judge not only to conduct circuit court proceedings, but also to review his own prior rulings in a case.\footnote{114}{See David R. Stras, *Why Supreme Court Justices Should Ride Circuit Again*, 91 MINN L. REV. 1710, 1721–22 (2007); MORRIS, supra note 107, at 93.} By 1869, circuit riding was placing such a strain on the justices that Congress created circuit judgeships to take on some of the caseload, specifically appeals from district courts.\footnote{115}{See Judiciary Act of 1869, ch. 22, § 2, 16 Stat. 44, 44–45; Morris, supra note 107, at 69–70.} Calls for additional reform continued in the following decades as the country’s population and industry expanded, increasing demands on the Supreme Court.\footnote{116}{Accordingly, in 1891, Congress severed the trial and appellate functions for most of the nation’s federal courts. Congress situated all trials in the district courts and vested appellate jurisdiction in the regional circuit courts, which were staffed with circuit judges.\footnote{117}{The new Circuit Court of Appeals for the Second Circuit, as it was then styled, promulgated its first set of local appellate rules in 1892 soon after it was established.\footnote{118}{There were thirty-four “general rules” and another nineteen rules on admiralty.\footnote{119}{The Second Circuit’s general practice rules adopted the Supreme Court’s rules of practice, “as far as the same shall be}}}} Accordingly, in 1891, Congress severed the trial and appellate functions for most of the nation’s federal courts. Congress situated all trials in the district courts and vested appellate jurisdiction in the regional circuit courts, which were staffed with circuit judges.\footnote{117}{The new Circuit Court of Appeals for the Second Circuit, as it was then styled, promulgated its first set of local appellate rules in 1892 soon after it was established.\footnote{118}{There were thirty-four “general rules” and another nineteen rules on admiralty.\footnote{119}{The Second Circuit’s general practice rules adopted the Supreme Court’s rules of practice, “as far as the same shall be}}}}
The New Second Circuit Local Rules

applicable, \textsuperscript{120} and otherwise regulated all matters great and small that might concern administration of court business, including: composition of the court’s seal, oath to be taken by court officers, preparation of the record on appeal, procedures for obtaining translations of documents in a foreign language, death of a party, format of briefs, oral argument, delivery of opinions, and citation of cases.\textsuperscript{121}

Since the 1892 adoption of the circuit’s first set of local rules, the number and subject matter of the rules have expanded and contracted to adjust to ever-changing modes of practice, technology, and litigation environments. However, the court has handled local rules revisions and amendments piecemeal, as evidenced by the fact that the wording of a handful of the 1892 local rules remained virtually unchanged in 2008.\textsuperscript{122} Little effort was made over the years to recalibrate the overall structure of the rules or to harmonize existing rules with amendments or additions. For example, the 1910 edition of the rules added four new “general rules” some of which addressed topics already the subject of existing rules; these new rules were tacked on at the end rather than integrated with existing rules.\textsuperscript{123}

To be fair, the court occasionally engaged in housekeeping efforts. After the 1911 abolition of the old circuit courts, the Second Circuit excised its rules of admiralty trial administration.\textsuperscript{124} However, the court’s attitude towards its local rules remained somewhat desultory, conceiving them as largely discretionary, to

\textsuperscript{120} Id. at 413 (Rule 8).

\textsuperscript{121} Id. at 412–22.

\textsuperscript{122} Compare id. at 412, 414, 420–21 (2D CIR. R. 1, 2, 3, 4(1), 9, 28 (1892)), with 2D CIR. LR §§ 0.11, 0.12, 0.13, 0.14(1), 0.19, 0.20 (2008). For example, Local Rule 9 from 1892 and Local Rule §0.19 from 2008 read identically: “Process. All process of this court shall be in the name of the President of the United States, and shall be in like form and tested in the same manner as process of the Supreme Court.”


\textsuperscript{124} See Act of Mar. 3, 1911, ch. 231, § 301, 36 Stat. 1087, 1169.
be applied flexibly as circumstances warranted. While this approach may have been intended to permit leniency toward litigants, it also led to a casual neglect of rulemaking standards. Little attention was paid to the format and arrangement of the rules, or to the consistent usage of style and language. Readability, accessibility and clarity all suffered.

By the time the Supreme Court adopted FRAP in 1968, the Second Circuit had trimmed its local rules to twenty-nine. In response to FRAP, the Second Circuit whittled its local rules further down to twenty-one to eliminate redundancy. That round of revisions also resulted in re-categorizing the rules to distinguish between rules relating to the organization of the court and rules supplementing FRAP, which were renumbered to correlate to the national rules. However, over the next forty years, with few exceptions, the Second Circuit local rules were not revised to take into account subsequent amendments to FRAP or other relevant statutory developments. Rules gradually accreted over the years, increasing back to twenty-nine by 1982, and to thirty-eight by 2008.

The Second Circuit was often among the most responsive and innovative of the circuit courts when called upon to meet challenges confronting appellate productivity. The Second Circuit developed the first plan to expedite the processing of criminal appeals and the first mediation program for civil appeals. Yet its

125 See Schick, supra note 29, at 86 (“[T]he practice of the Second Circuit is to apply the [local] rules flexibly.”).
128 See, e.g., FRAP 3(d) (amended in 1979 to require the district clerk to forward to the circuit court a copy of the notice of appeal in all appeals, rendering redundant a similar requirement in 2d Cir. LR 3(d) (2008)); FRAP 15.1 (adopted in 1986 to confirm existing practice in most circuits, thus superseding 2d Cir. LR 15.1 (2008)).
130 See generally Morris, supra note 107, at 170–71; Second Circuit Plan to Expedite the Processing of Criminal Appeals, 28 U.S.C.A., United States
The New Second Circuit Local Rules

local rules were not revised in any significant way in response to the 1991 Local Rules Project recommendations or the 1995 amended FRAP 47.\textsuperscript{131}

As of 2008, the Second Circuit’s local rules were organized in two parts: (1) Rules Relating to the Organization of the Court, and (2) Rules Supplementing Federal Rules of Appellate Procedure.\textsuperscript{132} The first section contained sixteen rules that were not numbered according to FRAP, as the Judicial Conference requires.\textsuperscript{133} Instead, these sixteen rules used a numbering system dating from 1968 that denominated the rules using a section symbol and decimals, for example, “§ 0.14 Quorum.”\textsuperscript{134} The first section of rules also largely dealt with internal administrative matters rather than the type of generally applicable direction to litigants that is the proper subject of local rules.\textsuperscript{135} The second section of rules contained twenty-two rules corresponding to FRAP counterparts, adopted at various times over the decades, inconsistently formatted and styled, a number of which were redundant or obsolete. Their deficiencies included misquoting FRAP, retaining rules that were superseded by statute, and failing to follow explicit FRAP directives.\textsuperscript{136}

The Second Circuit finally turned its attention to its local rules in 2008 when it embarked on a major technological upgrade in order to participate in the United States Courts’ Case Management/Electronic Case Files (CM/ECF) project, which

\begin{itemize}
  \item See supra note 52 and accompanying text.
  \item See supra note 68 and accompanying text.
  \item 2D Cir. L.R. § 0.14 (2008).
  \item See FRAP 47(a)(1).
  \item See, e.g., 2D Cir. LR 11 (2008) (purporting to quote language from FRAP 11(a) that had been amended in 1979); 2D Cir. LR § 0.26 (2008) (setting out rules for filing a category of petition eliminated with the repeal in 1996 of 28 U.S.C. § 636(c)(5)); 2D Cir. LR 30 (2008) (missing sanctions language required by FRAP 30(b)(2) since 1986). See also discussion infra Part II.M.
\end{itemize}
allows for the filing and accessing of electronic case files over the Internet.\(^\text{137}\) In anticipation of this undertaking, which would significantly impact the administration of the court’s business and require some degree of procedural reform, the court found it timely to launch a comprehensive revision of its local rules.

**C. Methodology of the Second Circuit Local Rules Revision Project**

From mid-2008 through 2009, the Second Circuit Local Rules Revision Project undertook a wholesale review of its local rules for revision, updating, and streamlining. The project aimed to take a fresh look at every rule and consider all suggestions for improvement and clarification. The court’s Rules Committee assigned a staff working group\(^\text{138}\) to take a first look at the local rules, and redraft them hewing to the strictures of FRAP and the Judicial Conference’s style Guidelines. According to the project’s methodology, the proposed revision then went through several rounds of vetting by, in order, the Attorney Advisory Committee to the Second Circuit Rules Committee, the Rules Committee, the full court, the public during a period of notice and comment, and finally, the full court again to assess and, if appropriate, incorporate any public suggestions.

As the reporter for the project, I initially undertook a number of preliminary tasks including:

---


\(^{\text{138}}\) The members of the Rules Committee at the time were Chief Judge Dennis Jacobs and Judges Jon Newman and Reena Raggi. The members of the staff working group were Catherine O’Hagan Wolfe, Clerk of Court, Andrew Contreras, Deputy Clerk of Court, Michael Jordan, Counsel to the Chief Judge, and the author, serving as initial drafter and reporter for the project.
The New Second Circuit Local Rules

• Researching relevant law, determining the history of existing court rules, and researching other courts’ local rules;

• Reviewing historical documents and court files relating to local rules promulgation;

• Interviewing current and former members of the court’s Rules Committee and staff in the Clerk’s Office, Circuit Executive’s Office, and Office of Staff Counsel; and

• Researching and communicating with the Federal Judicial Center regarding federal rulemaking guidelines.

I prepared a first draft of revised local rules strictly adhering to FRAP 47(a)(1) and the Guidelines’ style recommendations, with due attention to the Judicial Conference Local Rules Project critique of the circuit’s rules. In somewhat abridged form, the governing parameters for revision were:

• All generally applicable directions to litigants must be in a local rule.

• A local rule must not deal with internal administrative matters, but the court may elect to publish such information in Internal Operating Procedures (IOPs) or practitioner manuals.

• A local rule must be consistent with national rules and federal law.

• A local rule must not duplicate information already provided in national rules and federal law.

139 See supra note 42 and accompanying text.
A local rule must be numbered to correlate to the FRAP rule that covers the same topic.

A local rule must adhere to the style Guidelines and aim for clarity, consistency, readability, and brevity.

In addition, the first revision updated the substance of all local rules to reflect current actual practice and to accommodate the transition to electronic case filing.

The first revision also recommended a new format, arrangement, and numbering system, including relocating all administrative minutiae to a newly created category of IOPs. Headings and subheadings were added to orient readers, with subparts of rules following parallel organization and syntax. Extensive commentary accompanied the revision during the multiple rounds of vetting to explain all revisions and recommendations. The staff working group refined and augmented the revision, making significant changes to Second Circuit forms and other addenda to the local rules.

Once the staff working group had prepared a complete set of revised local rules and IOPs, it distributed the revision and explanatory commentary to the Attorney Advisory Committee to the Second Circuit Rules Committee. The members of the Advisory Committee organized themselves into subcommittees to address specific sections of the revision and then, in a series of meetings, shared their reactions and recommendations with the entire Advisory Committee and the staff working group. The work product of the subcommittees was consolidated into a comprehensive Advisory Committee report including proposed edits, analysis, and commentary.

The staff working group reconvened to evaluate and determine

---

140 IOPs correlating to FRAP were to be numbered accordingly, and those with no FRAP correlative were to be assigned a letter and appended at the end of the local rules. See 2D CIR. LR (2010).

141 The members of the Attorney Advisory Committee at the time were: Bruce R. Bryan, Daniel J. Capra, Ernest Collazo, Ira Feinberg, Michele Hirshman, Celeste Koeleveeld, Hon. Gerald E. Lynch, Karen McAndrew, William J. Nardini, Varuni Nelson, and Edward Zas.
whether to incorporate the Advisory Committee’s proposals into the revision. The staff working group then sent the court’s Rules Committee a new version of revised local rules and commentary, assimilating many of the changes suggested by the Advisory Committee and explaining why certain proposals were not incorporated. The Rules Committee also received a copy of the Advisory Committee’s final report. Assisted by the staff working group, the Rules Committee further revised the rules. The Rules Committee distributed its revision, along with the staff and Advisory Committee commentary, to the full court, which voted to publish the revised rules for public notice and comment. Public comments were few, and additional, non-substantive changes were made to the revision. The full court adopted the new rules effective January 1, 2010. The court, however, adopted only the actual rules and decided not to publish any of the commentary.

Over the course of 2010, the court and the appellate bar became familiar with the new rules in action. Their experience suggested the need for additional amendments—both technical and substantive. After a deliberative process echoing (albeit abbreviating) the earlier local rules revision project, the court published further revised rules, effective December 15, 2010, largely to fill gaps and clarify ambiguities.

II. COMMENTARY ON THE NEW SECOND CIRCUIT LOCAL RULES

This part presents commentary that explains how the Second Circuit local rules revision methodology was applied to alter and craft particular rules. The commentary focuses on the most critical reform objectives: (1) every generally applicable practice directive must be in a local rule and local rules are limited to that purpose; (2) local rules must be consistent with and not duplicate national law; and (3) the style and structure of local rules must be clear, consistent, and readable. This part discusses the rules in categories that correspond to how an appellate practitioner might interact with them chronologically during the course of pursuing an appeal.
A. Scope and Organization

The revised rules introduce LR 1.1 Scope and Organization to explain the new system of denoting and locating LRs and IOPs. In the revision, LRs and IOPs are numbered and titled to correspond to FRAP. If there is no FRAP counterpart, an LR is numbered to correspond to FRAP 47, and an IOP is assigned a letter and placed at the end of the rules. The rule also directs litigants to the court’s website for additional instructions and practice manuals. Arguably a redundant rule in itself, LR 1.1 was deemed a necessary prologue to a wholesale revision of the rules.

Local Rule 6.1 clarifies that the local rules and IOPs applicable to civil appeals are also applicable in bankruptcy cases.

B. Docketing the Appeal and Preparing the Record

Local Rules 12.1 Appeal Docketing Requirements in Civil and Agency Cases and 12.2 Appeal Docketing Requirements in Criminal Cases were introduced to provide notice of the obligation to file certain forms at the outset of the appeal and pay the docketing fee. The rules set firm deadlines, and warn parties that noncompliance may result in dismissal of the appeal. Previously, these instructions were publicized only in collections of guidelines and practice tips available in the clerk’s office or on the court’s website, in violation of FRAP 47(a)’s requirement that such generally applicable directions be in a local rule.

---

142 See 2D Cir. LR 1.1 (2010).
143 Only one local rule was without a FRAP counterpart—the rule on death penalty cases, now designated 2D Cir. LR 47.1. Other circuits similarly locate local rules not corresponding to FRAP after FRAP 47, and locate non-correlative IOPs at the beginning or end of their local rules. See, e.g., 5TH Cir. LR 47 (2009) (“Other Fifth Circuit Rules”); 8TH Cir. IOP (2007) (attached as an appendix at end of local rules).
144 See 2D Cir. LR 12.1, 12.2 (2010).
145 See 2D Cir. LR 12.1(a), (d), 12.2 (2010).
146 See, e.g., How to Appeal Your Civil Case and Civil Appeals Management Plan, in Appeals to the Second Circuit, supra note 5, at S170–85, S250–55.
Local Rule 12.3 Acknowledgement and Notice of Appearance in All Appeals clarifies and consolidates the court’s former requirement that a party submit: (1) a form acknowledging the docketing of the appeal and (2) a notice of appearance of record counsel or individual appearing pro se. This two now-supplanted forms significantly overlapped in the information they requested. Consolidating them reduces the paperwork and administrative burdens on the parties and the court. The new joint form must be filed at the outset of the case to accelerate the court’s access to important information about the appeal, including any necessary corrections to the caption and appellate designations, information about related cases, and the identity of and contact information for counsel of record. The form also satisfies the FRAP 12(b) requirement to file a representation statement. Most saliently, the consolidated form no longer seeks information about oral argument preferences and availability. The new rules defer those inquiries to a later point in the case, after the filing of the final appellee brief, when the parties are better situated to provide accurate and reliable responses.

Local Rule 12.3 also clarifies that all counsel of record must be admitted to practice in the court from the outset of the appeal; the deadline for filing the Acknowledgment and Notice of Appearance form provides time for counsel to apply for regular or pro hac vice admission as necessary. In addition, all counsel appearing in a case in any capacity must file a Notice of Appearance form at the time they enter the case. The admission requirement represents a

147 See 2D Cir. LR 12.3 (2010).
148 See Acknowledgement and Notice of Appearance form, U.S. COURT OF APPEALS FOR THE SECOND CIRCUIT (DEC. 22, 2010), http://www.ca2.uscourts.gov/clerk/Forms_and_instructions/forms_home.htm (follow “Forms” hyperlink; then scroll down to “Attorneys” and the link for “Acknowledgement and Notice of Appearance” download).
149 See FRAP 12(b) (requiring the attorney who filed the notice of appeal to “file a statement with the circuit clerk naming the parties that the attorney represents on appeal”).
150 See 2D Cir. LR 34.1(a) (2010); see also discussion infra Part II.I.
151 See 2D Cir. LR 12.3(a) (2010).
152 See 2D Cir. LR 12.3(b) (2010).
significant change from prior procedures that did not ask for a Notice of Appearance until counsel’s first brief, that did not attend to the admission status of counsel presenting oral argument until that day, and that generally neglected the admission status of counsel appearing in other capacities.

In practice, lax enforcement of admission requirements enabled non-admitted counsel of record to appear in a case for all purposes prior to filing the brief, and conceivably through resolution if the parties waived oral argument. Counsel serving in other capacities could easily evade admission requirements for the duration of the appeal. This system had several weaknesses, including the possibility that an appeal could be dismissed or otherwise resolved when a party is represented by a non-admitted attorney, that the court could decide the case on the basis of briefs submitted by a non-admitted attorney, or that it would be necessary to impose discipline on a non-admitted attorney. The urgency of reforming these procedures was sufficiently great that the new LR 12.3 and related amendments to the attorney admission rule at the time were implemented nine months ahead of the effective date of the other rules revisions.153

The court clarified its procedures for forwarding the record on appeal in new LR 11.1 Duties Regarding the Record. The new rule codifies the existing Second Circuit practice of requiring the district clerk to retain the record on appeal in all counseled appeals.154 Memorializing this practice finally puts the court in compliance with FRAP 11(e)(1) which requires each circuit court to announce in a local rule if its default practice is “that a certified copy of the docket entries be forwarded instead of the entire record.”155 The appellant’s duty in connection with this step of the appeal is to do “whatever is necessary” in connection with forwarding the docket entries.156

Local Rule 11.1 also lists two categories of cases where the

154 See 2D CIR. LR 11.1(a) (2010).
155 See FRAP 11(e)(1).
156 See 2D CIR. LR 11.1(a) (2010) (tracking the language of FRAP 11(a)).
The New Second Circuit Local Rules

district court still has to forward the record. First, in pro se cases, the court routinely needs ready access to the record on appeal as parties in those cases typically stint on providing appendices of record excerpts. Second, the court has decided, pursuant to FRAP 30(f), to authorize certain classes of appellants to proceed on the original record without appendices, in which case the procedure for the district court to retain the record on appeal does not apply.157

Local Rule 11.2 Exhibits Retained by the Parties replaces a dense and duplicative prior local rule dealing with the handling of exhibits on appeal. The old rule’s defects included: repeating parts of FRAP 11 and 30, quoting language from a superseded version of FRAP 11(a), employing a convoluted process for designating and transmitting retained exhibits to the circuit clerk, and imposing obligations on the parties that were superfluous after the advent of electronic filing and contrary to the circuit’s now-codified practice of having the district court retain the record on appeal.158 The new rule eliminates the requirement that parties forward retained exhibits to the circuit clerk with the record on appeal.159 Local Rule 11.2 also eliminates the requirement that parties deposit retained exhibits with the district clerk once a notice of appeal is filed, sparing the district clerk the burden of cataloguing, storing, and transmitting exhibits that the parties prefer to retain and that the circuit court may never ask to see.160

Local Rule 11.3 Duty of Court Reporters effects an even more significant change, placing squarely on court reporters the duty of timely transcript preparation and penalizing them for late delivery.161 This new rule brings the Second Circuit into compliance with FRAP 11(b) and Judicial Conference resolutions regarding late delivery of transcripts. In 1979, FRAP 11(a) was amended to circumscribe the appellant’s duties with respect to forwarding the record, recognizing that “[a]side from ordering the

157 See 2D Cir. LR 11.1(b) (2010); 2D Cir. LR 30.1(e) (2010) (authorizing appeals on the original record in in forma pauperis proceedings, social security cases, and immigration cases).
158 See 2D Cir. LR 11 (2008); see also FRAP 11(a) (1979) & 1979 amend.
159 See 2D Cir. LR 11.2 (2010); 2D Cir. LR 11.1 (2010).
160 Compare 2D Cir. LR 11.2 (2010), with 2D Cir. LR 11(c) (2008).
161 See 2D Cir. LR 11.3 (2010).
transcript within the time prescribed the appellant has no control over the time at which the record is transmitted, since all steps beyond this point are in the hands of the reporter and the clerk." 162

Because preparing and delivering the transcript of proceedings is entirely within the court reporter’s power, it was nonsensical and unfair to ask the appellant to do more than place a timely order for the transcript. Amended FRAP 11(b)(1) therefore places the burden on the court reporter to notify the circuit clerk of receipt of the transcript order, to request any necessary extensions of time, and to risk the wrath of the district judge for delays that the circuit clerk now must report. 163 In 1982, the Judicial Conference added another layer of incentives to improve transcript delivery times when it adopted a resolution authorizing fee reductions for late delivery of transcripts. 164

Despite these developments, the Second Circuit continued to require the parties to monitor transcript readiness and move for any extensions of time when a transcript was delayed. Local Rule 11.3

162 FRAP 11(a) (1979) & 1979 amend.
163 FRAP 11(b)(1)(A), (B), (D).
164 See REPORT OF THE PROCEEDINGS OF THE JUDICIAL CONFERENCE OF THE UNITED STATES 10 (Mar. 11–12, 1982). The resolution provides:

That for [a] transcript of a case on appeal not delivered within 30 days of the date ordered and payment received therefor, or within such other time as may be prescribed by the circuit council, the reporter may charge only 90 percent of the prescribed fee; that for a transcript not delivered within 60 days of the date ordered and, payment received therefor, or within such other time as may be prescribed by the circuit council, the reporter may charge only 80 percent of the prescribed fee. No fee may be charged which would be higher than the fee corresponding to the actual delivery time. In the case of a transcript which is subject to F.R.A.P. Rule 11(b), the reduction in the fee may be waived by the clerk of the court of appeals for good cause shown. Nothing contained herein should be construed as sanctioning untimely delivery, nor should this provision be considered the only penalty that could be imposed by the court or circuit council on habitual offenders.

Id. The resolution was reaffirmed in September 1990 and is still in effect. See REPORT OF THE PROCEEDINGS OF THE JUDICIAL CONFERENCE OF THE UNITED STATES (Sept. 12, 1990); see also Reporters, 28 U.S.C.A. § 753(f) (West 2010) (subjecting to the approval of the Judicial Conference the rates charged by court reporters for transcripts).
The New Second Circuit Local Rules

cures this nonconformity in Second Circuit practice. First, it requires the court reporter to estimate a completion date no later than thirty days after receipt of the transcript order form.\textsuperscript{165} Second, it requires the court reporter to request an extension if more time is necessary.\textsuperscript{166} Third, a court reporter is obligated to update the circuit clerk in fourteen-day intervals until a late transcript is filed, and must charge reduced fees unless the court has excused the delay.\textsuperscript{167} Although aspects of LR 11.3 repeat provisions of FRAP,\textsuperscript{168} because the new local rule radically departs from existing practice, repetition was deemed necessary to ensure compliance.

New rule LR 4.2 explains an appellant’s duties when a motion has been filed in the district court that extends the time to file a notice of appeal.\textsuperscript{169} The appellant must notify the court upon the filing of such a motion, and again upon its disposition.

\textbf{C. Electronic Case Filing}

New rule LR 25.1 Case Management/Electronic Case Filing (CM/ECF) addresses the special issues that arise when court documents are filed and maintained in electronic form.\textsuperscript{170} Local

\textsuperscript{165} See 2D Cir. LR 11.3(a) (2010).
\textsuperscript{166} See 2D Cir. LR 11.3(b) (2010).
\textsuperscript{167} See 2D Cir. LR 11.3 (c), (d) (2010). In adopting LR 11.3, the Second Circuit joined seven other circuits that have acknowledged and elaborated on FRAP 11’s regulation of court reporters. See 4TH Cir. LR 11(a), (b), IOP 11.1; 5TH Cir. LR 11.1; 6TH Cir. LR 11(b), IOP 11(c); 7TH Cir. LR 11(c); 9TH Cir. L.R. 11.1–.3; 10TH Cir. LR 10.1(C) & App. B; 11TH Cir. LR 11-1, IOP 11-1.
\textsuperscript{168} Compare 2D Cir. LR 11.3(a) (2010) with FRAP 11(b)(1)(A) (court reporter to state expected completion date); compare 2D Cir. LR 11.3(b)(1) (2010) with FRAP 11(b)(1)(B) (court reporter duty to request extension of time); compare 2D Cir. LR 11.3(c) (2010) with FRAP 11(b)(1)(D) (circuit clerk to notify district judge if transcript is delayed).
\textsuperscript{169} See 2D Cir. LR 4.2 (2010); FRAP 4(a)(4), (b)(3), 6(b)(2).
\textsuperscript{170} See 2D Cir. LR 25.1 (2010). This rule borrows much from the Administrative Office of the U.S. Courts’ model local appellate rules for electronic filing.
Rule 25.1 makes electronic filing the norm in the Second Circuit, to increase efficiency, reduce costs, and maximize the anticipated benefits of enhanced public access to court documents. The rule nonetheless comports with FRAP 25’s requirement of reasonable exemption from electronic filing for particular documents and particular cases upon showings of good cause. All counsel admitted to practice in the court must register as a Filing User with PACER, and pro se parties may do so with permission. By registering, a Filing User consents to electronic service of documents. A Filing User’s manual signature is no longer required; the personal log-in and password constitutes a signature. Signally, the new rule prohibits the submission of paper copies for every document other than a brief, an appendix, and certain motions, writs and petitions for rehearing.

The Second Circuit had inched toward electronic filing in 2005 when it permitted submission of PDF (Portable Document Format) versions of briefs and appendices. The success of this pilot program led to the May 2008 adoption of a rule requiring submission of a PDF for every document filed except appendices. The requirement of submitting PDF appendices was added in January 2009. In the 2010 comprehensive local rules revision, all of these provisions on PDF submissions were collapsed into LR 25.2 Submission of PDF Documents. Local Rule 25.2 retains the PDF submission requirement for cases that predate the CM/ECF system and for cases exempt from electronic filing.

---

171 See 2D Cir. LR 25.1(a)(2) (2010).
172 See FRAP 25(a)(2)(D); 2D Cir. LR 25.1(j) (2010).
173 See 2D Cir. LR 25.1(b) (2010).
174 See 2D Cir. LR 25.1(h) (2010).
175 See 2D Cir. LR 25.1(f) (2010).
176 See 2D Cir. LR 25.1(g) (2010); see also 2D Cir. LR 21.1, 27.1(a)(4), 30.1(b), 31.1, 35.1(c), (remove “and”) 40.1(b) (2010).
177 See 2D Cir. Interim LR 32(a)(10), 25 (2005); 2D Cir. Interim LR (2005) (on file with author).
179 See 2D Cir. Interim LR 25.2 (2009).
180 See 2D Cir. LR 25.2 (2010).
The New Second Circuit Local Rules

As in LR 25.1, LR 25.2 largely dispenses with paper copies, requiring only the original document to be filed. However, LR 25.3 Additional Paper Copies authorizes the clerk to request additional paper copies for any document filed, whether electronically or otherwise.

D. Motion Practice

Another dramatic change was to the rule on motion practice, now styled LR 27.1 Motions. The prior rule dated in large part from 1972, and had not been revised to respond to material amendments to FRAP 27 in 1998, 2002, or 2005. For example, the prior local rule repeated FRAP 27’s content and format requirements for motions, but failed to reflect the 2005 FRAP amendment requiring that motion papers follow FRAP 32’s typeface and style requirements. The new local rule eliminates repetition of FRAP and expressly reinstates the authority of the national rule’s form requirements. The new local rule also eliminates the old rule’s four separate mentions of the unavailability of oral argument for motions, which were rendered superfluous with the 1998 addition of FRAP 27(e) directing that motions “will be decided without oral argument unless the court orders otherwise.”

Most significantly, the new LR 27.1 expressly disfavors the two most common motions made in the Second Circuit—to extend the time to file a brief and to file oversized briefs—and sets more

181 See 2d Cir. LR 25.2(d)(4) (2010).
182 See 2d Cir. LR 25.3 (2010).
183 See 2d Cir. LR 27.1 (2010).
185 See 2d Cir. LR 27(a) (2008).
186 See 2d Cir. LR 27.1(a)(1) (2010).
187 H.R. Doc. No. 105–269, at 52 (1998) (providing the amended language of FRAP 27(e)). The new rules relegate the only mention of oral argument on motions to IOP 27.1 Oral Argument on Motions, to advise the date normally appointed for argument in those limited cases when the court orders it. See 2d Cir. IOP 27.1 (2010).
rigid standards for granting such motions. Under the old regime, these two types of motions constituted a significant percentage of all motions filed in the circuit and imposed substantial burdens on court personnel. The revision follows the example of other circuits that similarly discourage these requests.\footnote{See, e.g., 7TH CIR. LR 26 (2010); 9TH CIR. LR 28-4 (2010); 10TH CIR. LR 27.4 (2010); D.C. CIR. LR 27(h)(3) (2010).}

With respect to motions for extensions of time, LR 27.1(f) works in conjunction with LR 31.2 Briefing Schedule; Failure to File, a new rule on briefing schedules that establishes longer but firm time periods for the preparation of briefs to reduce the incidence of extension motions.\footnote{See 2D CIR. LR 27.1(f), 31.2(d) (2010); discussion infra notes 209–17 and accompanying text (explaining LR 31.2).} Local Rule 27.1 provides that motions seeking to extend the time to file will not be granted “[a]bsent an extraordinary circumstance, such as serious personal illness or death in counsel’s immediate family.”\footnote{See 2D CIR. LR 27.1(f)(1) (2010).} This greatly contrasts with the previous practice of nominally adhering to FRAP’s briefing scheduling but granting virtually automatic sequential 30-day extensions on request. Furthermore, a party seeking an extension of time can no longer assume that making the motion tolls the brief’s deadline—the brief is due at the time originally set until the court orders otherwise.\footnote{See id.} A motion for an extension of time must be filed “as soon as practicable after the extraordinary circumstance arises,” thus short-circuiting any attempts to reconstruct distant events as an emergency in view of an imminent brief deadline.\footnote{See 2D CIR. LR 27.1(f)(3) (2010).}

Local Rule 27.1 also expressly disfavors motions to file an oversized brief, and significantly changes the procedures for such motions.\footnote{See 2D CIR. LR 27.1(e) (2010).} The new rule eliminates the prior requirement that a party submit page proofs with a motion to file an oversized brief, and instead requires a party to explain the reasons for exceeding FRAP’s size limitations.\footnote{See FRAP 32(a)(7); 2D CIR. LR 27.1(e)(2) (2010); 2D CIR. LR 27(g).}
The New Second Circuit Local Rules

particularly inefficient, because it required parties to wait before filing a motion for an oversized brief until the brief was essentially completed. Consequently, counsel had to complete a brief of the desired length without knowing whether it would be accepted. This policy often required the party to write its brief twice, once to submit the page proofs and again to comply with the court’s ruling if the request was not granted in full. The new rule recognizes that a party’s need to file an oversized brief will be obvious long before page proofs are ready—for example, a multi-defendant criminal appeal from a lengthy trial, or a complicated civil or regulatory dispute. Requiring an adequate explanation for the brief’s additional length, rather than page proofs, encourages parties to resolve the size issue at the earliest opportunity, especially given the risks of waiting too long and then being pressed for time to reduce the brief size if the motion is denied. In any event, a party must move to file a motion for an oversized brief no later than fourteen days before the brief is due, and untimely motions will be evaluated under the same “extraordinary circumstances” standard as motions for an extension of time.195

The new affirmative obligation imposed on parties to notify their opponents when filing a motion has the potential to streamline motion practice.196 Previously, the local rule was silent on this subject, and the court’s motion form asked only if the movant had sought the other parties’ consent to the motion.197 Prior practice did not mandate either that the movant seek consent or even alert the other parties that a motion would be made. Now the movant must communicate with the other parties about the motion or state why the movant was unable to do so. In this communication the movant must inquire as to opposing counsel’s position on the motion and whether opposition papers will be filed, and then report this information to the court.198 This process is

196 See 2d Cir. LR 27.1(b) (2010).
197 See 2d Cir. LR 27 Motion Information Statement (2008), United States Court of Appeals for the Second Circuit, in FEDERAL COURT OF APPEALS MANUAL: LOCAL RULES 381 (West 2008).
198 See 2d Cir. LR 27.1(b) (2010). Three other circuits have similar local
intended to expedite the handling of motions, especially procedural motions that may be disposed of without awaiting a response from the non-moving parties.199

Three other new provisions were added to the motions rule. The first authorizes the Second Circuit clerk to decide routine, unopposed motions, thereby complying with FRAP 27’s requirement that if an appellate court delegates authority to its clerk to decide motions, it must do so by rule.200 The second provision lists mandatory procedures for filing emergency motions.201 These procedures provide for early notification to the clerk’s office and plain labeling and explanation of the emergency, in part to avoid situations where a grant of *ex parte* relief would be inconsistent with normal principles of due process and notice. A third new provision sets a 14-day time limit for seeking reconsideration of a decision on a procedural motion, a timeframe consistent with FRAP 40’s deadline for a motion for panel rehearing.202 This provision also repeats FRAP 27’s admonition rules that require the movant to contact opposing counsel and report whether opposition papers will be filed. See 4TH CIR. LR 27(a); 5TH CIR. LR 27.4; 10TH CIR. LR 27.3(c).

199 See FRAP 27(b) (“The court may act on a motion for a procedural order . . . at any time without awaiting a response.”).

200 See 2D CIR. LR 27.1(c) (2010). Except for the Sixth Circuit, each circuit has adopted a similar rule, but they have taken different approaches in explaining the clerk’s authority. The Second Circuit’s LR 27.1(c) follows the broad, categorical approach that six other circuits have employed. See 1ST CIR. LR 27.0(d); 3D CIR. LR 27.6; 4TH CIR. LR 27(b); 9TH CIR. LR 27-7; D.C. CIR. LR 27(e); FED. CIR. LR 27(h). Four circuits itemize the specific motions that the clerk may decide. See 5TH CIR. LR 27.1; 7TH CIR. IOP 1(c)(2); 10TH CIR. LR 27.3; 11TH CIR. LR 27-1(c). One circuit uses a blended approach, describing a category of motions the clerk may decide and providing specific examples of those motions. See 8TH CIR. LR 27B.

201 See 2D CIR. LR 27.1(d) (2010). Seven circuits have similar local rules addressing emergency motions. See 1ST CIR. LR 27.0(b); 3D CIR. LR 27.7; 4TH CIR. LR 27(e); 5TH CIR. LR 27.3; 6TH CIR. LR 27(c), IOP 27(b); 9TH CIR. LR 27-3; 11TH CIR. LR 27-1(b).

202 See 2D CIR. LR 27.1(g) (2010); FRAP 40(a)(1). At least six other circuits have local rules establishing a similar timeframe. See 4TH CIR. LR 27(b) (14 days); 8TH CIR. LR 27B(d) (14 days); 9TH CIR. LR 27-10(a)(2) (14 days); 11TH CIR. LR 27-2 (21 days); D.C. CIR. LR 27(e)(2) (10 days); FED. CIR. LR
that response papers filed after the original motion was decided do not constitute a motion for reconsideration. This repetition was necessary to clarify that the new rule effectively rescinds the court’s former standing direction that allowed the clerk to treat response papers as a motion for reconsideration when timely filed but arriving after the original motion was decided.

Many deletions from the old motions rule are also notable. For example, the revision eliminates provisions detailing the timing and mechanics of how motions are heard and decided, as these provisions are largely obsolete and concern internal court procedures as opposed to matters of appellate practice. The streamlined approach preserves administrative flexibility.

Newly catalogued as a motions-related rule, LR 27.2 Certification of Questions of State Law was relocated from the section of the old rules titled “Rules Relating to the Organization of the Court.” The rule allows the court on its own or a party’s motion to certify a question of state law to a state’s highest court. The rule required only minor revision to clarify that the court technically does not issue a stay when it certifies a question of state law, but rather retains jurisdiction and holds in abeyance that much of the case that is dependent on the results of certification.

27(k) (14 days).

203 See FRAP 27(b); 2D Cir. LR 27.1(g) (2010).
204 See Compilation of Standing Directions to the Clerk of the Second Circuit Court of Appeals, ¶ 27, Aug. 12, 2002 (on file with author).
205 See, e.g., 2D Cir. LR 27(b), (c), (f) (2008). The new rule no longer itemizes and explicates: motions to be heard at regular sessions of court, motions to be heard by a panel which has rendered a decision, motions for leave to appeal, motions to be determined by a single judge, pro se motions, and miscellaneous motions.
206 See 2D Cir. LR 27(b)–(f), (b), (j) (2008).
207 See 2D Cir. LR 27.2 (2010); 2D Cir. LR § 0.27 (2008).
208 See 2D Cir. LR 27.2 (2010).
209 See, e.g., Tunick v. Safir, 209 F.3d 67, 89, 90 (2000) (noting that the court would retain jurisdiction over the case during the certification process).
E. Briefing Schedules and Requirements for Briefs and Appendices

Effecting a radical change in the procedure for establishing briefing schedules, LR 31.2 Briefing Schedule; Failure to File, sets brief deadlines according to the parties’ proposed dates, within an outer time limit, removing the justification for routine extensions of time.\textsuperscript{210} This innovation also removes the need for the court to issue and docket scheduling orders, which had imposed a significant administrative burden given that few appeals were briefed in accordance with their original timetables. Second Circuit practice had long contravened FRAP 31’s briefing timetable, which allows forty days to appellant, followed by thirty days to appellee, for filing briefs.\textsuperscript{211} The Second Circuit routinely extended these deadlines at a party’s request, with the result that by 2009, the typical appeal took nine months to brief.\textsuperscript{212} Because the Second Circuit has endured significant backlogs over the last decade largely due to the explosion of its immigration docket, protracted briefing schedules did not materially affect the court’s operations and calendar. To the contrary, timely briefs often had the drawback of being stale by the time the court heard the case.

Nonetheless, the practice of granting serial extensions of time to file a brief was problematic for many reasons, including that: (1) multiple motions for extensions for time within each case and across all cases imposed significant burdens on court personnel and resources, and increased the risk of docketing errors; (2) irregular and indeterminate briefing schedules made the appellate process less predictable and efficient; and (3) the routine availability of extensions of time made the appellate process more vulnerable to manipulation for purposes of adversarial advantage and delay. In response to these concerns, in January 2009, the court implemented a pilot program for criminal appeals, adopting procedures for establishing briefing schedules similar to those later

\textsuperscript{210} See 2d Cir. LR 31.2(a) (2010).
\textsuperscript{211} See FRAP 31(a).
\textsuperscript{212} Telephone Interview with Catherine O’Hagan Wolfe, Second Circuit Clerk of Court, U.S. Court of Appeals for the Second Circuit (July 14, 2010).
The New Second Circuit Local Rules 517

codified in LR 31.2. The success of the pilot program led to the adoption of the new local rule extending the “pick-your-own-deadline” scheduling approach to all appeals.213

Local Rule 31.2(a) now requires the parties, rather than the court, to set the schedule through a notification procedure keyed to certain events. For appellant, the event in most cases is the delivery date of the last transcript, known as the “ready date.”214 The appellant must notify the clerk of its proposed brief deadline, which must be within ninety-one days of the ready date. Upon filing of the last appellant’s brief, the appellee must notify the clerk of its proposed brief deadline, which must be within ninety-one days after that filing.215 Later deadlines are available “only if the case involves a voluminous record or extreme hardship would result.”216 If a party fails to submit the required notification, it must abide by FRAP’s default—and much shorter—briefing deadlines.217 Reply briefs must be filed roughly in the same time that FRAP requires—fourteen days after the last appellee’s brief.218 Thus, in cases where the parties set the briefing schedule at the outermost acceptable limits, the typical appeal will take six-and-a-half months to brief.

The new brief scheduling rule also addresses motions regarding briefing,219 in a manner arguably redundant of the new LR 27.1.220 This redundancy was deemed necessary, however, to signal and reaffirm to parties that the new procedures do not countenance routine extensions of briefing deadlines. Similarly made plain is

213 See Notice to the Criminal Law Bar, Jan. 14, 2009 (on file with author).
214 See 2d Cir. LR 31.2(a)(1)(A) (2010).
215 See 2d Cir. LR 31.2(a)(1)(B) (2010). The version of the new local rules, effective on January 1, 2010, originally set the outer time limit for appellant’s and appellee’s briefs at 120 days, but concomitant efforts to reduce the court’s backlog were sufficiently successful to require reducing the outer limit to 91 days, effective as of the December 15, 2010 rules amendments. It is conceivable that future backlog reductions will result in recalibrating the outer limits of briefing deadlines to further shorten the duration of the average appeal.
218 See 2d Cir. LR 31.2(a)(2) (2010).
219 See 2d Cir. LR 31.2(c) (2010).
220 See supra notes 187–94 and accompanying text.
the court’s *sua sponte* authority to dismiss an appeal in the case of default under the rule.221

Local Rule 31.2 also establishes a new Expedited Appeals Calendar (XAC) to expedite handling of appeals from threshold dismissals of a complaint.222 The clerk’s office will automatically place on the XAC all appeals from a judgment or order of a district court dismissing a complaint solely for lack of subject matter jurisdiction under Fed. R. Civ. P. 12(b)(1), failure to state a claim upon which relief can be granted under Fed. R. Civ. P. 12(b)(6), or filing a frivolous complaint or for any other ground specified in 28 U.S.C. § 1915(e)(2).223 Briefing schedules are abbreviated to thirty days per side for an initial brief, and fourteen days for a reply brief. The new calendar appears to be a reaction to the U.S. Supreme Court’s recent tightening of pleading standards in federal civil cases.224 The Second Circuit seeks to return promptly to the district courts cases where it disagrees with the dismissal ruling, to get those cases back on track with as little disruption to the flow of the litigation.

Other requirements for briefs are materially unchanged in the new rules. Local Rule 31.1 Number of Copies of Brief to be Filed with Clerk reduces the number from ten to six to reflect changing needs in connection with electronic case filing.225 The first paragraph of LR 28.1 Briefs continues to warn parties to be concise and logical, but does so in a more succinct and less repetitive fashion.226 The Appellate Local Rules Report had

221 2D CIR. LR 31.2(d) (2010).
222 See 2D CIR. LR 31.2(b) (2010).
223 See 2D CIR. LR 31.2(b)(1) (2010).
226 See 2D CIR. LR 28.1(a) (2010); 2D CIR. LR 28(a) (2008). Consideration was given to rescinding this provision because case research unearthed only two instances in the thirty-one-year history of LR 28’s first paragraph when the court invoked the provision, suggesting that the provision may be unnecessary. See Jian Chen v. Gonzales, 216 F. App’x 121, 122 (2d Cir. 2007) (warning counsel of possible disciplinary proceedings for continuing to submit briefs that “fall[] far below the standards identified in Federal Rule of Appellate Procedure 28 and Local Rule 28”); Singh v. Gonzales, 211 F. App’x 33, 34 (2d Cir. 2007)
The New Second Circuit Local Rules

criticized the second paragraph of the old LR 28 as inconsistent with FRAP 28 because it imposed the additional content requirement of a “preliminary statement” naming the judge or agency member who rendered the decision below and a citation to that decision.227 The new local rule nonetheless retains the requirement because of the utility of the information it seeks, contrasted with the minor imposition on litigants. To ameliorate the inconsistency with FRAP 28, the revision eliminates the need for a separate “preliminary statement,” which has no FRAP 28 counterpart, and locates the additional content in the brief’s “statement of the case” that the national rule already requires.228

The rule for citing summary orders in briefs and other court filings, LR 32.1.1 Disposition by Summary Order, was reorganized and streamlined.229 The revised rule trims and relocates to an IOP the description of the provenance of the court’s practice of issuing non-precedential summary orders, now that the practice is well-established.230 The summary order legend was also relocated to an IOP, and shortened to track the rule’s simplified language.231 The revision also clarifies that the rule regulates only the citation of Second Circuit summary orders in that court, and not the citation of summary or unpublished dispositions of other courts, or the

(concluding that “meaningful appellate review [was] impossible” because of petitioner’s failure to follow the requirements of LR 28). Because the provision may nonetheless have some deterrent value, it was retained with only stylistic changes.

227 See FRAP 28(a); 2D Cir. LR 28(2) (2008); Appellate Local Rules Report, supra note 3, at 47 (criticizing former 2D Cir. LR 28 as “inconsistent” because “[i]f the [Judicial Conference] Advisory Committee had intended that additional areas be discussed in the briefs, it could easily have amended Appellate Rule 28 to include these items. Further, such variations among the courts may unduly confuse practitioners.”). This criticism was repeated in a 2004 Judicial Conference Report. See Marie Leary, Fed. Judicial Ctr., Analysis of Briefing Requirements on the United States Courts of Appeals: Report to the Judicial Conference Advisory Committee on Appellate Rules 4–14 (Oct. 2004).

228 See 2D Cir. LR 28.1(b) (2010); FRAP 28(a)(6).

229 See 2D Cir. LR 32.1.1 (2010); 2D Cir. LR § 0.23 (2008).

230 See 2D Cir. IOP 32.1.1(a) (2010).

231 See 2D Cir. IOP 32.1.1(b) (2010); 2D Cir. LR § 0.23 (2008).
citation of Second Circuit summary orders in other courts. As a practical matter, the Second Circuit cannot prevent how and under what circumstances its rulings, whatever the form, are cited in other courts. There is no mechanism to police this practice and no remedy available to the court when it happens. Another change of particular interest to counsel litigating against a pro se party is that citation of a summary order automatically triggers the obligation to serve a paper copy on the pro se party. Formerly, a paper copy was required only if the summary order was not publicly accessible on line. At the suggestion of the Attorney Advisory Committee, the new rule recognizes that reliance on an electronic database version is unfair given that many pro se parties may not have ready access to computers or the Internet, or sufficient computer skills to readily locate a summary order.

Requirements for appendices underwent even more revision than those for briefs. Local Rule 30.1 Appendix was restructured for clarity and readability by reordering the items to appear in the same order as FRAP 30. To deter a party from inappropriately including exhibits and other items that are not part of the record on appeal, the rule now expressly limits the contents of the appendix to materials listed in FRAP 30(a)(1) plus the notice of appeal. As with the number of copies of briefs, in light of electronic filing the number of copies of the appendix to be filed with the clerk is reduced from ten to three. Procedures for relying on the original record without an appendix are greatly simplified, also in light of electronic filing. Most significantly, the new version of this local rule adds a provision to comply with FRAP 30’s mandate that

---

232 Compare 2D Cir. LR 32.1.1(b) (2010), with 2D Cir. LR § 0.23 (2008).
233 Compare 2D Cir. LR 32.1.1(d) (2010), with 2D Cir. LR § 0.23(c)(1)(B) (2008).
234 See generally Lebron v. Sanders, 557 F.3d 76 (2d Cir. 2009) (per curiam) (recommending reconsideration of local rules that rely on pro se party access to electronic databases when opposing briefs cite unpublished opinions).
235 See 2D Cir. LR 30.1 (2010); 2D Cir. LR 30 (2008).
236 See 2D Cir. LR 30.1(a) (2010).
237 See 2D Cir. LR 30.1(b) (2010); 2D Cir. LR 31(b) (2008).
238 See 2D Cir. LR 30.1(e) (2010).
239 See 2D Cir. LR 30.1(f) (2010).
The New Second Circuit Local Rules

“[e]ach circuit must, by local rule, provide for sanctions against attorneys who unreasonably and vexatiously increase litigation costs by including unnecessary material in the appendix.” Since this mandate was added to FRAP in 1986, nine circuits complied by adopting a local rule providing for sanctions. The local rules revision brings the Second Circuit into compliance with the 1986 FRAP amendment.

Local Rule 32.1 Form of Brief and Appendix was modified to elaborate on pagination rules that will make it easier for court personnel to refer to a PDF version of these documents. Dividing an appendix into separate volumes is now expressly required when an appendix exceeds three hundred pages. Requirements for a special appendix were simplified to make it clear that one is required only when the appendix exceeds three hundred pages.

The court added a new provision to LR 29.1 Brief of an Amicus Curiae, acknowledging the FRAP 29 amendment that requires disclosure of party interests in amicus curiae briefs. Such disclosure will assist judges in assessing those submissions, and deter counsel from using an amicus brief to circumvent page limits. The Second Circuit twist on the amendment is that the disclosure must appear in the first footnote on the first page of the amicus brief.

F. Criminal Appeals

Local Rule 4.1 Continuation of Counsel in Criminal Appeals represents a substantial revision of the previous rule addressing the

---

240 FRAP 30(b)(2).
241 See 1ST CIR. LR 30.0(f); 3D CIR. LR 30.5; 4TH CIR. LR 30(a); 5TH CIR. LR 30.1.8, 32.5; 6TH CIR. LR 30(o); 8TH CIR. LR 30A(c); 9TH CIR. LR 30-2; 10TH CIR. LR 46.6(a); D.C. CIR. LR 30(b).
242 See 2D CIR. LR 32.1(a)(3) and (b)(3) (2010); 2D CIR. LR 32(b) (2008); see also H.R. DOC. NO. 99–179, at 16 (1986).
243 See 2D CIR. LR 32.1(b)(2) (2010).
244 See 2D CIR. LR 32.1(c) (2010).
245 See 2D CIR. LR 29.1(b) (2010); FRAP 29(c)(5).
246 See 2D CIR. LR 29.1(b) (2010).
duties of counsel in criminal cases. The revision removes language that may have suggested that counsel’s obligation to continue representing the defendant on appeal extends only to conviction after a trial. The revised rule clarifies that this obligation applies in all circumstances, including appeals after conviction on a guilty plea. The new rule cleans up the instructions for motions to withdraw as counsel on appeal, including how such motions interact with electronic filing requirements. Provisions of the old rule that referred to a superseded plan to implement the Criminal Justice Act were deleted.

In addition, two new provisions individually address motions to withdraw: (1) on the ground that the appeal is frivolous and (2) on the ground that the court has rendered an adverse decision. The frivolous appeal provision of LR 4.1 codifies circuit practice in response to the Supreme Court’s *Anders* decision, which delimited counsel’s duty to prosecute a criminal appeal that counsel has determined to be devoid of merit. *Anders* mandates that counsel request permission to withdraw in such cases, accompanied by a brief explaining the possible issues that could be raised on appeal and why those issues are frivolous. The local rule now formally requires the submission of *Anders* briefs, and compliance with any other mandates of subsequent case law on this subject. The adverse decision provision of LR 4.1 addresses appointed counsel’s obligations regarding petitions for certiorari, which had previously been delineated only in the Second Circuit’s Plan to Implement the Criminal Justice Act of 1964. The local

---

247 See 2D CIR. LR 4.1(a) (2010); 2D CIR. LR 4(b) (2008).
249 See generally 2D CIR. LR 4.1(b), (c) (2010).
251 See 2D CIR. LR 4.1(b) (2010).
The New Second Circuit Local Rules

rule codifies the existing procedures for counsel to move to be relieved of the obligation to file a certiorari petition if counsel has reasonable grounds to believe that the petition would have no likelihood of success.\(^{253}\)

The local rule on death penalty cases was substantially revised and renumbered LR 47.1.\(^{254}\) The old rule repeated provisions of both FRAP and statutory law, often in a manner that could lead to confusion as to how to proceed. For example, the old provision dealing with the duration of an automatic stay of execution misstated FRAP 41’s mandate rules and purported to continue the stay of execution beyond the court’s mandate. Once the court issues its mandate, however, it has no jurisdiction to stay execution of the death sentence.\(^{255}\) As originally worded, LR 47.1 did not effect a stay of the mandate and therefore could not, despite proclaiming otherwise, effect a stay of execution through Supreme Court review. This could set a trap for litigants who might overlook the need to move for a stay of the mandate. The new rule also recognizes the distinction between direct review of a federally-imposed death sentence and collateral review of a death sentence (whether imposed in state or federal court). The former situation does not pose the same exigencies because no execution date is set until after completion of direct review, and Federal Rule of Criminal Procedure 38(a) already imposes a stay.\(^{256}\) Moreover, the need for advance notice and monitoring of proceedings is much more acute in the latter case, when an execution date has been set and a stay of execution is sought.

In addition, the new death penalty rule reorganizes the subparts to put contemplated events in chronological order, and eliminates provisions regarding original petitions and certificates of appealability which repeated FRAP 22. Provisions dealing with

\(^{253}\text{See 2d Cir. LR 4.1(c) (2010).}\)

\(^{254}\text{See 2d Cir. LR 47.1 (2010); 2d Cir. LR § 0.28 (2008). The old rule was renumbered to correspond to FRAP 47, the protocol for local rules that have no FRAP counterpart. See 2d Cir. LR 1.1 (2010).}\)

\(^{255}\text{See United States v. Rivera, 844 F.2d 916, 921 (2d Cir. 1988) ("Simply put, jurisdiction follows the mandate.").}\)

\(^{256}\text{See Fed. R. Crim. P. 38(a) (imposing an automatic stay of a death sentence pending appeal).}\)
preparation and transmittal of the record were similarly excised as repetitive of FRAP and obsolete in an era of electronic case filing. 257 The details of the clerk’s internal docketing procedures and of the composition of death penalty case pools and panels now reside in IOP 47.1.

G. Habeas Corpus

Revised LR 22.1 Certificate of Appealability now comports with federal standards on habeas petitions, and specifically the requirement that ordinarily the district court must rule first on the appealability of a petition it denied. 258 The Second Circuit’s former local rule on certificates of appealability (COA) introduced ambiguity on this point by suggesting that a habeas petitioner may bypass the district court and request a COA in the first instance from the circuit court. 259 Four other circuits acknowledge in their local rules the requirement that a district judge first issue or deny a COA, and they ordinarily will decline to consider a COA application without a district judge ruling. 260 The revised LR 22.1

257 See FRAP 10, 11.


259 See 2D CIR. LR 22(a) (2008) (instructing that “where an appeal has been taken but no [COA] has been issued by the district judge or by this court or a judge thereof, the appellant shall promptly move in this court for such a certificate”).

260 See 1ST CIR. LR 22.0(a) (“ordinarily neither the court nor a judge thereof will act on a request for a certificate of appealability if the district judge who refused the writ is available and has not ruled first”); 3D CIR. LR 22.2 (instructing clerk to enter a remand if the district court does not make a determination); 6TH CIR. LR 22(a) (allowing an application for a COA in circuit court only after it has been denied by the district court); 9TH CIR. LR 22-1(a) (“A motion for a certificate of appealability (“COA”) must first be considered by the district court.”). Seven circuits are silent on the subject. See 5TH CIR. LR 22; 7TH CIR. LR 22, 22.2; 8TH CIR. LR 22A, 22B; 10TH CIR. LR 22.1; 11TH CIR. LR 22-1(b); D.C. CIR. LR 22; FED. CIR. LR (showing the repealed LR 22).
also adds a time limit for seeking a COA—twenty-eight days after the later of the district court denial or the filing of the notice of appeal.\footnote{See 2D CIR. LR 22.1(a) (2010).}

Local Rule 22.2 Second or Successive Applications Under §§ 2254 and § 2255 is a new rule intended to facilitate processing of motions to authorize a second or successive habeas application. The rule requires applicants to complete a detailed form describing all prior applications and to attach copies of those applications and any resulting district court decision.\footnote{See 2D CIR. LR 22.2 (2010).}

\section*{H. Civil Appeals Management Plan and Related Rules}

Local Rule 33.1 codifies the Second Circuit’s Civil Appeals Management Plan (CAMP).\footnote{See 2D CIR. LR 33.1 (2010).} CAMP dates back to a 1973 pilot program to mediate settlements of cases on appeal.\footnote{See Morris, supra note 107, at 171.} By 1980, CAMP had evolved into a routine pre-argument conference for civil appeals to explore the possibility of settlement and attempt to narrow the issues on appeal.\footnote{Civil Appeals Management Plan, 28 U.S.C.A., United States Courts of Appeals Rules, Second Circuit, at 487–93 (West 1980).} Despite the program’s status as a “generally applicable direction” to parties, it was never previously codified as a local rule. Its substance was distributed between two addenda to the local rules titled “Civil Appeals Management Plan” and “Guidelines for Conduct of Pre-Argument Conference under CAMP,”\footnote{Civil Appeals Management Plan and Guidelines for Conduct of Pre-Argument Conference under Camp, available at United States Court of Appeals} neither of which stated that it had the force of a local rule.

Fourth Circuit’s COA rule contains ambiguous language akin to the old Second Circuit rule in that it does not explicitly bar the circuit court from ruling on the COA when the district judge has not yet done so. See 4TH CIR. LR 22(a)(1) (permitting a petitioner to seek a COA in the circuit court “in cases in which the district court has not granted a certificate”).

\footnote{See 2D CIR. LR 22.1(a)(1) (2010).} Seven other circuits have adopted similar requirements of disclosure. See 1ST CIR. LR 22.1(a); 3D CIR. LR 22.5(a); 4TH CIR. LR 22(d); 6TH CIR. LR 22(b); 7TH CIR. LR 22(a); 8TH CIR. LR 22B; 9TH CIR. LR 22–4.
rule. Those two documents overlapped and duplicated information in FRAP and other local rules concerning civil appeals, including how to docket the appeal, forward the record, arrange for the transcript, apply for indigent status, and establish a briefing schedule. This imposed unnecessary burdens on counsel who needed to parse multiple sources of information about CAMP and the civil appellate process to identify any local variations from national requirements, and any CAMP commands that supplemented or varied from other local rules.

Local Rule 33.1 consolidates all information specific to appeal conference procedures and eliminates redundant material, reducing clutter and bringing the CAMP procedures into compliance with FRAP 47(a). The new rule has one new requirement—counsel must anonymously submit a post-conference survey, designed to provide information to the court about CAMP’s effectiveness.268

Another new rule clarifies the requirements for a procedural device often used in conjunction with CAMP—dismissal of an appeal without prejudice to later reinstatement. Local Rule 42.1 formally recognizes this expedient, which the court has used for many years to accommodate parties who wish to suspend activity on their appeal pending, for example, settlement negotiations or a decision in another case.269 The rule requires the parties to set an outside date for reinstating the appeal, modifying existing practice, which allowed stipulations that relied on indefinite dates. Indefinite dates were difficult for the clerk’s office to track and contributed to the years-long idling of some appeals.270 The requirement of a limiting date will facilitate tracking cases and

---

267 Prior to the 2010 local rules revision, the only statement that CAMP “has the force and effect of a local rule” were court decisions making that assertion. See, e.g., Adkins v. Gen. Motors Acceptance Corp., 562 F.3d 114, 117 (2d Cir. 2009) (quoting Lake Utopia Paper, Ltd. v. Connelly Containers, Inc., 608 F.2d 928, 929 (2d Cir. 1979)).

268 See 2d CIR. LR 33.1(c)(4) (2010).

269 See 2d CIR. LR 42.1 (2010).

270 Telephone Interview with Catherine O’Hagan Wolfe, Clerk of Court, U.S. Court of Appeals for the Second Circuit (July 14, 2010).
The New Second Circuit Local Rules

reduce the time that a case may languish.

I. Oral Argument

The court permanently adopted, with some refinement, new procedures for oral argument that had previously been implemented on an interim basis. Under LR 34.1 Oral Argument and Submission on Briefs, within fourteen days after the filing of the final appellee brief, each party is required to submit an Oral Argument Statement. Failure to do so will allow the court to infer that the party does not seek oral argument. And even if the parties request oral argument, the court reserves the prerogative to decline to hear argument and take a case on submission, as authorized under FRAP 34(a)(2). The revised rule eliminates, however, the provision that expressly invited parties, in cases where oral argument is deemed unnecessary, to file a statement of reasons to hear oral argument. FRAP does not require an opportunity to object to having a case heard on submission, and no other circuit court provides such an opportunity.

The Second Circuit’s Oral Argument Statement form also requests information about who will be arguing the case, and dates of the parties’ unavailability for argument. This replaces the inquiries no longer made in the revised Notice of Appearance form. The timing of filing the Oral Argument Statement form is keyed to the filing of the final appellee brief for two reasons. First, parties are better able to make an informed decision whether to

---

271 See 2d Cir. Interim LR 34 (2008).
272 See 2d Cir. LR 34.1(a) (2010). The interim rule had required the parties to file a Joint Oral Argument Statement that required them to consult as to the necessity of oral argument, but did not require the form to be filed until after the deadline for reply briefs. See 2d Cir. LR 34 (2008).
273 See 2d Cir. LR 34.1(b) (2010).
274 See 2d Cir. LR 34(d) (2008).
276 See supra note 148 and accompanying text.
forgo oral argument when briefing is close to complete. Second, the court needs to obtain the parties’ unavailability dates reasonably close to the likely argument date.

Once a case has been set down for oral argument, the court will grant requests for postponement only upon a showing of “extraordinary circumstances, and not by stipulation of the parties.”277 The same “extraordinary circumstances” standard appears in LR 27.1 on motions, where the rule provides as examples of such circumstances personal illness or a death in counsel’s immediate family.278 Local Rule 34.1 refines this definition to note that “[e]ngagement of counsel in another tribunal (other than the U.S. Supreme Court) is not an extraordinary circumstance.”279

The local rules revision also made permanent the previously interim rule establishing a non-argument calendar.280 Local Rule 34.2 Non-Argument Calendar makes explicit that oral argument is not available as a general rule in certain immigration cases.281 The court originally promulgated this rule in 2005 to address docket management issues arising out of the increasing and overwhelming number of asylum filings in the circuit.282 The local conditions that

277 2d Cir. LR 34.1(e) (2010). Every other circuit that has promulgated a local rule regarding postponement of oral argument requires some level of cause for granting relief. Five circuits require "good cause." See 3d Cir. LR 34.2; 4th Cir. LR 34(c); 5th Cir. LR 34.6; 6th Cir. LR 34(c); 11th Cir. LR 34-4(f). Three circuits require "extraordinary cause" or extraordinary circumstances." See 7th Cir. LR 34(e); 10th Cir. LR 34.1(A)(3); D.C. Cir. LR 34(g). See also 1st Cir. LR 34.1(d) ("grave cause"); 9th Cir. LR 34-2 ("good cause" except that a request made within fourteen days of argument date requires "exceptional circumstances").

278 2d Cir. LR 27.1(f)(1) (2010).

279 2d Cir. LR 34.1(e) (2010).

280 See 2d Cir. INTERIM LR § 0.29 (2008).

281 2d Cir. LR 34.2(a) (2010).

justified the rule have not abated, warranting its transition to permanent status. Besides some scrubbing of the rule’s language to excise jargon and eliminate administrative nonessentials, the biggest change to LR 34.2 is the declaration that the court may at some future time identify other classes of cases as appropriate for the Non-Argument Calendar.\footnote{See 2D CIR. LR 34.2(a)(2) (2010).}

\textit{J. Post-Disposition Matters}

Procedures for en banc and panel rehearing did not undergo any significant substantive change, although their status as interim rules was upgraded to permanent.\footnote{See 2D CIR. LR 35.1, 40.1 (2010); 2D CIR. INTERIM LR. 35, 40 (2008).} The language of LR 35.1 En Banc Procedure now clarifies that a simultaneous petition for both a panel rehearing and a rehearing en banc should be made in a single document.\footnote{2D CIR. LR 35.1(a) (2010).} A sanctions provision was added to LR 35.1 to parallel the one already included in LR 40.1 Panel Rehearing, which was revised substantially.\footnote{2D CIR. LR 35.1(d), 40.1(c) (2010).} In both rules, the sanctions provision no longer sets a ceiling on the amount that can be assessed against a party that files a frivolous petition for rehearing.\footnote{2D CIR. LR 35.1(d) (2010), and 2D CIR. LR 40.1(c) (2010), with 2D CIR. L.R. 40(c) (2008).} In addition, sanctions are no longer awarded in favor of the petitioner’s adversary.\footnote{See 2D CIR. INTERIM LR 40(c) (2008).} Because FRAP does not permit the adversary to respond to a petition unless ordered to do so,\footnote{FRAP 35(e), 40(a)(3) (2009).} it was anomalous to award sanctions for a meritless petition to an adversary who likely did not respond and did not incur related expense.

Matters of internal court administration concerning en banc polls and decisions were streamlined to abate duplication of statutory and national rule requirements, and relocated in an
adjoining IOP. One potentially contentious issue newly resolved in IOP 35.1 is when to determine eligibility to vote in an en banc poll. It is well-settled that only active judges may vote to determine whether a case should be heard or reheard en banc. However, given that there is often a time gap (sometimes substantial) between the date an en banc poll is requested and the date the en banc order is entered, it is necessary to specify when to determine active status. A judge may be active at the time of casting a vote in an en banc poll, but may have taken senior status by the time all votes are cast. The IOP determines the judge’s eligibility to have a vote counted as of “the date of entry of the en banc order.” This choice coincides with the most likely reading of the en banc statute, which allows for en banc hearings only when “ordered by a majority of the circuit judges of the circuit who are in regular active service.” Because the operative verb in the statute is “ordered,” the better reading of the statute is to determine a judge’s eligibility at the time an en banc hearing is ordered.

Even if the en banc statute were ambiguous, as a policy matter, the best option for determining a judge’s status for the purpose of counting votes in an en banc poll is the date of entry of the en banc order. Until an order is actually entered, judges can (and often do) change their vote. Majorities shift back and forth. Thus, measuring a judge’s eligibility at the latest possible date ensures that only active judges cast a vote. Furthermore, before the close of voting, a retiring judge could be replaced by a newly appointed judge who would be entitled to vote on the en banc poll. Because the circuit cannot have more active votes than it has authorized judgeships, the retired judge should give up the en banc vote to the new

290 See 2d Cir. IOP 35.1 (2010); see also 28 U.S.C. § 46(c) (1996) (establishing procedures for a rehearing en banc, including the basis for determining the majority necessary for ordering en banc consideration); FRAP 35(a) (2009) (same).
291 See 2d Cir. IOP 35.1(b) (2010).
293 See 2d Cir. IOP 35.1(b) (2010).
294 28 U.S.C. § 46(c) (emphasis added).
The New Second Circuit Local Rules

The instructions for applications under the Equal Access to Justice Act were renumbered LR 39.2, corresponding to FRAP 39 Costs. Half the old rule was purged as obsolete because of amendments to the underlying statutory authority. What remains deals only with the application’s format. The provision on the clerk’s entry of orders was renumbered LR 45.1 Clerk’s Authority to correspond to FRAP 45 Clerk’s Duties. The local rule’s lengthy and detailed enumeration of the clerk’s authority to sign orders was reduced to a single sentence, as the old rule no longer accurately reflected the court’s procedures and repeated multiple provisions of FRAP.

K. Attorney Regulation

In conjunction with the implementation of electronic filing, the court overhauled its system for storing and accessing attorney

295 Although no other circuit has adopted a rule addressing the issue of when to assess a judge’s status in determining whether to count his vote on an en banc poll, the Third, Fifth, and Ninth Circuits have introduced some ambiguity on the subject by allowing a judge who takes senior status after the en banc poll to continue to participate in the argument and the subsequent decision. See 3D CIR. IOP 9.6.4 (2010) (“Any judge participating in an en banc poll, hearing, or rehearing while in regular active service who subsequently takes senior status may elect to continue participating in the final resolution of the case.”); 5TH CIR. LR 35.6 (2009) (same); 9TH CIR. LR 35-1 (2009) advisory committee note 2 (“[A] judge who takes senior status during the pendency of an en banc case for which the judge has already been chosen as a member of the en banc court may continue to serve on that court until the case is finally disposed of.”). These three approaches arguably conflict with 28 U.S.C. § 46(c) by allowing a senior judge who was not a member of the panel, and did not hear the en banc case while in regular active service, to participate in the resolution of the case.

296 See 2D CIR. LR 39.2 (2010) (formerly 2D CIR. LR § 0.25 (2008)). The only change to LR 39.1 Reproduction Costs was to eliminate the clerk’s authority to increase copy rates from time to time to reflect prevailing rates, to clarify that the clerk does not have unilateral authority to do so and a rule amendment is required. 2D CIR. LR 39.1 (2010).

297 See 2D CIR. LR 45.1 (2010); 2D CIR. LR § 0.18 (2008) (repeating provisions of FRAP 27(b), 31, 36, 41, 42(b)).
admissions information and revised Local Rule 46.1 Attorney Admission to require periodic renewal of admission.298 The new system will improve the court’s ability to track and communicate with an increasingly mobile and national bar, and to vet the professional credentials of lawyers who appear before the court.299 Consistent with this goal, the rule also explicitly requires that attorneys keep the court apprised of changes in contact information.300

Another material change to the rule concerns pro hac vice admission. The old rule purported to allow pro hac vice admission only under “exceptional circumstances,”301 but counsel frequently sought admission at oral argument without making the necessary showing, and the court routinely granted such requests. This phenomenon was largely the consequence of the prior regime’s failure to inquire into counsel’s admission status until oral argument was imminent.302 As mentioned above, new LR 12.3 cures this lapse by requiring counsel of record to be admitted and to file a notice of appearance attesting to their admission status at the outset of the case.303 Non-admitted attorneys must immediately seek full admission or make a written motion for pro hac vice admission.304

Local Rule 46.2 Attorney Discipline radically reorganizes the previous rule, eliminating provisions that duplicate FRAP 46, repeat other local rules, or discuss purely internal administrative matters.305 The rule’s new architecture first discusses the

298 See 2D Cir. LR 46.1(a) (2010) (requiring renewal of admission every five years).

299 Two circuits similarly have implemented an attorney re-registration requirement, requiring renewal every five years. See 5TH Cir. LR 46.1; 11TH Cir. LR 46-2.

300 See 2D Cir. LR 46.1(b) (2010).

301 See 2D Cir. LR 46(d) (2008).

302 See 2D Cir. Notice of Appearance form (2008); see also supra text accompanying notes 151–53.

303 See 2D Cir. LR 12.3(a) (2010); see also source cited supra note 151.

304 See 2D Cir. LR 46.1(d) (2010) (Pro hac vice admission is available only to attorneys acting for indigent parties or able to demonstrate exceptional circumstances).

305 See 2D Cir. LR 46.2 (2010) (formerly 2D Cir. LR 46(f)–(h) (2008)).
provisions addressing the structure and procedures of the court’s grievance and disciplinary mechanisms.\textsuperscript{306} Second, the rule addresses specific disciplinary matters, such as those arising out of suspension and disbarment orders of other licensing authorities, and those arising out of criminal convictions.\textsuperscript{307}

In addition to discipline for attorney misconduct, the court may sanction parties for delay under LR 38.1 Sanctions for Delay.\textsuperscript{308} This rule was revised to expand the court’s authority beyond sanctions merely for late filing of a record, brief, or appendix.\textsuperscript{309} Under the new rule, sanctions are available for any action that unnecessarily delays an appeal, such as late filing of other documents, failure to file required forms and notices, or the filing of frivolous motions. The new rule also removes language specifying the types of sanctions that might be ordered. Use of the broad term “sanctions,” without qualification, in this provision and elsewhere throughout the rules, conveys that all the types of sanctions may be ordered, and avoids the implication that one sanctions provision is more or less punitive than another.\textsuperscript{310}

\textit{L. Local Rules Relating to the Organization of the Court}

The Second Circuit local rules revision dispenses with the category of rules formerly known as “Local Rules Relating to the Organization of the Court.” Most of the provisions listed in that section are redesignated as IOPs, with a few deleted as obsolete, and a few retained as local rules and renumbered to correspond to FRAP.\textsuperscript{311} The provisions redesignated as IOPs address basic descriptive information about the court such as its name, seal, term, and sessions.\textsuperscript{312} Also relocated to IOPs were administrative details about what constitutes a quorum, the location and hours of the

\textsuperscript{306} See 2D CIR. LR 46.2(a), (b) (2010).
\textsuperscript{307} See 2D CIR. LR 46.2(c), (d) (2010).
\textsuperscript{308} See 2D CIR. LR 38.1 (2010).
\textsuperscript{309} See 2D CIR. LR 38 (2008).
\textsuperscript{310} See 2D CIR. LR 27.1(h), 30.1(f), 33.1(g), 35.1(e), 40.1(d) (2010).
\textsuperscript{311} See 2D CIR. LR 27.2, 34.2, 39.2, 45.1, 47.1 (2010).
\textsuperscript{312} See 2D CIR. IOP A, B, C, D (2010).
clerk’s office, library access, court fees, and circuit judicial administration. The Appellate Local Rules Report had also criticized former Second Circuit LR § 0.17 as unnecessarily redundant of the Court of Appeals Miscellaneous Fee Schedule set forth in 28 U.S.C. § 1913. Redesignated as IOP G, this provision now simply refers parties to the statute without reiterating its fee schedule, eliminating the need to continually update the local rule to keep current with fee schedule revisions. However, despite the Report’s comment that local rules stating a court’s name are unnecessary, the Second Circuit retained its name provision.

The IOPs were also recalibrated to comply with the revision project’s parameters. For example, the new “quorum” provision reduces repetition of the federal quorum statute and of FRAP 27(b)’s provision permitting a single judge to issue procedural orders. The new “clerk” provision eliminates both outdated procedures for lending and preserving original papers, and confusing instructions for certifying the record to the Supreme Court (which is covered more comprehensively and authoritatively in the Supreme Court rules). The IOP on Circuit Judicial Administration was trimmed of excessive detail, providing greater flexibility in the administration of the circuit’s periodic judicial conference, the format of which can vary with budgetary and personnel constraints.

313 See 2D Cir. IOP E, F, G, H, I (2010).
314 APPELLATE LOCAL RULES REPORT, supra note 3, Appendix for the Court of Appeals for the Second Circuit 2.
315 2D Cir. IOP G (2010).
316 APPELLATE LOCAL RULES REPORT, supra note 3, at 1 (observing that rules setting forth a court’s name repeat 28 U.S.C. §§ 41, 43(a)).
317 See 2D Cir. IOP E (2010).
319 See FRAP 27(b).
320 See Sup. Ct. R. 12(7), 16(2); 2D Cir. IOP F (2010); 2D Cir. LR § 0.16 (2008).
321 See 2D Cir. IOP I (2010); 2D Cir. LR § 0.22 (2008). The revision is similar to the simplified approach taken by six other circuits. See 1ST Cir. LR 47.1; 3D Cir. L.A.R. MISCE. 105.0; 5TH CIR. OTHER IOPS; 8TH CIR. IOP L.B.3; 9TH CIR. IOP D; FED. CIR. LR 53.
The New Second Circuit Local Rules

M. Rescinded Rules

The revision eliminates a number of local rules superseded by FRAP amendments. The rescission of LR 3(d) comes thirty years late, as the rule was superseded by a FRAP amendment that took effect in 1979.\textsuperscript{322} Prior to 1979, FRAP 3(d) required the district clerk to forward to the circuit clerk a copy of the notice of appeal only in criminal cases and habeas corpus proceedings. Accordingly, Second Circuit LR 3(d) was designed to fill the gap and require the district clerk to forward a copy of the notice of appeal in all other civil cases as well. In 1979, FRAP 3(d) was amended to extend the forwarding requirement to all civil cases, thus eliminating the need for LR 3(d).

Also long overdue was the deletion of LR 9 listing five items necessary to an application for release of a criminal defendant pending appeal. This local rule had been redundant since the wholesale revision of FRAP 9 in 1994, which in combination with various national rules and statutes already requires a release application to include old LR 9’s first four items.\textsuperscript{323} As to LR 9’s fifth item—counsel’s certification that the appeal is not taken for delay—the Appellate Local Rules Report specifically criticized this requirement as both inconsistent with FRAP 9(b)’s instructions on the materials comprising a release application and duplicative of a virtually identical statutory requirement.\textsuperscript{324}

The now-expunged LR 15, dealing with party designations in National Labor Relations Board enforcement proceedings, was superseded in 1986 with the amendment of FRAP 15.1 to conform national practice to “the existing practice in most circuits.”\textsuperscript{325} Rescinded LR 21’s directives regarding extraordinary writs were similarly superseded when FRAP 21 was amended in 1996 to

\textsuperscript{322} See FRAP 3(d) & Note on 1979 Amendments.
\textsuperscript{323} See FRAP 9(b), 9(c), 27(a); FED. R. CRIM. P. 32(k), 46(c); 18 U.S.C. §§ 3143, 3145(c) (2006); 2D CIR. LR 9(1)-(4) (2008).
\textsuperscript{324} See 18 U.S.C. § 3143(b)(1)(B) (1992) (requiring a finding that “the appeal is not for purpose of delay” to release a defendant pending appeal); APPELLATE LOCAL RULES REPORT, supra note 3, at 18.
\textsuperscript{325} See FRAP 15.1 & Note to 1986 Addition; 2D CIR. LR 15 (2008).
adopt procedures already in force in most circuits. Now all that remains of the rule is the provision for filing paper copies under certain limited circumstances.

The court repealed LR 41, authorizing immediate issuance of the mandate upon disposition of: (1) cases decided in open court, (2) extraordinary writs, and (3) cases dismissed for default in filing. This rule was unnecessary because FRAP 41 permits the court on its own to shorten or extend the time to issue the mandate in any case. This rule was also inconsistent with FRAP 41 because, according to the Appellate Local Rules Report, the court must "make a determination in individual cases of whether the mandate should issue immediately. The [Second Circuit] local rule is inconsistent with this procedure in permitting the immediate issuance of the mandate in whole classes of cases without any individual examination." Furthermore, each of the three situations LR 41 marks for an immediate mandate are no longer pertinent. Cases which are decided in open court have become exceedingly rare, and the panel can always include in their order that the mandate must issue immediately. Mandamus or other extraordinary relief is also sufficiently rare as to make LR 41 irrelevant. As for defaults in filing, most are non-willful and parties cure them when given the opportunity. An immediate mandate would lead to motions for reinstatement and recall of the mandate, generating additional work for the court and staff without any offsetting efficiencies.

CONCLUSION

With the adoption of the 2010 Second Circuit Local Rules, counsel and parties in the Second Circuit will encounter much clearer direction in satisfying the procedural requirements of appellate practice. The goals of transparency, clarity, and

---

326 See 2D CIR. LR 21 (2008); FRAP 21 & Note to 1996 Amendments.
328 See 2D CIR. LR 41 (2008).
329 See FRAP 41(b).
330 APPELLATE LOCAL RULES REPORT, supra note 3, at 73.
accessibility have been largely achieved. Granted, a handful of the local rules continue to repeat national directives, but only to the extent deemed necessary to provide the bar with adequate notice of some of the more radical changes represented by this wholesale revision.\(^{331}\) Inconsistencies have by and large been eradicated. Additions to and departures from FRAP exist for the most part only when FRAP itself invites a local rule\(^{332}\) or when local variations justify legislating in an area.\(^{333}\)

The enduring challenge is for the Second Circuit to maintain the vitality and congruity of its rules both in their particular applications and when read as a whole. Three undertakings, of varying frequency, will enable the court to meet this challenge. First, the court should regularly monitor amendments to national rules, statutory developments, and Judicial Conference directives as they implicate the local rules. No less than annually, the court should determine whether any developments on the national level require corresponding attention to local rules. However, except in the face of a specific national mandate or urgent administrative need, amendments to the local rules should be prudently timed to avoid imposing on the bar a continual burden of learning and adjusting to new practice protocols. Thus, for routine, technical, or stylistic amendments, the court might consider issuing rules revisions no more frequently than every other year.

Second, the court should periodically audit the operation of its local rules. No less than every three to five years, Clerk’s Office staff and regular practitioners should be surveyed as to whether particular rules are having their intended effect and are responding to local needs in the most efficient manner. Rules should be regularly reassessed to determine if they have become obsolete, impose unnecessary or undue costs and burdens on parties or the court, or simply do not work in the manner intended. New rules might be proposed to address the changing litigation environment and the court’s dynamic docket management needs.

Third, no less than every ten years, the court should reappraise

\(^{331}\) See, e.g., supra notes 142, 170, 203 and accompanying text.
\(^{332}\) See, e.g., supra notes 155, 162, 200 and accompanying text.
\(^{333}\) See, e.g., supra notes 281–82 and accompanying text.
the entire body of local rules and internal operating procedures, and adjust them as needed. Although it may not be necessary to embark on the scopic overhaul that led to the 2010 revision, a decennial holistic survey of the rules will avert the degree of disorder and atrophy that requires such an overhaul.

Efficiency should be a circuit court’s ultimate goal in local rulemaking. The geographic, environmental, and cultural differences among the thirteen federal circuits suggest that some local rules variation will always be necessary. Local appellate rules variation is soundly motivated when it arises out of the court’s internal collegial relationships, its unique docket management needs, or the local legal culture and bench/bar relations. However, those differences are best honored by local rules that comply with FRAP strictures and Judicial Conference guidance. At this writing, no appellate court other than the Second Circuit has voluntarily scrutinized its local rules to ensure such compliance. With the adoption of the 2010 revised Second Circuit local rules, a roadmap for compliance is now available to the other circuits.
DEBIASING ADVERTISING: BALANCING RISK, HOPE, AND SOCIAL WELFARE

David Adam Friedman*

ABSTRACT

This article explores the nuances of “debiasing through law,” a regulatory approach proposed by Christine Jolls and Cass Sunstein. Proponents have described advertising regulation as a form of debiasing through law. Debiasing approaches regulation through less paternalistic means using strategies that improve consumer decision making. Proponents believe that the improved decision making will gracefully lead to better choices. These debiasing strategies find some of their roots in addressing human cognitive biases and bounded rationality.

Although debiasing through law purports to offer an attractive alternative to pure paternalism, this Article demonstrates that the net social welfare effects of this approach can prove difficult to anticipate. An examination of the recent tightening of the regulation of consumer endorsements in advertising illuminates the challenges and nuances of the debiasing approach.

This Article attempts to answer a myriad of questions. Does removing the consumer endorsement tool from the marketing arsenal ensure that consumers will make better decisions? The

* B.A., J.D., Yale University. Assistant Professor of Law, Willamette University. Thank you to Michael Abramowicz, Laura Appleman, Richard Birke, Caroline Davidson, Brannon Denning, Chris Jay Hoofnagle, Kristin Madison, Nancy Rapoport, Spencer Weber Waller, and Norman Williams for offering their time to give helpful guidance. Thank you also to the workshop participants at the Cumberland School of Law for comments and insights. The author expresses special gratitude to Christine Jolls for helpful suggestions.
truthful consumer endorsement can offer other consumers legitimate hope. When netted out, does taking “hope” off the market advance social welfare? Can regulators remove true information from the marketplace and expect to foster an improvement in social welfare? Can regulators preserve real choice while simultaneously improving decisions? Are regulators truly capable of harnessing cognitive and social science to engage in this sort of precision welfare engineering? Consumer-endorsement regulation presents a live proxy for the analytical dissection of debiasing through law.

Some scholars have expressed a keen skepticism toward this approach to regulation and toward the whole of behavioral law and economics. This Article presents a cautionary tale about the unpredictability of the net outcomes from deploying debiasing through law.
Debiasing Advertising

Table of Contents

INTRODUCTION

I. THE PREMISE OF DEBIASING THROUGH LAW
   A. One Key Challenge of Debiasing: The Endowment Effect
   B. The Mechanics of Debiasing
      1. The Product-Safety Analogy
      2. Product Safety Compared with Advertising Regulation
   C. Debiasing, Advertising, and Information Flow
      1. Advertising Regulation and Debiasing
      2. Regulating Information Flow
         i. Adding Information
         ii. Removing False or Misleading Information
         iii. Removing True Information

II. A CASE STUDY FROM ADVERTISING REGULATION
   A. The Power of Consumer Endorsements
   B. The New Regulatory Approach and Potential Impact
      1. Previous Regulatory Approach
      2. New Approach
   C. Benefits of the New Approach: Neutralizing Cognitive Biases
      1. Optimism Bias and Hope
      2. Availability Bias
      3. Framing Effect
      4. Representativeness Bias
      5. Confirmation Bias

III. POTENTIAL COSTS AND FAILURES OF THE NEW APPROACH
   A. Removing “Hope” from the Market
   B. The Cost of Failed Transactions
   C. Unintended Consequences: Price Effects
   D. The Stubborn Normative Question

IV. RECOMMENDATION TO POLICYMAKERS

V. CONCLUSION
INTRODUCTION

Jared Fogle not only stumbled across a fantastic way to lose weight, but he also had impeccable timing. The advocate for using Subway sandwiches as a delicious vehicle for weight loss managed to get discovered and signed to a promotional deal in 1997—over a decade before the Federal Trade Commission (FTC) tightened its guidance on “consumer endorsements.”1 Jared offered a powerful narrative to consumers about his experience consuming Subway products.2 Even though Jared Fogle’s experience may have been freakish, consumers received the message that at the outer boundary, consumption of Subway products could produce extraordinary weight-loss results.

The FTC’s 2009 revisions to the regulation of consumer endorsements, however, would have smothered in the cradle the welfare-yielding transactions that ultimately resulted from the Jared campaign. Reacting to the strong, psychologically persuasive pull of true information transmitted from a peer,3 the FTC literally

---


Debiasing Advertising

warns advertisers choosing to use consumer endorsements to prepare for litigation. Specifically, the FTC counsels advertisers to generate and retain empirical data proving that their consumer-endorsement advertisements are non-deceptive. The FTC recommends this practice in order to avoid the initiation of an FTC action. Had Jared Fogle achieved his Subway weight-loss feat today, Subway would have had substantially more risks to weigh in employing his endorsement. Had the threat of FTC action prevented this advertising campaign, Jared and Subway would have suffered welfare losses, but so would consumers who might have benefited from the message.

The revision of the FTC consumer-endorsement guide illustrates some of the complexities of evaluating this particular application of behavioral law and economics in regulation, or as Christine Jolls and Cass Sunstein have deemed it, “debiasing through law.” Debiasing through law purports to bring attractive features to the regulatory arena, achieving welfare-enhancing outcomes without paternalism. The debiasing approach involves influencing human decision making up-front and at the source, and enabling people to make better choices and presumably to enjoy better outcomes. Debiasing proponents wish to avoid the
“paternalistic” approach of removing actual choice from the
marketplace, while changing the environment to promote improved
decision making.8

A regulatory approach this delicate and dramatic, hinging on
inexact or controversial behavioral science, invites scrutiny.9 Some
Debiasing Advertising

scholars have expressed a keen skepticism toward this approach to regulation and toward the whole of behavioral law and economics. \(^10\) This Article presents a cautionary tale about the

unpredictability of the net outcomes from debiasing, using as a vehicle the anticipated effectiveness and potential unintended consequences resulting from the new FTC guide.

Additionally, Jolls and Sunstein posit with specificity that advertising regulation presents an actual pre-existing example of debiasing through law.\(^\text{11}\) Advertising regulation usually focuses on purifying or supplementing the flow of information that advertisers use to influence consumer decision making.\(^\text{12}\) Though Jolls and Sunstein argue that debiasing through law appears less intrusive and paternalistic,\(^\text{13}\) this Article argues that upon closer examination through the lens of a specific form of advertising regulation, the net effects of debiasing can prove difficult to anticipate. Examining the new consumer-endorsement approach in detail can help illuminate the strengths and weaknesses of the claims embodied by Jolls and Sunstein. Are regulators truly capable of this sort of precision welfare engineering? Can regulators preserve real choice while simultaneously improving decisions? Consumer-endorsement regulation presents a live example of debiasing in action.

Although the new approach to consumer endorsements may alter decision making and protect some consumers from entering into social-welfare-destroying transactions, this Article notes that this approach may cause consumers to forgo welfare-creating opportunities as well. Removing true information from the decision-making realm can lead to failed transactions and perhaps other welfare-reducing market distortions.

Moreover, in many scenarios, the consumer endorsement

\(^{11}\) Jolls & Sunstein, supra note 6, at 216.


\(^{13}\) \textit{See} Jolls & Sunstein, supra note 6, at 200–03.
Debiasing Advertising

actually comprises part of the substance of the actual offering. The truthful consumer endorsement can offer legitimate hope to the recipient—it presents the upper boundary of an expected range of outcomes from a transaction. When netted out, does “taking hope off the market” advance consumer welfare? Does removing a lonely person’s hope that a dating service might bring them down the course toward marriage advance consumer welfare? Does removing hope that eating healthier sandwiches will induce weight-loss advance consumer welfare?

Sometimes consumers contract for hope and are willing to accept disappointment. They expect to purchase a set of probabilities. This construct might apply to dining at unfamiliar restaurants, buying a new lawnmower, or buying an article of clothing. This might also apply to trying a new diet system, a gym membership, an acne-treatment cream, or a dating service. Consumers may indeed expect an offering to be imperfect or to perform below the highest level, but the information that the offering worked well for someone like them gives them the impetus to try it. As long as that hope is based on a realistic range of expectations, what is inherently wrong with offering that range of expectations?

Ultimately, advocates of debiasing still must wrestle with a core normative problem. What constitutes “good” decision making? What constitutes a “good” outcome? Debiasing through law generally emphasizes promoting the former with a more casual emphasis on the latter. Nonetheless, as the consumer endorsement circumstance demonstrates, in order to answer either question, a view must be formed about the presentation of risk—and the promotion of risk. One person’s risk is another person’s hope, aspiration, or thrill. Debiasing still leaves policymakers with another intrusive, paternalistic judgment call to make.

Part I of this Article describes the overall premise of debiasing through law, including its roots in behavioral law and economics. Then, Part I describes how debiasing works in theory and practice, using a comparison of product-safety regulation to advertising regulation. Focusing further on the underpinnings of advertising regulation, the Article explores the relationship between debiasing and regulatory control over the flow of commercial information to
consumers. Operating within the sphere of advertising regulation, Part II uses the case study of consumer endorsements to illustrate and evaluate this attempt to debias through law. Here, the Article discusses several cognitive biases potentially exploited by consumer endorsements and evaluates how the regulatory changes address them. Part III explores the potential downside to the new regulation, specifically, a reduction in social welfare due to various market distortions, including transactions that fail due to insufficient information and disruptions in pricing. This part also raises a normative question unsolved by debiasing through law—the degree to which risk should be present in the marketplace. Part IV prescribes a recommendation to policymakers that use debiasing techniques, suggesting that debiasing should be applied narrowly to certain markets because sweeping approaches tend to have more unanticipated potential pockets of downside. The Article concludes by reinforcing the notion that regulatory intervention in human decision making can yield positive results, but that the venues for deploying debiasing tactics should be carefully selected.

I. THE PREMISE OF DEBIASING THROUGH LAW

An understanding of the premise of debiasing through law requires a short description of behavioral law and economics because that is where debiasing finds its roots. Pure law and economics assumes that actors will behave rationally and respond predictably to incentives. Any errors that actors make in judgment will, according to the law and economics adherents, balance out in the aggregate and in the long run.14 Behavioral law and economics emphasizes that human beings are not disposed to behave rationally at all times and in all situations—human rationality is bounded.15 Delimiting those bounds can generate controversy,

14 This has been described as the "rational expectations hypothesis." See John F. Muth, Rational Expectations and the Theory of Price Movements, 29 ECONOMETRICA 315, 316–17 (1961).

15 Economist, policymaker, and controversial former Harvard University President Larry Summers once prefaced an unpublished article attacking the efficient-market and rational expectations hypothesis with the words, "THERE
Debiasing Advertising

particularly when policymakers attempt to incorporate and correct notions of bounded rationality. 16

People make certain judgment and perception errors consistently and in the same direction. These non-random errors do not “balance out” in the long run. Those who advocate for a debiasing approach suggest that regulators can address these biases and improve human decision making, however defined.

Much of the current regulatory apparatus attempts to ameliorate the outcomes that result from bounded rationality, presuming that core human behavior proves highly resistant to external efforts to change it. 17 Most of these traditional regulatory attempts to change behavior aim to “insulate” people from making injurious choices by restricting the choices available for them to make. 18 The new debiasing-through-law approach presumes the


16 For example, well-established research supporting the existence of the endowment effect, which predicts that people are reluctant to trade what they already have for an item of equivalent value, has come under some fire. See supra note 11. The endowment effect, the initial identification of which largely has been attributed to Richard Thaler, has been a firmament in behavioral law and economics since the early 1980s. See generally Richard Thaler, Toward a Positive Theory of Consumer Choice, 1 J. ECON. BEHAV. & ORG. 39 (1980) [hereinafter Thaler, Toward a Positive Theory]. See infra Part II for a more detailed description of the endowment effect, its implications for the Coase Theorem, and the research that casts doubt on both its existence and the explanation for its existence. See also Plott & Zeiler, The Willingness Gap, supra note 10; Plott & Zeiler, Exchange Asymmetries, supra note 10.

17 Jolls & Sunstein, supra note 6, at 200.

persistence of a slew of irrational human-behavioral quirks and modifies rules to accommodate or address these quirks.

Recently, applied behavioral law and economics has aspired to alter the behavior of actors without the harsh paternalism of removing people’s ultimate choices from the realm. The new emphasis of behavioral law and economics puts greater weight on improving the quality of individual decision making rather than on mitigating potential bad outcomes resulting from poor choices. The core of the premise rests on the notion that if cognitive psychology affords regulators with knowledge of the consistent cognitive biases of human beings, regulators can counter, de-program, or even leverage these biases. If policymakers know what heuristics or shortcuts people use to make decisions and how these heuristics can lead to error, they can generate awareness of these systematic errors, or counteract them, and improve decision making. Policymakers could employ regulatory tools to neutralize the decision-making “quirk” rather than dictate the outcome.

This Article analyzes a debiasing effort in the advertising regulation realm and assesses the effectiveness of that effort to enable extrapolation of some potential general lessons. This analysis requires an up-front foundational explanation of the mechanics of debiasing through law.

---

19 Sunstein & Thaler, Nudge, supra note 7, at 4–14.
21 Much of the cognitive psychology and behavioral economics research into the endowment effect has been performed in human laboratory experiments. For a survey of this research and an overview of the broader reaction to it, see Christopher Curran, The Endowment Effect, in 1 Bibliography of Law and Economics 819, 819–21 (Boudewijn Bouckaert & Gerrit De Geest eds., 2000). Though the results of the experiments should not be dismissed wholesale, the artificiality and errors of this setting should inject some caution into those who might rely upon them for assessing the effects of policy or tweaking policy.
22 See Jolls & Sunstein, supra note 6, at 200–01. See generally Camerer et al., supra note 7; Sunstein & Thaler, Not an Oxymoron, supra note 7; Thaler & Sunstein, Libertarian Paternalism, supra note 7; Sunstein & Thaler, Nudge, supra note 7.
Debiasing Advertising

A. One Key Challenge of Debiasing: The Endowment Effect

A full understanding of the cognitive forces at work in transactional decision making requires exposure to some brief foundational background about the endowment effect. The endowment effect describes a phenomenon where people tend to overvalue wealth or objects with which they are initially endowed. Advertising attempts to induce transactions and the endowment effect may present a force that the advertiser must counter through persuasive tactics. The true power and nature of the endowment effect remains unsettled to some, which concededly complicates the transactional decision equation, but only to a degree.23

This section describes the endowment effect debate with the caveat that a bounded understanding of the cognitive biases should perhaps make regulators a bit more cautious when attempting to neutralize or leverage them to improve human decision making.24 Setting this foundation facilitates exploration of the mechanics of debiasing through law, as applied to product-safety regulation, advertising regulation, and more broadly, control over the flow of information.

The endowment effect may drive an observed behavior where people tend to value something that they own at a higher mark than an equivalent item offered in exchange.25 This puts a drag on the tendency to exchange or trade. There may be a biological or evolutionary explanation for this phenomenon left over from an epoch when it was considerably more risky to give something up than to retain it.26 This prehistoric explanation has competed with a

---

23 With or without the weight of an endowment effect, advertisers must and still do tap into long-recognized cognitive biases to induce transactions.

24 See Stephen M. Bainbridge, Mandatory Disclosure: A Behavioral Analysis, 68 U. Cin. L. Rev. 1023, 1056 (2000) [hereinafter, Bainbridge, Mandatory Disclosure]. Regulators may be prone to the same biases and cognitive errors as those they regulate and protect.


purely economic explanation—actors in transactions are more reluctant to part with what they have if they perceive that there is counterparty risk.\textsuperscript{27} The knotty problem of the endowment effect, if not explained away by counterparty risk, presents a basic illustration of conflict between law and economics and behavioral law and economics.\textsuperscript{28} For example, the Coase Theorem is premised on the notion that initial allocation of property rights is irrelevant for ultimate efficient allocation of property, and that the sole friction that would prevent this ultimate state from materializing results from transaction costs.\textsuperscript{29} If the National Football League allocated tickets to the Superbowl through a random lottery, the Coase Theorem would predict that the recipients of the tickets would trade them to people who had a stronger natural preference to attend the game. Small-scale social experiments have indicated that this would not be the likely result. The lottery winners would be much more likely to use the ticket and attend the game because of the endowment effect.\textsuperscript{30} Recent research\textsuperscript{31} and conjecture have put the significance of that result in doubt, which presents some challenges and urges some caution for making policy from this social science.\textsuperscript{32}

\textsuperscript{27} That is, the risk of the other party not fulfilling the commitment. See Carney, supra note 10; Wright, supra note 10; Smith, supra note 10.

\textsuperscript{28} Though this Article does not have a mission to litigate the issue, the existence and impact of the endowment effect has been strongly called into question. See, e.g., Plott & Zeiler, The Willingness Gap, supra note 10; Plott & Zeiler, Exchange Asymmetries, supra note 10.


\textsuperscript{30} Daniel Kahneman, Jack L. Knetsch & Richard H. Thaler, Experimental Tests of the Endowment Effect and the Coase Theorem, 98 J. POL. ECON. 1325, 1344–45 (1990). Also note, the “status quo” bias may also intersect here. The status quo bias would also support the behavioral attitude to desiring to keep what one already has. Kahneman, Knetsch & Thaler, supra note 25, at 197–99.


\textsuperscript{32} See supra note 27 and accompanying text (discussing counterparty risk).
Debiasing Advertising

The degree of acceptance of the legitimacy of the endowment effect can be critical to understanding debiasing through law, because marketing and advertising purpose themselves to address it. Marketing and advertising attempt to induce a purchase: to induce consumers to part with wealth in exchange for an offering. If one anticipates a vital endowment effect, one can understand marketers’ need and purpose for massive doses of persuasive advertising. In order to drive transactions, marketers must use advertising to enhance the perceived value of an offering, an offering presumably presented at a competitive market price. Some offerings may require more persuasion than others to sell at a viable price for producers, when the brand or offering might be unfamiliar or appear risky to try.

Essentially, marketers and regulators (presumably acting on the behalf of consumers) engage in a tug-of-war over which psychological tactics constitute “fair game” in the inducement of a transaction.

See also David Dunning, George Loewenstein & Leaf Van Boven, *Egocentric Empathy Gaps between Owners and Buyers: Misperceptions of the Endowment Effect*, 79 J. OF PERSONALITY & SOC. PSYCHOL. 66, 73–75 (2000) (raising a new notion that empathy gaps between transactors may drive the endowment effect gap and that this gap may not be easily addressable through debiasing that employs empathetic efforts).

33 Marketing and advertising aim to move the consumer to part with a dollar to buy a can of soda, but also a specific brand of cola.

34 And also related prospect theory. See infra note 176.

35 The three leading advertising spenders in the United States in 2009 were companies managing established brands. Procter & Gamble Co. spent $4.2 billion, Verizon Communications spent $3.0 billion, and AT&T spent $2.8 billion. *Marketer Trees: 2010, Database of 100 Leading National Advertisers*, Ad Age (Jun. 21, 2010), http://adage.com/marketerstrees2010/.

36 And reduce perceived risks related to the transaction.

37 Necessarily factoring into this tug-of-war, the First Amendment also restricts regulation of commercial speech. The current doctrine is largely drawn off the line of cases evolving from *Central Hudson Gas & Electric Corp. v. Public Service Commission*, 447 U.S. 557 (1980).
B. The Mechanics of Debiasing

The techniques that advertisers use to counteract the endowment effect (or simply to persuade) often rely upon leveraging the optimism bias, the availability heuristic, the framing effect, and a host of other cognitive quirks that they can tap. Advertising regulation purports to make this contest fair, however defined. Though false information certainly cannot be conveyed as part of advertising, misleading information can be trickier to identify and may present the more challenging regulatory dilemmas. Regulators must balance the need to promote conveyance of information and the free flow of welfare-producing transactions against notions of fairness—no small challenge.

Debiasing through law purports to offer a subtle solution. The debiasing premise would dictate that policymakers not ban outright certain types of advertising, which would effectively have the hard (as opposed to soft) paternalistic result of effectively removing choice from consumers. Certain products may only sell well using certain techniques. This may be for the better or the worse, but if advertising techniques for certain products are eliminated, choice through product availability may diminish, if not completely, almost so.

Debiasing proponents would argue for counter-programming campaigns of sorts that would foster improved consumer decision making about the purchase. As Jolls and Sunstein note:

[T]he law of deceptive advertising is thus a form of debiasing through substantive law; like our proposed approach here, it adopts a middle ground between inaction

---

38 Presumably, they are drawing upon both formal and informal understandings of human cognitive biases. See infra Part II for descriptions and illustrations of these cognitive quirks.


Debiasing Advertising

or naïve informational strategies on one hand, and... strategies of heightened liability standards or outright bans, on the other. Of course, the law of deceptive advertising concerns limits on affirmative statements (Advertisements) firms choose to undertake, while [debiasing through law] concerns requirements that firms affirmatively provide information in particular ways; but in both cases the underlying attempt is to use law to reduce the severity of consumer error.41

Debiasing aims to strike a balance by generating improved consumer decisions in an arena where consumers are bombarded by marketers with both discrete bits and vivid narrative presentations of commercial information.

1. The Product-Safety Analogy

Product-safety law provides a stark illustration of the stakes that those who would use debiasing through law for advertising are trying to address, and a helpful bridge to analyzing commercial advertising regulation. Both product-safety regulation and advertising regulation try to control information and product availability while maximizing consumer choice and minimizing consumer risk or danger.42 Product-safety law may provide some clues about why the FTC may have come up short in their goals with the new regulations for consumer endorsements, when both are viewed through the approach of debiasing through law.43 (Conversely, an examination of consumer endorsements may ultimately inform a better understanding of debiasing through law.)

With product safety, regulators attempt to inform consumers of the risks of product danger in a manner that overcomes the natural optimism bias.44 However, a few complications emerge that prevent an enabling of an accurate risk assessment by the consumer:

41 Jolls & Sunstein, supra note 6, at 216.
42 In sum, these regulations optimize consumer welfare.
43 See Jolls & Sunstein, supra note 6, at 207.
44 See infra Part II.C.1 for additional discussion of the optimism bias.
Various commentators . . . have emphasized [that] optimism bias may lead many consumers to underestimate their personal risks even if they receive accurate information about average risks—a problem that may arise whenever the key piece of information to be disclosed to individuals is a probability estimate (such as the risk of harm from a product) rather than a certain outcome.45

Given these consumer proclivities, what can product-safety regulators do? The options product-safety regulators can consider start with a total a-paternalistic (or libertarian / caveat emptor) approach, putting the product on the market and allowing consumers to learn about the product’s safety attributes purely through word-of-mouth or direct experience.46 That is, an absence of regulation that requires no disclosure and no tweaking of the product to make it safer, which would require torts-driven private assessments of risk management.47 Leaving the calculations entirely to private actors could be costly—and much of our product-safety regimen maintains well-established social and regulatory roots that reflect the notion that the tort system alone cannot promote a desired aggregate level of safety and that prevention of actual harm through regulatory mandates may be required.48

45 Jolls & Sunstein, supra note 6, at 207.
46 In reality, this option does not often exist. Automobile manufacturers, for example, must achieve certain safety thresholds to let their products go to market. Above that minimum standard lies a zone for consumers to trade off price, features, and safety. See Howard Beales, Richard Craswell & Steven C. Salop, The Efficient Regulation of Consumer Regulation, 24 J. L. & ECON. 491, 492–93 (1981).
47 Note that in the absence of regulation, tort law would predict that producers would seek the proper mixture of disclosure, enhanced safety, and acceptable risk. There is a vast literature on product liability law (and associated law and economics) that this Article recognizes but does not have adequate room to address.
48 For an historic call to action to prevent consumer injury, see generally RALPH NADER, UNSAFE AT ANY SPEED: THE DESIGNED-IN DANGERS OF THE AMERICAN AUTOMOBILE (1965). For an overview of government intervention to promote food and drug safety, see generally PHILIP J. HILTS, PROTECTING AMERICA’S HEALTH: THE FDA, BUSINESS AND ONE HUNDRED YEARS OF
Debiasing Advertising

The next approach on the spectrum could be categorized as “weak-form” debiasing. This approach would be the provision of raw statistical risk information to the consumer, while leaving the product on the market. Under most circumstances, this approach proves somewhat ineffective because consumers make “systemic mistakes in the assessment of probabilities.”

Advocates of the “strong-form” (or the standard) debiasing approach would posit that making available a “concrete instance of the occurrence” would be most effective, while not intrusive, nor completely hands-off. A “truthful narrative[] of harm [would provide] a more modest and measured approach to optimism bias” than approaches that might be more paternalistic or “over-the-top” in their messaging to consumers. These techniques are guided toward improving decision making unobtrusively without the heavy hand of removing choice from the market. Choice generally, and choice to transact, has an inherent value and should not be casually eliminated by regulators if people can consider the choice in a properly informed manner.

REGULATION (2004).

---

49 Jolls & Sunstein, supra note 6, at 210.

50 I.e., “[I]n the consumer safety context the law might require that the real life story of an accident or injury be printed in large type and displayed prominently . . . .” Id. at 213 (emphasis added). For example, presenting “truthful narratives of harm.” Id. at 215. Presentation of so-called worst-case scenarios can cause consumers to either unduly overreact or simply tune out. Id. at 214–15.

51 See id. at 210.

52 Id. at 215. An aggressive approach might be exemplified by the well-known ad where an egg was displayed, with the narrator proclaiming, “This is your brain,” and then after the egg hits a hot frying pan, continuing, “This is your brain on drugs.” Shaila K. Dewan, The New Public Service Ad: Just Say “Deal with It,” N.Y. TIMES, Jan. 11, 2004.


54 With product safety, choice can be eliminated by a direct mandate to remove the product from the market entirely, raising disclosure requirements to a level where consumers will almost universally reject the offer (effectively removing choice), or raising safety features to an unreasonably protective level that makes the offering cost-prohibitive (again, effectively removing choice).
and Rose Friedman put the phrase “Free to Choose” into the popular lexicon.\textsuperscript{55} Thomas Jefferson noted, “[f]reedom is the right to choose: the right to create for oneself the alternatives of choice. Without the possibility of choice, and the exercise of choice, a man is not a man but a member, an instrument, a thing.”\textsuperscript{56} In the consumer-transaction context, freedom can be expressed even through the bountiful variety found in a grocery cart.\textsuperscript{57}

Moving along the paternalism spectrum, using product safety as the baseline illustration, regulators can employ “insulating” strategies, which begin to slide toward “hard” paternalism and removal of ultimate choice for the consumer.\textsuperscript{58} These strategies aim to protect the consumer from defects by elevating safety standards to burdensome levels. The costs of the insulated product can cause potential transactions to fail because risk has been almost entirely stripped from the offering. This paternalistic approach assumes that consumers should not be permitted to trade off price or other features for risk. The example \textit{reductio ad absurdum} would be if regulators required automobiles to be designed to prevent all fatalities, rendering them too expensive for everyday use. Individuals would be prevented from making the “un-insulated,” “risky,” “dangerous” choice of purchasing and owning a car, but would lose the freedom to choose a precious means of transportation.\textsuperscript{59}

At the far end of the spectrum, the ultimate strategy that would

\begin{itemize}
\item See Jolls & Sunstein, \textit{supra} note 6, at 207–08.
\item \textsuperscript{55} See generally MILTON FRIEDMAN & ROSE D. FRIEDMAN, \textit{FREE TO CHOOSE: A PERSONAL STATEMENT} (1980).
\item \textsuperscript{56} Iyengar & Lepper, \textit{supra} note 53, at 349. This quote has also been attributed to modernist American poet and Librarian of Congress, Archibald MacLeish. \textit{See}, e.g., DEREK HUMPHRY, \textit{LAWFUL EXIT: THE LIMITS OF FREEDOM FOR HELP IN DYING} 8 (1993); FRANK OBERLE, \textit{FINDING HOME: A WAR CHILD’S JOURNEY TO PEACE} 275 (2004).
\item \textsuperscript{57} But see generally BARRY SCHWARTZ, \textit{THE PARADOX OF CHOICE: WHY MORE IS LESS} (2005) (concluding that an excess of choice can generate unhappiness for the chooser).
\item \textsuperscript{58} Jolls & Sunstein, \textit{supra} note 6, at 207–08.
\item \textsuperscript{59} For a rich discussion of the inevitable tragedies of cost-benefit analysis in a resource-constrained environment, see generally PHILIP BOBBIT & GUIDO CALABRESI, \textit{TRAGIC CHOICES} (1978).
\end{itemize}
truly eliminate risk exposure for the consumer would be an outright ban of an offering. This ultimate paternalistic approach eliminates all choice for the consumer. It fails to improve consumer transactional decision making and it leaves absolutely no room for absorbing any level of acceptable risk.60

2. Product Safety Compared with Advertising Regulation

Product-safety regulation provides a crisp illustration of the regulatory approaches available for protecting consumers—ranging from zero regulatory protection, to enhancement of decision-making through debiasing through law, to reducing choice through insulation from outcomes, to eliminating choice completely. As noted previously, advertising regulation can be treated as a form of debiasing through law. At the very least, advertising regulation provides a unique testing ground for debiasing through law.

Chilling consumer endorsements (or effectively eliminating them), certainly reflects an effort to enhance regulation, to control more tightly the flow of information from the marketer to the consumer.61 At first glance, a crackdown on consumer endorsements resembles an effort to control information to enhance consumer decision making. It does not resemble an insulation strategy, as the offerings endorsed are not required to change, nor are offerings outright banned. One could look at this effort as a soft, non-intrusive (but still softly paternalistic) means of improving the considerations people mix into decisions to transact—a classic example of debiasing through law.62

A second look, one which this Article will explore in greater detail, reveals that a crackdown on consumer endorsements—effectively regulating commercial information flow by removing true information from that flow—can certainly have the power to

---

60 See generally Jolls & Sunstein, supra note 6, at 207–08.
61 The changes in consumer-endorsement regulation will be explored in significant detail infra Part II.
62 For better or worse, by removing a tool from the marketer’s arsenal, regulators have made it incrementally more difficult for marketers to counter the influence of the endowment effect on consumers.
“improve” the quality of consumer decision making, but can also have other possible side effects that might prove less desirable when considering total social welfare.

For example, there are some products or services that present an arguably “acceptable” amount of consumer risk. As will be discussed, occasionally a consumer endorsement is designed to induce the consumer to take a calculated risk, and sometimes the product or service offering is focused around risk per se. If regulation of a marketing practice is too heavy-handed and the practice becomes too risky to employ, the practice may disappear entirely. If the practice disappears entirely, certain types of offerings dependent on the practice will disappear. This might increase social welfare with certain transactions and it might subtract with others. If advertising regulation is debiasing through law, advertising-regulation initiatives should be closely scrutinized to see if debiasing actually works as intended by proponents.

C. Debiasing, Advertising, and Information Flow

Regulators are concerned with preventing “bad,” social-welfare-destroying transactions from reaching completion. Note that “bad” transactions present themselves affirmatively. Injuries, failed results, and complaints resulting from disappointed

---

63 The FTC’s commissioned studies on the weight-loss industry may demonstrate that the effective elimination of consumer endorsements in that industry, and the subsequent disappearance of the offering, might enhance consumer welfare. HASTAK & MAZIS, WEIGHT LOSS, supra note 3; HASTAK & MAZIS, DIETARY SUPPLEMENTS, supra note 3.

64 The Subway sandwich campaign discussed at length, infra Part II, may provide an example of where effectively eliminating consumer endorsements might reduce consumer welfare. On-line dating services that use married couples who met through their service may provide another illustration. The Article later suggests a middle ground where the FTC, instead of leveling a blanket threat, addresses industry-specific behavior under the broader enforcement powers of Section 5 of the FTC Act, or similarly through the Food and Drug Administration.

65 And again, vice-versa—advertising regulation can inform the broader picture of debiasing, just as the breastfeeding advertising campaign did.
Debiasing Advertising

expectations are tangible to regulators.\textsuperscript{66} Therefore, regulators often try to address what they can easily see,\textsuperscript{67} or tackle problems related to something dramatically unpleasant that recently happened.\textsuperscript{68}

For example, by tightening down on consumer endorsements, the FTC attempted to purify the information flow from producers to consumers in order to create a transactional environment where consumers become less prone to acting on misperceptions and erroneous extrapolations from true individual consumer experiences. As the Article explains, this novel debiasing approach—an approach that involves using information and presentation to improve decision making—can yield imperfect results, illustrations of which are available. The new consumer-endorsement regulations attempt to improve decision making with a seemingly benign design.\textsuperscript{69} But the welfare impact of the change

\textsuperscript{66} This applies across the regulatory board. The Food and Drug Administration’s drug approval process has at certain junctures been regarded as too cautious. \textit{See} Robert Pear, \textit{Faster Approval of AIDS Drugs Is Urged}, N.Y. TIMES, Aug. 16, 1990, at B12. The modern process can trace some of its legacy to the thalidomide tragedy in the 1950s. \textit{See} FRIEDMAN & FRIEDMAN, \textit{supra} note 55, at 202–05; \textit{see also generally,} Linda Bren, \textit{Frances Oldham Kelsey: FDA Medical Reviewer Leaves Her Mark on History}, 35 FDA CONSUMER MAG., Mar.–Apr. 2001, \textit{available at} http://permanent.access.gpo.gov/lps1609/www.fda.gov/fdac/features/2001/201_kelsey.html. Approving a dangerous drug is visible. Libertarians have argued that the social costs of over-caution in drug approval can be less visible, but that over-caution may be just as socially expensive. FRIEDMAN & FRIEDMAN, \textit{supra} note 55, at 204–10; Gary S. Becker, \textit{Get the FDA Out of the Way and Prices Will Drop}, BUSINESSWEEK, Sept. 16, 2002, \textit{available at} http://www.businessweek.com/magazine/content/02_37/b3799028.htm.

\textsuperscript{67} \textit{See, e.g.,} Omnipresent advertisements for “free credit reports” that confused consumers resulted in FTC Action. Free Annual File Disclosures Rule: Final Rule and Statement of Basis and Purpose Amending the Rule To Prevent Deceptive Marketing of Free Credit Reports, Pursuant to Section 205 of the Credit Card Accountability Responsibility and Disclosure Act of 2009, 16 C.F.R. \S\ 610 (2010).

\textsuperscript{68} \textit{See, e.g.,} Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. 111-203, 124 Stat. 1376 (2010), passed in part “to protect consumers from abusive financial services practices . . . .” that revealed themselves in the years just prior to passage of the legislation. \textit{Id.}

\textsuperscript{69} \textit{FTC Publishes Final Guides Governing Endorsements, Testimonials,
may be ambiguous and uneven across markets. From a normative perspective, these regulations may turn out to be more rigidly paternalistic than intended. They may indeed remove choice from the marketplace, more than regulators might anticipate. 70 These results can inform a broader understanding of debiasing through law.

Voluntary transactions in the marketplace generate positive social welfare, net of any externalities, because the assumption is that the parties would not have entered into the trade if each were not put in a better position. The assumption is that the trade represents the best available option for each party to improve its position and that the end result will prove to be Pareto optimal.71

Welfare economics finds that regulatory action is justified only when a specific market imperfection is evident and when the benefit from correcting that market imperfection exceeds the cost of properly regulating the market. If a market imperfection is not apparent, or if the benefits of correcting that imperfection are small, then intervention risks causing more harm than good to social welfare.72

supra note 1. “Under the revised Guides, advertisements that feature a consumer and convey his or her experience with a product or service as typical when that is not the case will be required to clearly disclose the results that consumers can generally expect.” Id. On the surface, this sounds like a means of facilitating commerce fairly, but as this paper explores, more implications must be considered.

70 For example, if regulators provide consumers with information about the risks of eating chocolate, if the information is too alarming, chocolate will cease to be a real choice. Regulators can scare decision makers away from a choice that then renders the choice extinct, even for others who would choose it in spite of exposure to and true absorption of the information. Other distortions can happen from removing information from the marketplace. For example, if regulators forbid price-based advertising, total prices may rise, which may reduce choice and welfare. See infra Part III.

71 “[A] policy change is socially desirable [and meets the Pareto criterion] if, by the change, everyone can be made better off, or at least some are made better off, while no one is made worse off.” RICHARD E. JUST, DARRELL L. HUETH & ANDREW SCHMITZ, THE WELFARE ECONOMICS OF PUBLIC POLICY 14 (2004).

Debiasing through law describes a type of regulatory intervention, a means of correcting imperfections. Engaging in this type of intervention requires an understanding of the underlying nature of the regulated transactions.

Debiasing should disrupt or improve “bad” transactions. A “bad” transaction can be rooted in contracting norms as well as welfare economics. 73 “Bad” transactions often result from “bad” information polluting the transaction environment, distorting expectations about what will be exchanged. 74 A “bad” transaction could also arise from scenarios where the parties, ex-ante, err (or engage in deceit, mild or aggressive) in their valuation of what they are surrendering or receiving. 75

These ex-ante errors are significant because the law of contract is notoriously reluctant to undo or revisit transactions that are merely unbalanced or one-sided. Generally, a party must show that there is something seriously deficient with the process of the formation of the transaction for the agreement to be undone. The defenses of duress, 76 undue influence, 77 and unconscionability, 78 for example, all require inquiries into the transaction process in order for a party to succeed. 79

For the purposes of this Article, “bad” transactions can be translated into welfare terms. When Party X makes a decision based on a distorted perception from information proffered by Party Y, and the distorted decision devalues transactional outcome for Party X, Party X obviously loses welfare. That creates a social concern. If Party Y is the counterparty and creator of misperception in this exchange, and Party Y gains more in welfare than Party X loses, this exchange might enhance welfare in a one-


73 See Beales et al., supra note 46, at 505–06.
74 Id.
75 Id.
77 Id. § 177.
78 Id. § 208 cmt. b.
79 Improper threats, the exercise of dominion over a party, and “defects” in the bargaining process are threaded throughout these defenses.
round game. But in a multiple round game, the destruction of trust will cast a pall over future transactions, ultimately reducing welfare, as fewer welfare-creating transactions and exchanges will likely result.\(^{80}\) Regulation of advertising should attempt to maximize social welfare with fairness as the constraint.\(^{81}\) Over-regulating advertising might have the same welfare-reducing effects as the open permission of bad transactions, if information proves insufficient for consumers to engage heartily in social-welfare-yielding transactions.

1. Advertising Regulation and Debiasing

Before entering the narrower problem of consumer-endorsement regulation, taking a step back to consider advertising regulation more broadly can provide context. Advertising regulation, as embodied in the Federal Trade Commission Act\(^{82}\) (and “baby” FTC Acts at the state level),\(^{83}\) focuses broadly on unfair or deceptive practices that generate transactions through the inducement of ex-ante errors on the part of the consumer.\(^{84}\) Presumably, an atmosphere of deception promoted by an advertiser would lead to a “bad” transaction—a failure to yield net instant or long-run social welfare, an unjust redistribution of surplus to the advertiser from the consumer, or perhaps, all of the above.

Regulators often focus concern on transactions that are induced


\(^{81}\) See generally Craswell, supra note 12.


\(^{84}\) In 1981, Richard Craswell surveyed the constitution of “unfair acts or practices,” including “withholding material information” (which has direct applicability to the advertising environment) and “unsubstantiated advertising claims.” See generally Craswell, supra note 12, at 107–09. For a contemporaneous analysis of the challenges of defining deceptive advertising, see Beales et al., supra note 46, at 495–501, 516–21.
Debiasing Advertising  

by one party through the exploitation of a hard-wired predisposition of a consumer to make certain cognitive errors or systematic misjudgments.\textsuperscript{85} There are some tactics of psychological manipulation that may unfairly induce transactions. Obviously, the regulation of the information presentation that induces cognitive error must be done with a high level of care, and not just for purposes of compliance with the First Amendment.\textsuperscript{86}

When advertising is disaggregated, portions of it will be pure presentation of the product, the price, and availability. These essentials comprise the core of advertising. When these essentials are subtracted from the whole of a piece of advertising, the remaining “residue” is pure persuasion. The attractive people depicted drinking beer on a tropical beach, the happy family depicted playing the board game, the free offer, the call to “buy now because supplies are limited,”\textsuperscript{87} all reflect tactics that persuade, yet convey almost none of the core substantive information about the product.

Much of advertising regulation and debiasing through law focuses on controlling, managing, or addressing this “residue.”\textsuperscript{88} It bears emphasizing again that Jolls and Sunstein contend that debiasing through law aptly describes the approach of advertising regulation.\textsuperscript{89} Of paramount importance for protecting consumer decision making is controlling the purity of market information and its presentation. However, if regulatory control over the

\textsuperscript{85} For an example of the multiple sociological and cognitive manipulations involved with so-called “free offer” advertising, see generally David Adam Friedman, \textit{Free Offers: A New Look}, 38 N.M. L. REV. 49 (2008).

\textsuperscript{86} See supra note 37.

\textsuperscript{87} This tactic triggers the scarcity effect. For an empirical demonstration, see generally Akanbi Adewole, Jerry Lee & Stephen Worchel, \textit{Effects of Supply and Demand on Ratings of Object Value}, 32 J. PERSONALITY & SOC. PSYCHOL. 904 (1975).

\textsuperscript{88} The use of the term “residue” should not imply that persuasion is unimportant to marketers. Persuasion tactics are extremely important and ingrained in our commercial culture. Entities invest heavily in creating brand identities that attempt to persuade. See Desai R. Desai & Spencer Weber Waller, \textit{Brands, Competition and the Law}, BYU L. REV. (forthcoming) available at http://ssrn.com/abstract=1545893.

\textsuperscript{89} Jolls & Sunstein, \textit{supra} note 6, at 216.
market information tightens unduly, consumers may also be put in a suboptimal position. Consumers benefit from the unilateral (or occasionally interactive) communications to which advertisers expose them in all media, presuming that the content of these communications is truthful and the persuasive residue “fair.” Availability, price, and differentiation all guide the consumer toward opportunities to make an exchange that could enhance total welfare and make both parties better off.

Nonetheless, “bad” transactions present a real concern and the changes to consumer-endorsement regulation are but one microbe in a broad federal, state, and self-regulatory environment that targets this concern. Although regulation may throw out good transactions with the bad, some of the prevented transactions may well indeed be bad transactions. Studies launched by the FTC have demonstrated that in certain instances, when consumers are presented with a consumer telling the story of an exceptional result from the offering, it raises the expected experience to be above that of the typical experience.

Accordingly, there is a subtle result—a precarious balance—that regulators should seek. Certainly, “bad” transactions should be prevented, but preventing them should be done in the most effective and efficient way. The new regulatory approach does not impose a de jure legal prohibition on consumer endorsements, but they are, from their very design, aimed to severely impede them to the point where consumer endorsements might be almost too prohibitive to employ. As this Article describes in Part II, by ratcheting up the costs and raising the risk of legal exposure related to this method of delivering information, the FTC effectively addresses all consumer endorsements in the same manner. All of the “bad” transactions that would have been generated by consumer endorsements would simply not occur. Taken alone, this would be a positive social development.

However, at least two types of “good” transactions that might

90 “Private” or self-regulation would include industry norms and third-party groups like the Better Business Bureau and Underwriters Laboratory. See Beales et al., supra note 46, at 501–02.

91 See HASTAK & MAZIS, WEIGHT LOSS, supra note 3; HASTAK & MAZIS, DIETARY SUPPLEMENTS, supra note 3.
Debiasing Advertising

be induced through this method would get swept out with the bad: (1) The transactions that required the consumer endorsement “jolt” to boost the consumer incentive to transact, and then put the consumer in a satisfactory position *ex post*; and (2) The transactions that were really about “hope” as much as anything else. Some transactions may fit into both of these typologies.

The second category of transactions (“hope”) presents a difficult category of transactions for regulators to address. Regulators are not in any way taking the offering feature of hope *completely* off the market, but with consumer endorsement restrictions, they are removing the opportunity for producers to use true outcomes from their offerings to present consumers with a hope to which they can aspire.92 The “hope” in this instance for the consumer would be the possibility of achieving the actual outcome that the consumer endorser experienced.

Concededly, examples of potentially “bad” transactions that use hope and consumer endorsements to induce transactions are commonplace. Weight-loss program offerings and “business-opportunities” programs have potential to magnify the effectiveness of the advertising tactic.93 With dietary supplements, the FTC produced two empirical studies to demonstrate the potential cognitive distortive effect on expected product results.94 With so-called “get-rich-quick” business-opportunity offerings, the expectations and results are also easily quantifiable.95 If the new consumer-endorsement guides merely altered the nature of this particular family of transactions and precisely rebalanced the cognitive playing field within this family, that could be easily justifiable. Other socially-valuable communications, however, could get swept away with a broad, blanketed approach.

For illustration, consider the chiropractor who wishes to induce customers to try controversial chiropractic therapy through her clinic.96 She publishes a website with true testimonials of

---

92 Because the tactic has become more expensive and riskier to implement, producers are disincentivized from using “hope.”
93 See HASTAK & MAZIS, WEIGHT LOSS, supra note 3.
94 *Id.*; HASTAK & MAZIS, DIETARY SUPPLEMENTS, supra note 3.
95 See HASTAK & MAZIS, WEIGHT LOSS, supra note 3.
96 Assume that her clinic is on a state border, opening her interstate
exceptional results, designed to lure skeptics to have hope that her services can help. Perhaps this inducement might be able to help someone. Arguably, the new consumer-endorsement guides would help to mitigate any questionable trade practices, but other narrowly-targeted mechanisms exist for that purpose. Questionable trade practices can be addressed with precision rather than through overly-broad regulation.

Or consider an illustration that might be closer to home for readers of a law review article. A law school that showcases “student profiles” on its website, or profiles of recent alumni, may be transmitting experiences that are exceptional, not average. The profiles may indeed present true experiences, but law schools certainly design them to induce the potential applicant to take a closer look at the school, apply, and matriculate if accepted.

Consumer endorsements are certainly not without social value in given situations, but they should nonetheless be monitored. “Bad” transactions can certainly be induced by the presentation of outlier consumer endorsements, and regulators should be on heightened alert when these endorsements are used. However, by implementing the new consumer-endorsement guides, the FTC may have painted with too broad a brush. The FTC could simply have relied on its authority under the FTC Act to commence civil enforcement actions in specific “problem” industries without transmitting a general, strong warning about the tactic.

Perhaps the FTC should apply more scrutiny toward marketers that use this advertising tactic, but should operate with a discretionary stiletto rather than the proverbial hatchet. Some forms of hope have value and are less risky to leave unfettered in the marketplace. Other forms of hope can be quite damaging if expected, then dashed.

2. Regulating Information Flow

Debiasing through law, when meshed with the consumer
regulatory arena, distills the control of information transmitted through advertising or other disclosures. Depending on the situation and ultimate goal, regulators have mandated the *addition* of information to the flow, the removal or banning of *false* information, but also, interestingly, the *removal* of *true* information from the marketplace.\(^9^9\) To contextualize properly where consumer-endorsement regulation fits into the consumer law picture, this Article next presents examples of each of these regulatory mandates, starting with mandates to add information into advertising.

**i. Adding Information**

People can make decision errors due to the absence of relevant information that would help them make better choices and debiasing can address this problem.\(^1^0^0\) Sometimes adding this information provides raw data for the decision maker to consume, other times this information can be skillfully woven into the fabric of the decision-making process. Adding a narrative about a death resulting from smoking, for example, gives consumers raw information that they likely already have (smoking is hazardous), but also adds the information in a manner that tends to stick.

These decision errors may not necessarily be curable, but providing the information (or mandating its provision) to the decision maker represents a method of debiasing. The Truth in Lending Act,\(^1^0^1\) which provides for mandatory, uniform disclosures from lenders, provides an example of a regulatory debiasing mechanism for adding information.\(^1^0^2\) Whether these

\(^9^9\) Beales et al., sought to define regulatory “information remedies” to solve consumer market failures, categorizing the remedies as “(a) removing restraints on information; (b) correcting misleading information; (c) encouraging additional information.” Beales et al, *supra* note 46, at 513–31. The authors note that “remedying deficiencies in the information market is in some ways a more complex and subtle task than regulating product markets directly.” *Id.* at 514.

\(^1^0^0\) Jolls & Sunstein, *supra* note 6, at 207.

\(^1^0^1\) Truth in Lending Act, 15 U.S.C.A. § 1601 et. seq. (West 2010).

\(^1^0^2\) For a cautionary tale about “glibly” using mandatory disclosure as a panacea in the securities regulation context, see Bainbridge, *Mandatory*
pieces of additional information are heeded by decision makers, or whether they improve decision making, can be a situational question. Over-disclosure can overwhelm consumers with too much information. The marginal information may also prove insufficient to overcome the optimism bias.

Even if the information is absorbed, the additional information may yield lower utility for the recipient. Ignorance can, indeed, be bliss. Mandating disclosure about the FDA’s standard for acceptable number of rodent hairs permissible in peanut butter (through a big orange sticker on the jar bearing a picture of a rat) may extinguish a potential transaction that would otherwise create welfare. This would be an inverted restatement of the old saying, “What you don’t know won’t hurt you.”

Mandating disclosure of information to the consumer (adding information) constitutes a classic form of debiasing. Adding information changes how consumers calculate the risks and benefits of a potential transaction and make decisions. The initial disclosure, supra note 24, at 1058–59 (2000).


Sometimes we are better off not knowing things. This near-aphorism is unremarkable when understood as an observation about our everyday lives. Do you really want to know the day, time, and circumstances of your own death? Do you want to know the details of your children’s (or parents’) love lives? Do you want to know whether your neighbors are scrupulous in paying their taxes? To these, and many questions like them, we answer almost reflexively, “I’d rather not know.” We are happier, indeed better off by many measures, if ignorant. . . . Despite the advantages knowledge often confers, ignorance is sometimes preferable because it shields us from unpleasant realities, keeps us from facing difficult choices, or immunizes us against attack by others.
Debiasing Advertising

burden of the costs borne from the added regulatory burden falls on advertisers. Policymakers would hope that the welfare gains from consumer empowerment would exceed any losses created by this extra burden.\textsuperscript{106}

\textit{ii. Removing False or Misleading Information}

Section 5 of the FTC Act forms the basic core of advertising regulation at the federal level and speaks to the core mission of the FTC in its control over the information flow.\textsuperscript{107} The FTC Act declares unlawful “unfair or deceptive acts or practices in or affecting commerce.”\textsuperscript{108} Although this sweeps broadly, this section vests the FTC with the power to compel the removal of certain materially-false bits of information from the marketplace and seek penalties for those who put the information into circulation.\textsuperscript{109} Prohibiting advertisers from making false or misleading claims falls into this category.\textsuperscript{110}

Removing nakedly false data from the decision-making pool

\textsuperscript{106} In 1981, Richard Craswell described the FTC’s approach toward regulating disclosure measured against the FTC Act’s “unfairness” standard.

[\textit{T}]he failure to disclose information . . . will be [deemed] unfair if: (1) consumers currently lack the information in question; (2) consumers would choose differently if they had the information, thus facing sellers with a different set of demand curves; and (3) the benefits of better consumer decisions and improved seller performance are not outweighed by the costs of supplying the information.

Craswell, supra note 12, at 123. Included in the third category are the costs of collecting and disseminating the additional information, and the risks that disclosure “might [somehow] mislead consumers even further.” \textit{Id.}

\textsuperscript{107} “A deceptive claim that constitutes a false statement ‘is clearly at the core of most people’s understanding of deception.’” Beales et al., supra note 46, at 496.


\textsuperscript{109} Though as Beales et al. note, this definition of deception proves too narrow as it ignores broader concerns about the effect of other “less patently deceptive” claims on consumer decision making. Beales et al., supra note 46, at 495–96.

\textsuperscript{110} See Craswell, supra note 12, at 123–27 (addressing “unsubstantiated advertising claims”). Sellers must have a reasonable basis for believing all claims they make at the time they make them. \textit{Id.} at 123.
represents an almost pure form of debiasing through law. If advertising regulation and debiasing through law are closely related, this specific portion of the FTC’s mandate is the closest cousin. When the advertising practices are patently unfair, an economic rationale undergirds this basic form of consumer protection.111

iii. Removing True Information

Under the broad auspices of Section 5, “unfair” acts are unlawful. Conceivably, the disclosure of true information could be unfair, thus unlawful. Consumer endorsements might fall into this category, if offered in a way that leads consumers to misleading conclusions.112

Certain norms direct suspicion at regulation that removes true information from the marketplace.113 First Amendment

111 Transactions induced by false claims “contribute nothing to consumer welfare.” Beales et al., supra note 46, at 496. Of interesting note is 15 U.S.C. § 45(a), which enables with specificity regulation of “Made in the U.S.A.” labeling. This is a disclosure regulation that protects producers of products of United States origin from false claims of origin by foreign competitors. Consumers would probably not lose utility from use of a product with a falsely-labeled origin if they did not know about it.

112 Craswell, in his early framework, might have bucketed this problem, the consumer endorsement problem, within a category he deemed, “difficult . . . to analyze,” “unfair persuasive techniques.” Craswell, supra note 12, at 139–51. Specifically, Craswell notes that “one-sided advertising” had been deemed unfair because although the claims were true, they were so overwhelming that they washed over countervailing consumer knowledge. Consumers were not deceived by false information, they were unfairly persuaded. The examples cited were FTC actions involving advertising about cigarettes and food products linked to child nutrition. Id. at 147–51. “True, but unfair,” could apply to consumer endorsements, but Craswell expresses difficulty in reconciling this bucket. Note that Craswell peers into the future with a plea to social science for some assistance in assessing the fairness of persuasion techniques. “It may eventually be possible to develop a rationale for these cases, by improving our understanding of consumer decision making and the psychological effects of advertising.” Id. at 150.

113 Although, seller claims (or affirmative statements lacking a reasonable basis when made that are later even proven true) are vulnerable to be deemed
Debiasing Advertising

jurisprudence reflects that anti-paternalistic norm. As Justice Stevens noted in *Liquormart v. Rhode Island*, “[t]he First Amendment directs us to be especially skeptical of regulations that seek to keep people in the dark for what the government perceives to be their own good.” In *Liquormart*, the Court struck down two Rhode Island statutes that removed true information from entering the marketplace. The Rhode Island laws prohibited advertising the retail price of alcoholic beverages away from the point of sale. This price information was core offering information, not even persuasive “residue.”

Nonetheless, the specific situation matters. Removing true information from the marketplace can assist consumers in certain decision-making environments. Of recent note, Professor Elizabeth Warren, in advocating for the establishment of a Consumer Financial Protection Agency (CFPA), argued for mandating simplification and streamlining of disclosures and forms so that consumers would have easier means for comparing financial products. This amounts to removing information to simplify the transactional environment. “Over-disclosure” can occasionally confuse and overwhelm consumers in an information-heavy environment. Warren wrote:

The truth, of course, is that no consumer “chooses” to accept the tricks and traps buried within the legalese of financial products. Rather, consumers must choose among various products with one feature in common: dozens of pages of incomprehensible fine print.

The CFPA will not limit consumer choice. Instead, it will focus on putting consumers in a position to make choices for themselves by streamlining regulations, making disclosures smarter, and making financial products easier to understand and compare. The Agency will promote plain vanilla contracts—short, easy to read mortgages and credit

“unfair.” *Id.* at 123.


115 *Id.* at 503.

116 *Id.* at 489–90.

117 See generally Ben-Shahar & Schneider, *supra* note 103.
card agreements. The key principle behind the new agency is that disclosure that runs on for pages is not real disclosure—it’s just a way to hide more tricks. Real disclosure means that a lender has to be able to explain what it is selling so that the customer can read it and understand it. Once consumers can understand the risk and costs of various products—and can compare those products quickly and cheaply—the market will innovate around their preferences.118

Warren’s proposal amounts to advocating the removal and resequencing of true information to enhance consumer decision making. The proposal presents another example of debiasing through law.119

Of course, consumer-endorsement regulation could be classified as an effort to debias through law by removing true information from the market environment. Unlike the alcohol at issue in 44 Liquormart or the financial products at issue with the CFPA, however, consumer-endorsement regulation controls a tactic that cuts across all products and services. Across the board, the FTC has effectively deemed that consumer endorsements have the power to injure consumers, and that true narratives and testimony about a usage experience must be tightly monitored, if not in reality nudged out of the commercial-information sphere. From a policy standpoint, addressing the information problems offering-by-offering might have more appeal. Alcohol, mortgages, and sandwich shops all intersect with advertising differently and all present different types and degrees of transactional risk for the consumer. Why not factor in these differences when addressing the regulatory problem of consumer endorsements rather than firing what effectively amounts to a blanket salvo?


119 See also generally Lauren E. Willis, Against Financial-Literacy Education, 94 IOWA L. REV. 197 (2008) (arguing that social detriments of promoting consumer financial literacy could exceed the benefits, implying that removing this information might be helpful to consumers).
Debiasing Advertising

Next, this Article hones in on these recent restrictions on consumer advertising embodied in the 2009 revised consumer-endorsement guides. Examining this case study in detail should illuminate the strengths and weaknesses of the claims embodied by advocates of the debiasing-through-law approach. Are regulators capable of this sort of precision welfare engineering? Can regulators preserve real choice while simultaneously improving decisions? Consumer-endorsement regulation provides an example of debiasing through law in action.

II. A CASE STUDY FROM ADVERTISING REGULATION

Consumer endorsements have a long history and heritage in advertising in the United States. The power of a real peer, a real product or service user, to induce a transaction has been acknowledged by marketers, regulators, and consumer activists.

Following the statutory mission, the FTC prophylactically seeks to eliminate and prevent unfair and deceptive trade practices and advertising, which it has power to regulate under Section 5 of

---

120 Within the definitional scheme of deceptive advertising regulation that Beales et al. describes, it is challenging, but possible to envision where the new consumer-endorsement regulation would slot in as a cure. Beales et al. largely focus on false claims, production of inaccurate consumer beliefs, and failures to disclose. See Beales et al., supra note 46, at 496–501. Most likely, they would see this regulation as addressing inaccurate consumer beliefs—and much of their discussion of the remedy for this focuses on improved disclosure. Beales et al. do not focus on the removal of true information as a remedy. As discussed infra, consumer endorsements have a complex impact on consumer decision making and social welfare.


122 See HASTAK & MAZIS, WEIGHT LOSS, supra note 3; HASTAK & MAZIS, DIETARY SUPPLEMENTS, supra note 3 (demonstrating power of consumer endorsements to create audience recall of message).

123 See, e.g., Greenberg Statement, supra note 2, at 14–20.
the FTC Act.\textsuperscript{124} Given that producers aggressively seek to induce transactions against the inertia of the closed human wallet,\textsuperscript{125} and then against competing offers of other producers, the advertising arena can be quite fierce tactically.

Virtually all advertising promotes a call to consumer action. The call to action can be pure—a mere dissemination of information about price, qualities, and availability.\textsuperscript{126} Questions start to emerge when the call starts to employ devices that move beyond communicating the “pure” component of the call to action—into persuasion. Strategic bundling and pricing of an offer can cause the consumer to value the package differently if laid out one way versus another way. The medium of the message can convey information that does not directly relate to the product. A trusted celebrity endorser or a consumer endorser can add a layer of credibility and identity to the message.

The question is whether adding these extra layers of persuasion to the plain message are unlawfully deceptive or instead a positive addition of useful content. For example, inducing the sale of a food processor by communicating that a free set of steak knives is included, instead of simply offering the processor and the knives together at the same price, may not yield incremental social welfare, but would most probably be lawful.\textsuperscript{127} Also, consider that a marketer can lawfully employ a trusted B-list actor to promote a

\begin{flushright}
\textsuperscript{125} See discussion of endowment effect, \textit{supra} Part I. Though the conclusions and implications are now wrapped in some degree of controversy, Kahneman and Tversky demonstrated through classroom laboratory experiments that people value a loss of X greater than an absolute gain of X. \textit{See generally} Daniel Kahneman & Amos Tversky, \textit{Prospect Theory: An Analysis of Decision under Risk}, 47 \textit{Econometrica} 263 (1979) [hereinafter Kahneman & Tversky, \textit{Prospect Theory}]. At the front end of a transaction, the implication is that inducing just an even exchange of value requires the marketer to overcome a level of cognitive impulse for the consumer not to part with what they already have. Kahneman and Tversky’s loss aversion work has further implications for analysis of consumer-endorsement regulation, specifically, in assessing the consumer’s reaction to the results of a transaction.
\textsuperscript{126} The call to action can also be long-run—the creation of a brand image that communicates a message about the values behind a product.
\end{flushright}
Debiasing Advertising

safety-alert system for older people. Even if disclosed as a compensated endorser and not a user, the somewhat misplaced trust in the actor is designed to rub off onto the no-name alert system. The former case we may accept because it is practically difficult to control that form of bundling manipulation. The latter case could be troubling on the surface, but what if the offering is an honest one and the target audience needed the comfort of the celebrity endorsement to enter into a “good” transaction?

The consumer endorsement takes advantage of many of the same biases as the rest of advertising. Advertisers count on tapping into an array of consumer biological and cognitive biases to induce transactions, and, in turn, mitigate the biases that inhibit them to act. As noted, the advertising arena, a petri dish for psychological and cognitive games, provides a natural testing ground for debiasing through law. In particular, consumer endorsements have been wide open for debiasing “treatment” and the FTC has, in fact, accorded them this treatment. Section A begins with a description of the power of consumer endorsements, using a well-known advertising campaign as a working example; Section B describes the change in regulatory approach and its potential impact in general terms; and finally, Section C analyzes the various biases that the new approach uses to effectively neutralize the power of consumer endorsements.

A. The Power of Consumer Endorsements

In 1998, morbidly-obese Indiana University college student Jared Fogle began a journey toward achieving iconic celebrity and cultural-touchstone status. Eating nothing but weight-loss

---

128 This genre and application of celebrity endorsement ad has become so common that actor Sam Waterston spoofed it on Saturday Night Live, pitching “robot insurance” (as a “compensated endorser”) to senior citizens to protect them from “robot attacks.” Saturday Night Live: Laura Leighton/Rancid (NBC television broadcast Nov. 18, 1995).

129 See 16 C.F.R § 255.5 (2011).

130 Bob Levey, Column, Q&A with Bob Levey, WASH. POST (Mar. 18, 2003), http://discuss.washingtonpost.com/wp-srv/zforum/03/rMetro_levey_031803.htm. Fogle has also been spoofed on Saturday Night Live and even in an
friendly Subway\textsuperscript{131} sandwiches, Fogle ultimately shed and kept off over two hundred pounds.\textsuperscript{132} Two years later, sharp-eyed regional Subway franchise owners received word of Fogle’s astonishing achievement and retained him to tell his story in their advertisements. Not long after the regional commercials demonstrated tremendous effectiveness, Subway’s corporate headquarters recruited “Jared,” as he would come to be known, for a national campaign.\textsuperscript{133} Over the next decade, “Jared” lured not just crash dieters, but all kinds of health-conscious consumers through Subway’s doors, helping to spur Subway’s location presence to equate that of McDonald’s, the ultimate food-industry franchising pioneer, by 2009.\textsuperscript{134}

Although a number of factors likely drove the astonishing growth of Subway since 2000, the “Jared” campaign, according to CEO Jeff Moody, provided the “absolute cornerstone” for “Subway’s success.”\textsuperscript{135} Moody opined:

I believe the company would have been successful without Jared but not nearly to the degree that we are because he took these healthy products and gave them a face and said something that people could relate to. I mean he’s not a celebrity because he set out to be, he’s a celebrity [be]cause . . . he resonated with people that had similar issues to him and said, well if that guy can do it, I can do it . . . .\textsuperscript{136}


\textsuperscript{131} Subway, Inc. franchised 22,000 sandwich shops in the United States as of 2008. \textit{See Jared’s Revenge, supra} note 1.

\textsuperscript{132} \textit{Id.}

\textsuperscript{133} CHIP HEATH & DAN HEATH, \textit{MADE TO STICK} 218–221 (2007). Heath and Heath offer additional compelling explanations for the Jared campaign’s enduring success.


\textsuperscript{135} \textit{Jared’s Revenge, supra} note 1.

\textsuperscript{136} \textit{Id.}
Debiasing Advertising

offer. Fogle, at least at the outset of his arrangement with Subway, would qualify as a “consumer endorser” by the FTC. Under the Guides Concerning the Use of Endorsements and Testimonials in Advertising that were in effect until 2009 (“Previous Guides”), Subway could use Fogle’s outlier endorsement, provided that the company disclaimed the endorser’s experience as an atypical result.

Due to the recent changes in the Guides, this highly-effective type of consumer endorsement campaign would be significantly more impractical to launch effectively now. Had Subway wished to launch a consumer-endorsement Jared-style campaign today, under the new guides (“Revised Guides”), the company would have had to choose between two regulatory compliance paths. Subway could either choose to disclose scientifically-derived, empirically-sound, “typical” results alongside Jared’s results, significantly cluttering the message or Subway could simultaneously broadcast an “atypicality” disclaimer, similar to the one previously required but without any accompanying safe-harbor protections. The FTC warns that in this new regime, the disclosure route should also be backed up by empirical evidence that the totality of the advertisement is non-deceptive. This looming threat of

---

137 Arguably, Jared Fogle has now achieved a level of fame that would also subject him to regulation as a *celebrity* endorser for Subway under 16 C.F.R § 255.5 (2009).

138 The disclaimer provided a safe harbor. As an alternative, under the Previous Guides, advertisers could present the expected results alongside the exceptional results presented by the consumer endorser.

The Guides are administrative interpretations of the law intended to help advertisers comply with the Federal Trade Commission Act; they are not binding law themselves. In any law enforcement action challenging the allegedly deceptive use of testimonials or endorsements, the Commission would have the burden of proving that the challenged conduct violates the FTC Act.

*FTC Publishes Final Guides Governing Endorsements, Testimonials, supra* note 1.

139 See 16 C.F.R. § 255.2 (2009). The FTC warns advertisers to take significant precautions should they pursue the route of a typicality disclaimer:

The Commission tested the communication of advertisements containing testimonials that clearly and prominently disclosed either
enforcement puts a heavy burden on the prospective use of this tactic. The FTC explicitly advises advertisers to maintain this evidence to avoid the risk of the initiation of an FTC action taken with respect to such an advertisement.\textsuperscript{140}

The Revised Guides aim to put a serious brake on the practice of using consumer endorsements, given evidence that consumers unduly raise their expectations about the salutary effects of a good or service after exposure to a peer’s true, but exceptional, experience with the offering.\textsuperscript{141} Consumer endorsements indeed would be costlier to employ after the revision, given the empirical compliance requirements, and from a regulatory standpoint, would certainly be riskier to use.

At first glance this regulatory approach may seem like a simple, straightforward, common-sense adjustment to the flow of commercial information aimed at consumers. The FTC is charged with protecting consumers from “unfair or deceptive acts or practices,”\textsuperscript{142} and clamping down on advertiser presentation of misleading material would certainly be within that authority. If presenting a consumer endorser’s true, but atypical, experience

\begin{quotation}
“Results not typical” or the stronger “These testimonials are based on the experiences of a few people and you are not likely to have similar results.” Neither disclosure adequately reduced the communication that the experiences depicted are generally representative. Based upon this research, the Commission believes that similar disclaimers regarding the limited applicability of an endorser’s experience to what consumers may generally expect to achieve are unlikely to be effective. Nonetheless, the Commission cannot rule out the possibility that a strong disclaimer of typicality could be effective in the context of a particular advertisement. Although the Commission would have the burden of proof in a law enforcement action, the Commission notes that an advertiser possessing reliable empirical testing demonstrating that the net impression of its advertisement with such a disclaimer is non-deceptive will avoid the risk of the initiation of such an action in the first instance.
\end{quotation}

\textit{Id.}\textsuperscript{140} \textit{Id.}\textsuperscript{140} Like the consumers in the Hastak & Mazis studies. \textit{See} HASTAK & MAZIS, WEIGHT LOSS, \textit{supra} note 3; HASTAK & MAZIS, DIETARY SUPPLEMENT, \textit{supra} note 3.\textsuperscript{141} Federal Trade Commission Act, 15 U.S.C.A. § 45(a) (West 2010).
effectively misleads the public, regulators may choose to debias consumers by drastically restricting or altering the practice.

Here, the FTC has taken a less-common tack in protecting consumers from receiving bits of information that are entirely true, but potentially so powerful in their truth that the dissemination of these bits need to be tightly-controlled. This “you can’t handle the truth” quirk in the regulatory approach opens the door to some serious questions, both in this instance and generally.

Part of the overall question involves determining the nature of the entirety of the commercial offering and what the consumer endorsement intends to convey. In addition to selling healthy sandwiches, is Subway selling the hope to consumers (not the promise) that if they transact regularly with Subway they will lose significant amounts of weight? Is hope an appropriate product to sell? Or is Subway merely trying to induce consumer transactions through cognitive error, using a strange, but true, story as aspirational bait?

Framed another way, if an internet dating service presents the testimonial of a married couple that met through use of its offering, is the dating service legitimately selling hope, attempting to induce transactions through cognitive error, or both? Another potential way of framing “hope” is to suggest that Subway and the dating service are selling an expected range of outcomes. Perhaps, presenting a product or service in a way that taps into that expected “range of outcomes” presents as an entirely different offering in

---

143 See generally Turner, supra note 105. Shielding truthful information to preserve a broader truthful narrative can be important in the jurisprudence of evidence, to purify jury decision making or protect other interests. See, e.g., Fed. R. Evid. 407 (subsequent remedial measures); Fed. R. Evid. 408 (compromise and offers to compromise); Fed. R. Evid. 409 (payment of medical and similar expenses); Fed. R. Evid. 411 (liability insurance). In other contexts, people shield the vulnerable from truth when the truth has little value and great detriment.

144 See A FEW GOOD MEN (Columbia Pictures 1992).

145 The Commission employed professionals to write two reports to study the effects of consumer endorsements and disclosures. While the reports showed that consumers might expect the demonstrated results in an advertisement to be atypical, the reports did not probe about the range of expectations for a set of individual consumers. See HASTAK & MAZIS, DIETARY SUPPLEMENTS, supra note 3; HASTAK & MAZIS, WEIGHT LOSS, supra note 3.
the marketplace. Consumers may know that experiences vary and may want to be aware of the positive possibilities. They may weigh their expectations optimistically—but is it the feeling of optimism that they are buying? Are these transactions a milder form of a lottery ticket, where the above-average expectation of winning may drive the purchase in spite of disclosed odds, but the thrill of the range of expectations is part of the product? Or is it a less transparent lottery ticket? If Powerball advertised lottery tickets as short-term “Grade D-” bonds instead of presenting the visual narrative of an ordinary person receiving an oversized check, would the perceived ex-ante value of a lottery offering be diminished? One would almost have to argue that consumers are buying hope, not a certainty, even if consumer expectations might be prone to optimistic distortion.

If the presentation of a wider range of possible outcomes from a transaction induces more transactions, some positive social welfare might be created, simply through exchange. Ex-ante hope, based on a truthful experience relayed by a peer consumer, might have innate welfare value, and ex post disappointment, though quite real, is something with which consumer transactions in general are rife. Regulators cannot fix all disappointments, however initially induced. A shift in regulatory policy that effectively quashes consumer endorsements may remove some distortions of expectation by smoothing out certain cognitive biases, but not at zero welfare cost. If this holds true, a brief assessment of the impact of the Revised Guides can illuminate the pitfalls of debiasing through law.

B. The New Regulatory Approach and Potential Impact

1. Previous Regulatory Approach

Under the Previous Guides, advertisers could present the

146 Gambling on lottery tickets can be destructive after a point, but the common practice has been to showcase winning jackpot tickets, which might have longer odds than finding a spouse on a dating website. For a vibrant view into lottery psychology and sociology, see LUCKY (Big Beach Films 2010).

147 See infra Part II.C.
Debiasing Advertising

experiences of consumer endorsers under limited conditions. The most basic requirement was that the consumer endorser, if represented directly or implicitly to be an actual consumer, must have been an actual consumer. Use of a consumer endorsement implies that the advertiser “represent[s] that the endorser’s experience is representative of what consumers will generally achieve with the advertised product in actual, albeit variable, conditions of use.” The Previous Guides required that advertisers possess substantiation that the advertised product’s result would perform at the level the consumer endorser enjoyed, or that “the advertisement . . . either clearly and conspicuously disclose[d] what the generally expected performance would be in the depicted circumstances.” The other option, commonly availed, is for the advertiser to “disclose the limited applicability of the endorser’s experience to what consumers may generally expect to achieve.” Quite often, this safe harbor would be employed by the advertiser simply slapping a disclaimer into the advertisement denoting something to the effect of, “Results not typical” or “Results may vary.”

2. New Approach

The Revised Guides make this safe-harbor option more impractical for advertisers. If advertisers use consumer endorsers, substantiation must be developed and possessed to support the typicality of the endorser’s experience, and if the endorser’s experience is atypical, the true typical experience must be disclosed. No longer can a marketer promote an endorser who had an outlier experience and cure it magically with a mild disclaimer.

At first blush, the Revised Guides may appear to advance consumer welfare. No longer will consumers be deceived by outlier results. But as this Article discusses, consumers may not always be “tricked” by these results. Consumers may know that they are buying the chance to achieve the consumer endorser’s result. That is, consumers may be buying the “hope” that they can

---

148 See, e.g., 16 C.F.R. § 255.2(b) (1980).
149 See, e.g., 16 C.F.R. § 255.2(a) (1980).
150 Id.
achieve the result—and this purchase may be rational. Consumers do retain protection without the Revised Guides. If an ad is deceptive or unfair, the FTC already has the power to address it.151

These recent changes effectively eliminate the consumer option to select hope as a product feature in this context. How does taking hope off the market advance consumer welfare? Does taking away a person’s hope that their acne might clear like a peer consumer’s did in the television ad advance consumer welfare? Does removing a lonely person’s hope that a dating service might bring them down the course toward marriage advance consumer welfare? Some consumer advocates imply that by showcasing a high-end result for consumers, this hope must almost always find its roots in and lead to deception.152

This guidance change effectively puts an end to a long-established practice that American consumers have become accustomed to seeing every day, and as a commercial cultural change alone, demands notice. Specifically, this could be the death of campaigns that use consumer endorsers to engage peers with optimism that use of a product or service might have an outlier outcome like that of a Jared Fogle.

When advertisers communicate through consumer endorsers, they are trading on the aspirations of the target consumer to achieve that result. “If that person got thin, toned their abdominal muscles, cleared up their skin, found a life partner and was able make $100,000 per year working from home, or simply get rich, why can’t I?”153 One might interpret the implicit dialogue between advertiser and consumer that way. Or it might be a mere conveyance of information that this is an achievable (but not certain) result from use of the product or service—and that with an associated disclaimer, a consumer should internalize that this true information presented to them presents a possible but not necessarily a probable result. Other scenarios may appear more benign. For example, consumer endorsers can be employed by local restaurants, owners of bed-and-breakfasts, and businesses

152 See Greenberg Statement, supra note 2.
153 Human appearance, mating, and money all seem to present ripe opportunities for the use of a consumer endorsement.
with lower accessibility to a brand. The consumer endorser can provide information to other consumers about the feel of the user experience.

One common thread woven through all of these consumer-endorsement scenarios is that they involve a marketer attempting to transmit information about a product through the experiences of an actual user. This information at its most pure and benign conveys a true experience and facilitates a transaction to happen with proper expectations on the part of the consumer. At its most nefarious, advertisers can use this approach to seize a true consumer experience and use it to create unrealistic expectations that induce a “bad,” welfare-destroying consumer transaction.

C. Benefits of the New Approach: Neutralizing Cognitive Biases

The Revised Guides should effectively eliminate the ability of marketers to exploit a number of cognitive biases to effectuate transactions. Part C explores some of the major biases that consumer endorsements play on and that the Revised Guides address. For consumers, generally, this represents a positive development standing alone, but as this Article explores in Part III, a downside must also be weighed.

1. Optimism Bias and Hope

Where hope is commoditized, cognitive psychology must be nearby. Where commerce meets cognitive psychology, behavioral law and economics cannot be avoided. These “Jared-style” consumer-endorsement campaigns exploit the power of a few widely accepted psychological phenomena absorbed by the behavioral law and economics literature, notably optimism bias and the availability heuristic. Optimism bias, famously epitomized by humorist Garrison Keillor as the “Lake Wobegon” syndrome,\textsuperscript{154}

\begin{flushright}
\textsuperscript{154} As broadcaster Garrison Keillor famously repeats about his blessed, fictional upper midwestern town, “Lake Wobegon, where all the women are strong, all the men are good-looking and all the children are above average.”
\end{flushright}
refers to the inclination of individuals to believe that there is an above-average chance of good things happening to them and a below-average chance of bad things. The optimism bias drives people to get out of bed in the morning and encourages risk-taking. One study demonstrated that half of consumers believed that their own households carried “below average” risk, with the other half believing that their households were average. Basic arithmetic tells us that result reflects undue optimism in sum.

People typically think that their chances of a range of bad outcomes, from having an auto accident to contracting a particular disease to getting fired from a job, are significantly lower than the average person’s chances of suffering these misfortunes—although, again, this cannot be true for everyone.


For an overview of the optimism bias, see David A. Armor & Shelley E. Taylor, When Predictions Fail: The Dilemma of Unrealistic Optimism, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 334–47 (Thomas Gilovich, Dale Griffin & Daniel Kahneman eds., 2002). This cognitive bias can be credited for getting businesses off the ground, despite high rates of start-up failure. See generally Arnold C. Cooper, William C. Dunkelberg & Carolyn Y. Woo, Entrepreneurs Perceived Chances for Success, 3 J. BUS. VENTURING 97 (1988).

Jolls & Sunstein, supra note 6, at 204; Wesley A. Magat & W. Kip Viscusi, INFORMATIONAL APPROACHES TO REGULATION 93–95 (1992).

Jolls & Sunstein, supra note 6, at 204. See also Christine Jolls, Behavioral Economics Analysis of Redistributive Legal Rules, 51 VAND. L. REV. 1653, 1659 (1998) [hereinafter Jolls, Behavioral Economics Analysis]. This phenomenon transcends personal judgments and translates even into professional judgments. Another recent study concluded that litigators are prone to routinely overestimate the success of the outcomes of their cases before the fact. See generally Jane Goodman-Delahunt, Pär Anders Granhag, Maria Hartwig & Elizabeth F. Loftus, Insightful or Wishful: Lawyers’ Ability to Predict Case Outcomes, 16 PSYCHOL. PUB. POL’Y & L. 133 (2010). Goodman-Delahunt et al. even employed a debiasing technique, exposure of the subject to what Jolls and Sunstein deemed “truthful narratives of harm,” to alter the thought process, as opposed to cold exposure to statistics, or overly-aggressive tactics that turn off reception. See Jolls & Sunstein, supra note 6, at 215. Goodman-Delahunt et al. asked some of their lawyer subjects “to generate reasons why they might not achieve their stated goals. This manipulation did not
Debiasing Advertising

Much of this view of the optimism bias can be framed as a relative one. Individuals measure their risk against that of other individuals. Some literature shows that individuals are overly optimistic (or not as pessimistic as they should be) in assessment of stand-alone event probabilities.\textsuperscript{158} Jolls and Sunstein summarize numerous studies:

reporting that professional financial experts consistently overestimate likely corporate earnings, while business school students overestimate their likely starting salary and the number of [job] offers they will receive.\textsuperscript{159} People also underestimate their own likelihood of being involved in an automobile accident, and their frequent failure to buy insurance for floods and earthquakes is consistent with the view that people are excessively optimistic.\textsuperscript{160}

The frame of the optimism bias tells us that consumers may be cognitively wired to view the stated experience and outcome of another consumer as one they can readily attain—even if that stated experience is at the upper boundary of what one could hope to realize from the advertised product or service.\textsuperscript{161}

Optimism bias can be applied to this problem in two different ways—first, there are some aspirations induced by advertisers that are not necessarily detrimental to consumers. Consumers need to have some level of optimism about the results of a transaction in order for the transaction to be induced. If the information is not unfair or deceptive, this transaction should yield welfare. Communication of the attainability of a result might be the primary purpose for advertising. Fogle transmitted the message that improve calibration [with the real outcomes].” Goodman-Delahunty et al., supra, at 133.

\textsuperscript{158} Jolls & Sunstein, supra note 6, at 204.
\textsuperscript{159} Id. at 205; Armor & Taylor, supra note 155, at 334–35.
\textsuperscript{160} Jolls & Sunstein, supra note 6, at 205; Jolls, Behavioral Economics Analysis, supra note 157, at 1660–61.
\textsuperscript{161} A brain-imaging study may have shed some light on why this might be. When considering the potential negative outcomes from a decision, the brain may not retain the memory of negative past events as well as it does positive events. See generally Tali Sharot, Alison M. Riccardi, Candace M. Raio & Elizabeth A. Phelps, Neural Mechanisms Mediating Optimism Bias, 450 Nature 102 (Nov. 1, 2007).
consumption of the Subway offering could lead to significant weight loss. According to Subway’s management, this message induced millions of additional consumer transactions, presumably creating welfare.

The second application of the optimism bias, which the FTC is at least implicitly concerned about, is the dangerous type of consumer endorsement claim that deliberately induces systemic cognitive errors. The FTC aspires to eliminate advertisements that yield so much unjustified optimism that they deceive the public and result in bad transactions that are ultimately damaging to welfare. If hope, or a range of outcomes, is on the market, welfare-seeking regulators should not want undue optimism to cause people to overvalue systematically the probability that they will receive the higher range of the outcome.

Even though this more nefarious use of the optimism bias likely yields cognitive errors that result in consumer disappointment, the absolute magnitude of the disappointment of the result after the fact compared with expectations before the fact is not well-understood.162

If the optimism bias implies that the hard-wired hope of above-average results will be routinely dashed as results regress to the mean, it becomes more important to consider hope—and lost hope. The ancient Greeks held hope out as a sacred value. After Pandora “unleashed the affliction, disease and death upon mankind”,163 from her box, only one thing remained, hope. The import of hope, and the concept that hope would be a cruel thing to dash, has deep roots. The ancients knew that hope had a special value as a force.164

Moving to the modern problem of false hope in the commercial context, there is a notion that true information can lead to false hope. Disappointment per se can be socially costly, which means that fostering a high level of hope, or hope for a higher-than-

162 The FTC-commissioned studies fail to provide this analysis.
164 For a brief discussion of hope and social trust, see GEORGE CVETKOVICH & TIMOTHY C. EARLE, SOCIAL TRUST: TOWARD A COSMOPOLITAN SOCIETY 69–75 (1995).
Debiasing Advertising

typical result, can be destructive. If a weight-loss company induces purchase of its program through hope for a certain result, and this result is not obtained, this may lead to lost consumer emotional investment in hope, as well as the more transparent lost financial investment. More than the financial investment would be lost; hope would be lost, too. This can put the consumer in a position where the utility loss exceeds what would normally be expected from squandered value. Kahneman and Tversky famously found that people are cognitively predisposed to assign a heavier value to a loss of X than to the value of a gain of X.\textsuperscript{165} Manufacturing and selling hope of an above-typical result and failing to deliver can diminish utility more dramatically than in a transaction where the typical result would be expected.

But again, it is important to note that traditional FTC discretion would be valuable in regulating the sale of hope via consumer endorsements. Certainly, the hope dashed from failing to achieve an above-typical result from an expensive weight-loss company program differs from that of the hope dashed from the Subway program, which consists of a series of continuing, non-obliging transactions. Yet, consumer endorsements for both are deterred by the new regulations.\textsuperscript{166} Why would we regulate hope the same way for two very different cases? Subway operates, interacts, markets, and transacts with consumers in a vastly different manner than weight-loss programs. The transactions are serial and smaller, and the program is simple. Weight-loss programs tend to require a longer contractual commitment and a deeper financial investment.\textsuperscript{167} The damage levels from the potential betrayal of hope from a Subway program are likely below that of a typical weight-loss program.

If hope is valuable, taking a sweeping approach to forbid firms from creating it and marketing it as an attribute of their product or

\textsuperscript{165} See Kahneman & Tversky, \textit{Prospect Theory}, supra note 125.

\textsuperscript{166} Subway may be less deterred than a small weight-loss company if it has more resources to defend the advertising in an action.

service should have a downside somewhere. Basic cognitive science behind the optimism bias and consumer human valuation of gains and losses can inform our understanding of the trade-offs of a sweeping change. Assessing the impact of taking hope off the market in this wholesale manner can lead us to a clearer picture of the regulatory impact.

2. Availability Bias

Working along with the optimism bias, and perhaps even a subject of greater concern about consumer endorsements, is the availability bias.\textsuperscript{168} Put simply, people tend to use the information most readily available to them as a heuristic for what the broader picture looks like. Consumers are deluged with data of all types in the marketplace and must find efficient shortcuts to navigate through them.

Behavioralists often present the availability bias through the example of people overestimating risk and taking disproportionate steps to mitigate or avoid such risk.\textsuperscript{169} Fear of domestic terrorism may provide one example of overestimation of risk, as the media coverage after a tragedy on the scale of the terrorist attacks of September 11, 2001, made terrorism risk more “available” to the public.\textsuperscript{170} Sunbathing may present a more serious risk to public health, but the information is not nearly as dramatically and graphically available.\textsuperscript{171} The public has more access to information about homicide through the media and less access to information about suicides, which leads to a commonly-held, but incorrect conclusion that homicide is more prevalent than suicide.\textsuperscript{172} That conclusion could foster the formation of misguided public-health attitudes.

\textsuperscript{170} Id. at 25.
\textsuperscript{171} Id.
\textsuperscript{172} Id.
Debiasing Advertising

Also undergirding the availability bias is “salience,” which strongly relates to consumer-endorsement regulation. A salient experience will prove much more available when assessing the probability that an event or result will occur. For example, directly experiencing an earthquake will cause a person to have a more accessible data point when assessing the risks of experiencing a future earthquake. “Vivid and easily imagined” risks like dying in a plane crash are over-weighted in comparison to other risks like death by asphyxiation from choking on food. The consumer endorsement is employed to make a positive outcome both the most accessible piece of information and the most vivid one.

If presented with a strong consumer endorsement, the availability bias predicts that the consumer will heed that information because the advertiser has centered it right in front of them. The question still remains whether they will discount the more spectacular claims made by consumer endorsers as a possible result, rather than a probable result (on a spectrum of results). Much of this depends on the total pool of information that is available and competing with the consumer’s attention in a given environment.

Where consumer endorsements are the centerpiece of an advertisement, they become the available focus of the consumer. The Revised Guides remove consumer endorsements from that centerpiece, potentially neutralizing errors attributable to the availability bias.

3. Framing Effect

Prospect theory, promoted by Kahneman and Tversky, contends that humans evaluate the impact of decisions based on “an initial reference point, rather than on the basis of the nature of the outcome itself.” Tied closely into prospect theory is the

---

173 Id.
174 Id.
175 Id.
176 See generally Daniel Kahneman & Amos Tversky, Prospect Theory, supra note 125.
177 Jolls & Sunstein, supra note 6, at 205.
endowment effect and loss aversion. These theories and observations reinforce the notion that people value losses heavier than gains.178

From a practical perspective, the related “framing effect” demonstrates that perceptions and decision processes can be manipulated by framing certain outcomes as losses and others as gains.179 Advertisers and policymakers put the framing effect to work constantly.

Jolls and Sunstein identified a fairly recent controversy that emerged in the public-health arena involving framing and the use of advertising to promote behavioral changes.180 Breastfeeding advocates tangled with the baby formula industry over a public-service advertising campaign that graphically emphasized the risks of choosing not to breastfeed.181 The risks presented in the campaign included a higher incidence of leukemia and childhood asthma.182 One ad supported by the activists presented images of pregnant women roller-skating with a voiceover: “You’d never take risks while you’re pregnant. Why start when the baby is born?”183 Representatives from the baby-formula industry blanched at this tactic and urged that the campaign present a message that emphasized the positive benefits of breastfeeding, rather than a parade of potentially terrible consequences.184 Indeed, the breastfeeding advocates were trying to emphasize downside risks (losses), which people tend to value more highly, while the baby formula industry, in pushing for a softer touch, wished to emphasize the benefits, which people tend to value lower.185

178 This path of thinking, constructed around bounded rationality, creates outcomes that are different from what one would anticipate from traditional expected utility theory. Id.
180 Jolls & Sunstein, supra note 6, at 206.
182 Id.
183 Id.
184 Id.
185 Jolls & Sunstein, supra note 6, at 206.
Debiasing Advertising

two interest groups were sparring over the framing of the consequences of the decision to breastfeed.

With the Revised Guides, the FTC has implicitly recognized that “truthful narratives” are quite powerful in this consumer endorsement context as well.\textsuperscript{186} Truthful narratives, like the one deployed by the breast-feeding advocates, perform a debiasing function, a non-intrusive means of guiding choices, but debiasing advocates generally think of these narratives as a positive interventional tool.\textsuperscript{187} With consumer endorsements, a different view of the “truthful narrative” is presented—one where the truthful narrative could be harmful to the consumer. In fact, this view even offers a few commissioned studies to attempt to prove the harm.\textsuperscript{188}

The truthful narrative in this instance communicates a remarkable story to consumers about the results that one of their peers has achieved from using a product or service. Jared Fogle told the world his story. His narrative (Subway’s narrative) was about how an obese but otherwise ordinary young man from Indiana achieved weight loss (in what appears to be a less painful way than conventional dieting) with the disciplined consumption of savory Subway products.

From a debiasing perspective, removing information from the marketplace and specifically, information packaged in this narrative form, demonstrates the power of a narrative in the negative. The narrative produces the framing in a vivid way, and the new regulation silences that narrative and erases the outlier reference point.

With the Revised Guides, at least in the first round of the game, no choices have been paternalistically removed from the consumer’s menu. All that has been removed is a tactic that regulators purport misleads consumers into making cognitive errors, engaging in “bad” transactions by latching on to compelling narratives. As this Article will explore further in Part III, a broad

\begin{itemize}
\item \textsuperscript{186} Id. at 215.
\item \textsuperscript{187} See generally Jolls & Sunstein, supra note 6; Sunstein & Thaler, Nudge, supra note 7.
\item \textsuperscript{188} See Hastak & Mazis, Dietary Supplements, supra note 3; Hastak & Mazis, Weight Loss, supra note 3.
\end{itemize}
...stroke such as this one that removes this information might constitute “overshooting,” to use the parlance of Jolls and Sunstein. Removal of consumer endorsements might prevent “good,” social-welfare-accreting transactions from occurring. Additionally, the consumption of some goods and services yields value through the actual risk of the consumption experience per se. There might be something valuable to knowing that a dating service merely could lead to a match ending in marriage, for example. Risk is purchased everywhere. Consumers have ranges of expected outcomes for their experiences and quite often they are purchasing this range.

The broad-stroke regulatory deterrence of consumer endorsements might at best have an ambiguous impact. Framing is a tool that marketers may use to induce transactions, and certainly consumer endorsements pin the consumer around the reference point of the endorser. It does not logically follow and there is no available proof, however, that in the aggregate, framing tactics by advertisers in this manner reduces social welfare. Debiasing this particular effect is of unproven social value and unknown social risk.

4. Representativeness Bias

Representativeness bias presents another cognitive quirk that the Revised Guides would seem to neutralize. People have great difficulty comprehending that certain events that may seem to occur in a cluster are actually occurring randomly. Classic examples of this are studies of the misperception of “hot hands” in

---

189 See Jolls & Sunstein, supra note 6, at 230–31.

190 The Hastak and Mazis studies were narrowly limited to three offerings, weight-loss programs, dietary supplements, and money-making opportunities. HASTAK & MAZIS, WEIGHT LOSS, supra note 3; HASTAK & MAZIS, DIETARY SUPPLEMENTS, supra note 3.


192 See SUNSTEIN & THALER, NUDGE, supra note 7, at 27.
Debiasing Advertising

basketball free throws, the misperception of precision targeting of bombings during air-raid campaigns, and from epidemiology, the misperception of the formation of local cancer clusters. The human brain subconsciously makes sense of a random sample through a larger extrapolation that is part of a greater scheme. Much of stereotyping is the product of this phenomenon.

If consumer endorsers present the real, but exceptional, results that they have enjoyed with an offering, the recipients of the message may be prone to the distortions of the representativeness bias. The real, but outlier, results may be the genuine product of randomness. Presented alone, consumers may extrapolate from them the conclusion that they are more typical of the set of experiences that people will have with the offering. By chilling consumer endorsements, the Revised Guides make it difficult for advertisers to highlight the outliers, which may neutralize this bias from working in this circumstance.

5. Confirmation Bias

The confirmation bias describes the natural tendency of people to seek and to accept information that confirms previous perceptions and to eschew other information.

If one were to attempt to identify a single problematic aspect of human reasoning that deserves attention above all others, the confirmation bias would have to be among the candidates for consideration. Many have written about this bias, and it appears to be sufficiently strong and pervasive that one is led to wonder whether the bias, by itself, might

---


195 SUNSTEIN & THALER, NUDGE, supra note 7, at 31. See also Cancer Clusters: CDC Activities: Fact Sheet, CENTERS FOR DISEASE CONTROL AND PREVENTION, http://www.cdc.gov/nceh/clusters/factsheet.htm (last visited July 30, 2010).
account for a significant fraction of the disputes,
altercations, and misunderstandings that occur among
individuals, groups, and nations.196

Wishful thinking pervades human decision making, and hope
springs as eternal in the consumer transactional context as it does
anywhere else. Cognitively, humans are predisposed to seek
information that confirms or reinforces preconceived notions about
their worldview.197

Consumers seeking a desirable result may grasp on to that
result as the likely outcome from a certain course of action, simply
because it might confirm a belief that there is a cheap, easy means
to achieving a result. If a person wishes to lose weight, and an
advertiser presents him with a narrative about an easy system for
losing weight, along with an extremely positive actual outcome,
confirmation bias should influence the person. The information
that confirms the result that the weight-loss seeker wishes to be
true will be accorded more heft.

Put differently, people have demonstrated that they will
maintain a high evidentiary standard for acceptance of ideas or
conclusions that they have a predisposition to dislike and a lower
standard of proof for ideas or conclusions that they prefer. The
revised approach toward consumer endorsements removes from
the information sphere one of the many tools that advertisers can
use to trigger the confirmation bias. Simply showing through a
consumer endorsement that an offering might solve a problem or
meet a need can induce a transaction—if the consumer falls prone
to seeking confirmation that there is a way to solve a certain
problem or meet a certain need.

Generally, debiasing through law, or in this case, through
advertising regulation, can disarm a host of potentially distortive
cognitive triggers, including but not limited to the ones mentioned
here. However benign the regulatory intention, some negative
spillover effects can result from this disarming, which this Article
addresses in Part III. Any drag on the flow of information can
change how players act in the marketplace. With consumer

196 Raymond S. Nickerson, Confirmation Bias: An Ubiquitous Phenomenon in Many Guises, 2 Rev. Gen. Psychol. 175, 175 (1998).
197 See generally id. (exploring in detail the entirety of confirmation bias).
Debiasing Advertising

endorsements, a seemingly well-intended policy change may present some other side effects. These side effects may well be tolerable, just like a medication’s side effects. But just as with medication, awareness of these side effects is necessary for evaluating whether the prescription is appropriate.

III. POTENTIAL COSTS AND FAILURES OF THE NEW APPROACH

As noted, the removal of true information from the marketplace provides an illustration of how debiasing can lead to unanticipated results, and hints that debiasing techniques should be targeted narrowly and cautiously. Having noted the salutary effects of the consumer-endorsement debiasing effort, this section will explore three primary areas of concern, all of which have a common thread.

First, consumer endorsements effectively remove the presentation of “hope” in the marketplace and some offerings are hope. Removing “hope” or any offering from the marketplace runs the risk of reducing social welfare and removing choice from the marketplace. Closely related, the hard paternalism that debiasing through law tries to avoid rears its head if the elimination of consumer choice proves to be the practical result.

When the commercial information flow is chilled, particularly the flow of true information, welfare-creating, Pareto-efficient transactions that may have occurred, might fail to occur. Forcing an offering entirely out of the marketplace can create such a loss, but so can consumer uncertainty caused by a restricted flow of data. The cost of transactions lost in this manner must be considered.

Another potential effect that may result from the restriction of information in the market may be higher prices and deadweight loss. Although the impact from chilling consumer endorsements is uncertain, other restrictions in consumer markets, like a ban on disclosing price of alcoholic beverages away from the point of sale, have resulted in higher consumer prices and market disruption. Debiasing through law, toying with the information flow, and trying to improve consumer decision-making processes
can have secondary impacts of this sort. Again, this should guide the policymaker toward more cautious approaches when deploying debiasing strategies. “Nudging” requires a soft touch.

Finally, debiasing is not well equipped to address several related normative problems. Defining “good” decision making and defining “good” social outcomes will be matters for policy debate that transcend the debiasing through law toolkit. Nonetheless, to create policy, paternalistic judgments must be made about what is “good.”

With consumer endorsements, a view must be formed about the presentation and promotion of risk. Risk is naturally associated with a downside or potential losses. But risk can also embody something positive. As noted earlier, one person’s risk is another person’s hope, aspiration, or thrill. Policymakers still must make a potentially paternalistic judgment call on this dimension.

A. Removing “Hope” from the Market

The tighter regulation of consumer endorsements may remove a set of hope-based offerings from the market. The core offering may remain on the market, but without the consumer endorsement attached, the “hope” component of the offering may need to be eliminated. With a new, looming threat of civil enforcement, the costs and risks of using the consumer-endorsement marketing tactic make the approach riskier and costlier for an advertiser to deploy. This can destroy welfare if the underlying potential transaction would create welfare.

An illustration distills this point. Consider the example of a dating service. A dating service may use consumer endorsements to emphasize the number of successful matches the service has facilitated—and perhaps even the number of marriages that it has enabled. This could be effectively done through the presentation of a beautiful (or accessible) newlywed couple describing the narrative of how the site enabled their marital bliss.

This sort of presentation would deliver to the consumer a compelling narrative about the uppermost boundary of the range of outcomes sought by someone using this service to find a life partner. The range of results comprises this offering—ranging
Debiasing Advertising

from mere entertainment from browsing the site to an ultimate goal (finding a date or mate). Part of the offering might be the excitement about having access to this range of outcomes. Promoting the boundaries highlights this excitement and promotes hope. Hope is part of this product. If regulation removes hope from the market, the market transaction for that hope obviously cannot occur. Less “hope” will be produced and consumed, and less welfare accordingly created.

All offerings present a range of expected outcomes from consumption or use of the offering. Some ranges are tighter than others. An offering comprised of a can of a strongly-branded popular beverage would present a very tight range of expected outcome for consumers. A new four-star restaurant might present a much broader range of expected outcomes—and this might be part of the draw for potential consumers. The thrill of the risk of trying something novel and the hope that it will surpass expectations might be part of the restaurant experience. With a chain restaurant, consumers are not purchasing the thrill of a broad range, they are purchasing a comparatively narrow range of outcome. Consumers seek more certainty rather than hopes or thrills in this instance.

If an offering in this circumstance was entirely centered on hope, and depended heavily on consumer endorsements to promote that hope, the entirety of the offering may disappear. This might be a positive social development, if the promoted offering is of dubious value, like a “make-money-from-home” system that offers outlier testimonials. If the offering retains value and transactions can yield social welfare, the effective elimination of the offering would prove socially costly. A conceivable example might lie in an effective weight-loss program that can only sell at scale through the use of consumer-endorsement marketing tactics.

Taking hope “off the market”—with hope embedded in part of an offering or the entirety of an offering—will eliminate some welfare-producing transactions. Chilling the practice of describing hope through the narrative of a consumer endorsement can lead to that result. Broadly debiasing through law can similarly yield to results of concern. Addressing this particular problem in a more targeted way may prove more socially valuable, though the regulatory enforcement effort may be more complex and
challenging. This type of regulatory brain surgery must be performed with care because the consequences can prove unpredictable.

The consumer-endorsement restrictions also put downward pressure on social welfare simply by removing information from the marketplace. As the next section explains, when true information is removed from the marketplace, uncertainty looms, parties are less likely to transact, and transactions fail.

B. The Cost of Failed Transactions

Any regulatory intervention that removes true information from the marketplace should anticipate that the market will put a damper on the communications that enable transactions to happen. Debiasing through law may seem more nuanced, but tampering with information in a broad way to improve consumer decision making may have repercussions in the marketplace.

In the basic advertising scenario, banning deceptive or misleading information from the marketplace should prevent resource misallocation. Before an advertisement-influenced transaction, consumers would be willing to transact and part with something of value for something else to which they assign greater subjective value. If an advertisement conveys misleading or false information or makes misrepresentations, this false conveyance tampers with the consumer’s pre-transaction subjective valuation. If the transaction happens, the consumer will receive less than what was expected. This leads to a resource

---

198 Adam Smith wrote:

Whoever offers to another a bargain of any kind, proposes to do this: Give me that which I want, and you shall have this which you want, is the meaning of every such offer; and it is in this manner that we obtain from one another the far greater part of those good offices which we stand in need of. It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages.

Debiasing Advertising

misallocation, deadweight loss, diminished social welfare, and unjust outcomes.

A regime with no regulation of advertising, that of pure caveat emptor or lax enforcement of current laws and regulations, would prove economically chaotic. Transactions costs would rise, as consumers would need to perform more due diligence on those with whom they transact.\(^{199}\) Brand identity and professional licensing would matter more—even more than they do currently—in markets as consumers would seek heuristics for counterparty trustworthiness.\(^{200}\) In an environment of increased uncertainty about the validity of claims, one would expect consumers to change their purchase equation. Consumers would not be willing to pay as much for an offering made with uncertain representations. The uncertain representations would translate into a lower expected subjective value by the consumer for the offering.

Generally, the transactions costs would rise because of the consumer uncertainty—and because of the extra information gathering required to build trust in the offering. The current system of policing advertising aims to reduce these transactions costs by purifying information. The FTC, the states, other advertising regulators like the Food and Drug Administration (FDA), and private enforcement mechanisms\(^{201}\) work together toward this goal—reducing social costs by addressing dishonesty.

George Akerlof’s renowned article, *The Market for*
“Lemons,”202 presented a “structure . . . for determining the economic costs of dishonesty.”203 Akerlof held out the individual used car market204 as an example of a market where transactions fail due to perceived information asymmetry. Sellers have the best information about the true value of their own used cars. The sellers know more about the true condition, history, and performance of their own vehicles. Buyers do not have access to that information and therefore will discount the price that they will be willing to pay for the used car, because they do not have complete trust in sellers. Sellers expect the “true” value for the car, but buyers, in their uncertainty, won’t pay that value. Often, this gap yields a transactional failure, as the parties cannot reach agreement on price.205

Given that wealth and utility are created when parties successfully exchange goods, this “lemon” phenomenon points toward a common scenario where an information gap obstructs welfare from being created. When applied to advertising regulation, Akerlof’s findings have multiple implications. Eradicating deception from advertising leads to economic certainty, efficiency, and just results. Goods and services are exchanged at their stated qualities and subjective expectations from the exchange are met. Further, confidence and trust in the system at large is restored, furthering an environment hospitable for exchange.

However, as would be the case with most regulatory activity, overzealous or overbroad rulemaking and enforcement in combat
of deceptive advertising would potentially produce deleterious effects just as economically damaging as under-enforcement. When providers of products and services are compelled to disclose less information—or banned from disclosing true information—because of the risk that they may run afoul of new and aggressive regulation, consumer uncertainty about the nature of the offering rises. When consumers know less about an offering, they may be less likely to transact. Consumers would adjust downward the value that they are willing to exchange for the offering. Producers would supply less as a result of a diminished expected return. Deadweight loss would increase.

The chilling of information exchange reduces the ability of the buying and selling parties to communicate at a distance about the offering, thereby reducing the potential for exchange. If the information is truly false and deceives, the exchange should not take place. However, the risk of punishment for stepping over the line may prevent advertisers from treading anywhere near the acceptable line to ensure compliance. Larger, established entities may preserve their competitive position at the expense of smaller entities in this environment. Those with established reputations should prevail under the notion of trust.206 Even though debiasing through law may appear to be relatively unobtrusive and oriented toward improving individual decision making, in this particular instance the secondary effects may hit the market more broadly, leaving the welfare impact open to question. Removing true information can remove a true avenue of valuable commercial communication.

C. Unintended Consequences: Price Effects

Debiasing through law, advertising regulation, and specifically, the removal of true information from the marketplace, can have an upside in correcting consumer biases. Potential downsides lurk,

206 Note how entities tend to disclose when they were established as a means of presenting durability and trustworthiness. For instance, the author’s law school’s website advertises that the institution was established in 1883. COLLEGE OF LAW – WILLAMETTE UNIVERSITY, http://www.willamette.edu/wucl/ (last visited Jan. 25, 2011).
however, with debiasing, just as they would with traditional paternalistic regulation that directly eliminates consumer choices.

Here, this Article briefly discusses the subtle price distortions that could result from the implementation of the Revised Guides. Enhanced regulation of advertising could simply lead to less advertising because of the increased cost of advertising \textit{per se}. If consumer endorsements are the preferred route to marketing an offering, eliminating that tactic will force the advertiser to use a second-best approach. This second-best approach will likely be less effective, all things being equal, and therefore, more expensive. If all marketers had to play by the same rules, consumers may bear some of this burden. Generally, “if... [increased] advertising regulation also succeeded in reducing the extent of advertising by raising the cost of advertising, firms in many industries might have an incentive to seek such regulation because it would result in less competition and higher retail prices.”\textsuperscript{207} The resulting decrease in social welfare and increase in deadweight loss could provide yet another example of an unanticipated secondary effect from debiasing.

Further, economists have observed in several studies that “the prices of goods and services in places that restrict advertising tend to be higher than those in places that do not restrict advertising.”\textsuperscript{208} These linkage studies between advertising and pricing have ranged across consumer goods, from eyeglasses and drugs to alcohol.\textsuperscript{209} Restrictions on information in the marketplace “often tend to inhibit competition, with consequent efficiency losses.”\textsuperscript{210}

Though it may be difficult to predict what will happen to retail

\textsuperscript{208} Id. at 255.
\textsuperscript{210} Beales et al., \textit{supra} note 46, at 514.
Debiasing Advertising

pricing in the industries heavily affected by the Revised Guides, an empirical trail heightens suspicion that the restriction of information may yield unanticipated changes. The new consumer-endorsement approach does not restrict display of prices but it does effectively restrict a major form of information from reaching the marketplace. Debiasing may seem non-intrusive and “soft” but the impact can be unpredictable and not always positive for social welfare.

D. The Stubborn Normative Question

Advocates of debiasing through law claim the approach to be comparatively unintrusive to traditional regulation, less paternalistic, and protective of ultimate consumer choice. To have meaning, these expressed values must be colored by the real goals of any debiasing effort, which by and large, appear to be about minimizing risk or tightening the range of expected transactional outcomes. Generalizing this approach across the broad range of commercial transactions can be normatively challenging—the risks involved with different types of transactions are situationally unique.

Is it always correct to minimize risk through disclosure or restriction of information? As noted, somewhere in every consumer transaction, hope itself is for sale. In routine transactions, the result rarely diverges from the expected range. In some transactions (tickets to a new Broadway mystery musical, for example), the range of expected results might be part of the sought experience. Minimizing risk and closing the range of expectations through a mandatory disclosure about the mystery ending of the mystery musical may prove socially destructive, though it would in some sense protect consumers from seeing something unwanted.

As modern American political strategists are deftly aware, if “hope” can be credibly attached to a political purchase, voters might just buy it.211 Do voters expect that all of their hopes will

211 Governor Bill Clinton of Arkansas closed his 1992 Democratic Presidential Nomination speech by declaring (in a wrap-around reference to his birthplace of Hope, Arkansas) that he “still believe[d] in a place called Hope.” Bill Clinton, Governor, Acceptance Speech to the Democratic National
truly be realized? Lotteries market themselves this way.212 In the gambling and lottery instances, rational consumers would know that “the house” will win, on average, but they are often buying the thrill.

Sometimes consumers contract for hope and are willing to accept a certain amount of disappointment. They are expecting to purchase a set of probabilities. This applies to staying at unfamiliar hotels, buying a new lawnmower, or buying an article of clothing. Consumers may know that goods and services are not perfect but the information that it worked for someone like them gives them the hope to try it. What is normatively wrong with offering that range of expectations? Advocates of debiasing still must wrestle with the questions that old-fashioned regulators must answer—namely, which risks are acceptable to expose to consumers and which are not? Debiasing addresses this question by intervening in the decision-making process and paternalistic regulation addresses this question by directly removing the risky choices from the marketplace.

This is not to say that regulators should surrender in their efforts to match appropriate venues and approaches for managing consumer risk. Consumers demonstrate emotional discomfort with the risks associated with certain transactions, particularly high-stakes transactions. This can diminish the willingness to transact.213 As noted through the implications of Akerlof’s work, and in the discussion of social welfare, promoting Pareto-superior transactions should be a policy goal. Policymakers need to sort through these risk-involved scenarios on a granular level but must be careful not to develop a sweeping view that seeks to mitigate all risk—because not all risks are the same.


Debiasing Advertising

IV. RECOMMENDATION TO POLICYMAKERS

Policymakers must be selective and careful when they intervene and engage in debiasing through law. Advertising provides the perfect backdrop for assessment of debiasing techniques.

The new approach to consumer-endorsement regulation, as put forth in the Revised Guides, provides an enlightening case study. A sweeping approach toward concealing true information about consumer experiences could have a different impact across different consumer markets. Consumer-endorsement regulation, though controlling a narrow information tactic, could be tightened to address certain industries more carefully.

Industry-targeted regulation of information presentation might be most sharply demonstrated with the pharmaceuticals markets. For example, the FDA accorded special attention to direct-to-consumer advertising of pharmaceuticals, as spending on this form of advertising in this market grew from $12 million in 1989 to $2.38 billion in 2001.214 A broad array of guidelines and rules were adopted prior to and during this period, specific to direct-to-consumer pharmaceuticals advertising because of the unique risks presented by the products in this marketplace.215 The products are complex, confusing, and present large consequences resulting from the user experience. Information about the products may be entering from a variety of authoritative avenues, like physicians and pharmacists. This is a market where we would want risk to be minimized by enabling people to use the products in an informed manner. Finding clear answers proves difficult. But comparatively, pharmaceuticals may prove to be an area more worthy of debiasing attention than the sandwiches Jared Fogle promoted for Subway.

Consumer-endorsement regulation shows us that some industries require more attention than others and that overregulation of consumer endorsements in certain industries can

215 See generally id.
be welfare-destroying. This provides a lesson more broadly for regulators who intend to use the new tools highlighted in the popular book, *Nudge*, and who intend to apply the concepts of debiasing. Selectively deploying debiasing strategies proves key. This type of informational intervention should be performed with caution and elegance, given the intersections between the uncertainties of markets and human behavior.

V. CONCLUSION

Debiasing through law can be a powerful and creative approach toward regulation and increasing consumer welfare, but application requires care. The revised FTC approach to consumer-endorsement regulation provides a cautionary tale about a seemingly benign regulation that may ultimately prove to sweep too broadly. If regulators had attempted to pursue specific industries or markets, rather than chilling an entire tactic, total welfare—and consumer welfare—could potentially have been further enhanced.

Debiasing approximates much of the law of advertising regulation, as Jolls and Sunstein note. The regulation, as reflected in the Revised Guides, illustrates an approach where debiasing could have been effected more artfully, either through an industry or product-oriented approach. Though the FTC seemed to be

---

216 Treating consumer endorsements monolithically ignores the notion that some consumer endorsements may be more deceptive than others. Craswell “has suggested that the laws against deceptive advertising could best be understood as applying a form of cost-benefit analysis to the interpretation of advertisements.” Richard Craswell, *Regulating Deceptive Advertising: The Role of Cost-Benefit Analysis*, 64 S. CAL. L. REV. 549, 550 (1985). Craswell advocates applying a case-specific approach toward advertising regulation, inquiring into the “total injury” that an advertisement might cause, and whether the advertiser could have taken affirmative steps to prevent the injury. Craswell describes his approach as similar to the “Learned Hand test.” See generally id.

217 SUNSTEIN & THALER, NUDGE, supra note 7.

218 Fortunately, the FTC has retained enforcement discretion, but unfortunately, the warning shot from the Revised Guides was broadcast without discrimination.

219 The Revised Guides are still very new. This Article can only speculate about the future potential impact with the tools available.
Debiasing Advertising

aiming to eliminate or heavily burden a tactic, in an effort to improve consumer decision making, it effectively may have taken products off the market by “overshooting.” Consumer welfare in this instance could have been maximized with a more careful approach, one driven by typical cases, not the worst cases like the weight-loss industry.

This Article can draw several conclusions from the revised consumer-endorsement regulation that can be applied more generally to the underlying principles of debiasing through law. The revised consumer-endorsement regulations will likely increase the cost of inducing transactions in the arenas where this advertising is most commonly used, reducing welfare in some markets. In industries or markets that require a bit more protection for the consumer, like pharmaceuticals, the restriction of true information provided by real consumers about their experience might be welfare-enhancing. If laid down broadly, regulations such as these could have a mixed impact on social welfare, depending on how the regulation affects the variety of transactions that lead consumers to expect a range of results.

The key lesson for policymakers from the consumer endorsements example is that great care should be taken when tampering with the market information flow. Debiasing through law focuses on using information to enhance decision-making, which can certainly happen. But the potential downsides and side-effects are as real as the downsides of traditional, paternalistic regulation, and should not be discounted.
ACCOUNT ME IN: AGENCIES IN QUEST OF ACCOUNTABILITY

Dorit Rubinstein Reiss*

INTRODUCTION

Not long after taking office, Lisa Jackson, appointed by President Obama as Administrator of the Environmental Protection Agency (“EPA”) revamped the evaluation procedure for the EPA’s Integrated Risk Information System (“IRIS”) database.1 The database was an initiative undertaken by the EPA to improve the

* Associate Professor of Law, UC Hastings College of the Law. I would like to thank Oded Na’aman for his very useful suggestions; Heather Field for sharing her immense expertise about the IRS; also, Marsha Cohen for her aid in conducting the research and Ashutosh Bhagwat, Marsha Cohen, David Coolidge, Vibeke Lehmann Nielsen, Elizabeth Magill, Jerry Mashaw, Anne Joseph O’Connell, Reuel Schiller, and Glen Staszewski for their helpful comments and suggestions on previous drafts. I am very grateful to Peter Barton Hutt, Richard Merrill, Deborah Wolf, and other interviewees to whom I promised confidentiality (and therefore cannot name them) for generously sharing their stories and their valuable time with me. Finally, I wish to thank Fatemeh Shahangian for her excellent research assistance. All errors are, of course, my own.

1 The EPA’s website explains that:

The Integrated Risk Information System (IRIS) is a human health assessment program that evaluates quantitative and qualitative risk information on effects that may result from exposure to environmental contaminants. IRIS was initially developed for EPA staff in response to a growing demand for consistent information on substances for use in risk assessments, decision-making, and regulatory activities. The information in IRIS is intended for those without extensive training in toxicology, but with some knowledge of health sciences.

scientific accuracy and the quality of information available on risks associated with certain chemicals; it is used, for example, when the EPA determines whether to establish air and water quality standards regulating certain chemicals.\textsuperscript{2} IRIS assessments are neither rules nor adjudications; rather, they are background materials later used in making rules. Therefore, they are invisible to the Administrative Procedure Act ("APA").\textsuperscript{3} The procedures surrounding the IRIS database—as with everything related to the project—were designed by the EPA and are not mandated by law.\textsuperscript{4}

In spite of this potential freedom to choose any or no accountability mechanisms, the procedure the EPA adopted for IRIS—from the start of the system—including extensive steps of review, and numerous opportunities for input and checks by external actors.\textsuperscript{5} The additional procedures put in place by Administrator Jackson aimed at achieving a process that is “more transparent and timely, and . . . will ensure the highest level of scientific integrity.”\textsuperscript{6}

The adopted process exposed the EPA’s suggestions to external peer review, in addition to notice and comment, by using the same process required for informal rulemaking under the APA;\textsuperscript{7} it also subjected these suggestions to review by the Office of Information and Regulatory Affairs ("OIRA") in the White House’s Office of Management and Budget ("OMB"), which is required for significant rules.\textsuperscript{8} Most of these elements were part of the IRIS


\textsuperscript{3} Administrative Procedure Act, 5 U.S.C.A. §§ 551–559 (West 2010).

\textsuperscript{4} U.S. Gov’t Accountability Office, supra note 2, at 10.

\textsuperscript{5} See infra Part III.B (describing these procedures in detail).


\textsuperscript{7} See Administrative Procedure Act, 5 U.S.C.A. § 553 (West 2010).

system from the start; in one form or another, the agency voluntarily chose to expose its decisions to extensive scrutiny from various outside parties.

Similarly, as part of the reform of the Internal Revenue Service ("IRS") after 1998, Commissioner Rossotti and the IRS staff worked hard to increase the agency’s responsiveness and transparency. The agency put substantial efforts into making the Internal Revenue Manual more accessible, as discussed more extensively in Part III.C. It invested in improving customer service and, consequently, the number of calls answered rose dramatically, and the quality of the IRS’ response received very positive reviews. Improvement in service was required under the 1998 legislation reforming the IRS. However, Congress had previously passed other reforms requiring improvements in customer service—for example taxpayer rights provisions in the 1980s—but without sincere agency commitment and agency initiated efforts, those had limited effect. This time, the agency was committed to improving customer service and put in place changes increasing transparency and responsiveness.

Finally, before the courts started requiring that agencies answer each comment submitted to them, and before agencies took the spirit of the Freedom of Information Act seriously, the FDA, under the direction of its Chief Counsel, Peter Baron Hutt, adopted procedures that involved responding to comments submitted to it and an approach to transparency that involved making as much

---


information as possible publicly available.\footnote{See infra Part III.A.}

Often, we assume that agencies are the villains in the “accountability game,” the quest to be accountable. Much of the literature about accountability sees agencies as obstacles. It is often taken for granted that agencies will avoid accountability as much as they can, and that pressure to accept accountability will be required as a matter of course. Other scholars, in response, emphasize the multiple and conflicting pressures for accountability placed on agencies, and see them as victims of too much accountability. This approach sees agencies as merely passive actors in this area, subjected to accountability mechanisms against their will and with no real control or influence on their accountability environment.

While there is much truth to both perspectives, they each miss an important part of the picture. As the examples above suggest, agencies are not always the enemies of accountability. Nor are they always helpless, passive pawns, crushed under the oppressive weight of accountability. Agencies can also be autonomous and important actors in the accountability game, creating new forms of accountability, or accepting and adapting pre-existing forms. They often willingly join in and strive to be accountable. They may well invest substantial efforts in increasing their accountability.

Not all agencies do this all the time, and not all agencies do it well. But in today’s administrative environment, agencies need accountability, and being sophisticated actors, they work at achieving it. Their efforts happen for a number of reasons—internal and external—and not just because of cost/benefit considerations.

This Article examines such actions by agencies and addresses the reasons they take the actions they do. Following this introduction, the article proceeds in four parts. Part I provides the background, reviewing current literature and demonstrating the tendency to place agencies in either the villain or the victim camp, as well as discussing the very few studies that focus on agencies’ own contribution to the accountability regime surrounding them. Part II reiterates that agencies seek accountability and addresses
possible explanations for such behavior. The explanations highlighted are rational choice (agencies seek to be accountable because the costs of not being accountable are too high), power of ideas, and internalization of the idea that accountability is part of the administrative agents’ mission. Part III provides a small number of more detailed case studies of agencies which sought to increase their accountability, and includes examples from the FDA, the IRS, and the EPA. This discussion shows that the phenomenon is a real and common one, and provides empirical support for the explanations addressed in Part II for this behavior. Finally, Part IV discusses implications for theory and practice of accountability in the administrative state.

I. BACKGROUND: ACCOUNTABILITY IN THE ADMINISTRATIVE STATE

Accountability of administrative agencies is an ongoing concern in the administrative state. Agencies exercise tremendous power and engage in numerous activities.¹³ Not surprisingly, controlling them has been a constant preoccupation of scholars, politicians, and citizens, and an extensive literature discussing


Today’s citizens awake in the morning to breakfasts of bacon and eggs, both certified as fit for consumption by the United States Department of Agriculture (although the Department of Health and Human Services would urge you to eat a breakfast lower in cholesterol). Breakfast is rudely interrupted by a phone call; the cost of phone service is determined by a state regulatory commission. When our citizens drive to work, their cars’ emissions are controlled by a catalytic converter mandated by the Environmental Protection Agency. The cars have seat belts, padded dashboards, collapsible steering columns and air bags required by the National Highway Traffic Safety Administration. When our citizens stop for gasoline, they pay a price that is partly determined by the energy policies (or a lack thereof) administered by the Department of Energy. To take their minds off the bureaucracies regulating their lives, the bureaucratic citizens turn on their radios. Each radio station is licensed by the Federal Communications Commission, and all advertising is subject to the rules and regulations of the Federal Trade Commission.
accountability exists.\textsuperscript{14} This “accountability literature” has covered much ground and taught us much; however, it tends to treat the accountability environment the agencies face as something that agencies either manipulate to achieve their goals or something to which they are subject. In the terms mentioned above, it tends to treat agencies as either villains or victims.\textsuperscript{15}


\textsuperscript{15} This terminology is inspired by, though it is not directly drawing on, Julian Le Grand’s classification of the way those drawing and operating the welfare state are viewed. See Julian Le Grand, Knights, Knaves or Pawns? Human Behavior and Social Policy, 26 J. SOC. POL’Y 149, 153–60 (1992) (describing how loss of faith in the welfare state led to seeing the officials
A. Agencies as the Villains

Much of the writing about agencies today portrays them as the enemy, or the villain, of the accountability story. There are several varieties of this approach. The most neutral one, the one least hostile to agencies, explains the need for accountability as a principal-agent problem: Congress created agencies to do its bidding. Agencies may have their own interests and prefer to follow their own preferences\(^\text{16}\) (or the preferences of the industries they are captured by)\(^\text{17}\) rather than follow the wishes of Congress.


Without accountability, agencies would be free to go their own ways and ignore Congress. Under this theory, scholars of the administrative state should address mechanisms of oversight over agencies, examine them, evaluate them, and suggest improvements.  

A more extreme version of the agencies as villains narrative focuses, instead, on examples of agency abuse and misconduct and uses that to demonstrate that agencies, whenever they can, make

---


themselves as unaccountable as possible and actively avoid the control mechanisms put in place by others. Many of these studies therefore suggest increased controls or improved enforcement of existing controls.

For example, in a number of recent studies, scholars have demonstrated that agencies avoid some of the procedures put in place by the APA. In a very recent article, Michael Kolber demonstrated that the FDA tended to use the procedure known as “direct final rulemaking,” in which an agency publishes a rule without going through the notice and comment process beforehand, not as it was intended, i.e., for non-controversial rules where notice and comment is a waste of time, but instead for rules expected to be controversial. These findings about the FDA are also reflected in an article by Lars Noah criticizing that agency for cavalierly ignoring legal and statutory requirements (even though Noah acknowledges that these “subversive” actions are part of the reason the agency has done “fairly well” in protecting the public health in the face of limited resources,

---

19 For a general discussion of direct final rulemaking, see Ronald M. Levin, Direct Final Rulemaking, 64 GEO. WASH. L. REV. 1 (1995) (detailing the history and analyzing the strengths of direct final rulemaking), and Ronald M. Levin, More on Direct Final Rulemaking: Streamlining, Not Corner-Cutting, 51 ADMIN. L. REV. 757 (1999) [hereinafter Levin, More on Direct Final Rulemaking] (defending direct final rulemaking and emphasizing the necessity for restraint instead of abandonment). In a critical vein, see Lars Noah, Doubts About Direct Final Rulemaking, 51 ADMIN. L. REV. 401 (1999) (arguing that the procedure of direct final rulemaking is invalid under existing law).

20 The point of direct final rulemaking is to do away with notice and comment in cases where there will be no comments submitted because the rule is non-controversial. See Levin, More on Direct Final Rulemaking, supra note 19, at 758–60. For further discussion of the waste of time resulting from use of notice and comment in certain cases, see Dorit Rubinstein Reiss, Tailored Participation: Modernizing the APA Rulemaking Procedures, 12 N.Y.U. J. LEGIS. & PUB. POL’Y 321 (2009) [hereinafter Reiss, Tailored Participation].


controversial issues, lack of leadership, and problematic legislative directives).²³ In a similar vein, Kristine Hickman, in a recent article, demonstrated that the Treasury does not follow the APA notice and comment rulemaking procedures in over 40 percent of its rulemakings.²⁴

Likewise, Ashutosh Bhagwat demonstrated that the Federal Communications Commission (“FCC”) for a decade labeled its actions in relation to tariffs “enforcement policy” rather than acknowledging it was rulemaking.²⁵ Using the Chaney doctrine,²⁶ it was therefore able to avoid judicial review until the D.C. Circuit finally called it on the issue.²⁷

Other studies used analysis rather than empirical methods to make the same points. For example, one writer claims that “government has no sense of accountability.”²⁸ Dobkin, focusing on the Immigration and Naturalization Services, sees agencies as lacking accountability by acting behind the scenes, which leads to “lawlessness.”²⁹

These are the type of studies that resonate most powerfully outside the academic community, mostly because they reflect stories of abuse that come up in the news and fit the general American tendency to distrust bureaucrats.

For example, in 2007 the Inspector General of the Department of the Interior started expressing concerns about the operations of the Department’s Minerals Management Service, an agency that collects government royalties from oil companies drilling on public lands. The Inspector General’s initial concern was that since the

²³ Id. at 902–03.
²⁷ Bhagwat, supra note 25, at 169–70.
Account Me In

Clinton administration, the agency has allowed oil companies to underpay. Later investigations showed a culture of accepting gifts from industry representatives, sexual relationships with representatives from oil and gas companies (and it’s not often you find an agency literally in bed with the regulated industry), and abuse of alcohol, cocaine, and marijuana in industry-organized parties. While the newspaper reports treated it as a classic example of lack of accountability, it should be remembered that it was an internal administrative control that discovered all this—the Department of the Interior’s own Inspector General, Earl E. Devaney—and it was the administrative machine that stepped in to punish the problem agency.

The same department’s lack of accountability was severely criticized during the oil spill in the Gulf of Mexico. The agency was criticized for not exercising sufficient oversight over the operations of British Petroleum (BP), the company owning the well that spilled over. Once again, the government—specifically, the Department of the Interior—took corrective steps, and very extreme ones. The agency was substantially reformed, and the reforms included a fundamental restructuring. For example, a royalty-in-kind program subject to many abuses was abolished, an independent Marine Board was ordered to review the agency’s inspection program for offshore facilities, and inspections of deepwater operations were ordered.

32 See Power, supra note 31.
Scholars are not the only ones adhering to the “agency as accountability villains” story. Politicians regularly attack agencies for their lack of accountability. That, for example, was at the heart of Representative Elliot Levitas’ strong promotion of the legislative veto, which would have given Congress control over agency rules. The same view was at the heart of the new and vigorous attempts to reintroduce the “Regulations from the Executive In Need of Scrutiny (REINS) Act” promoted by, for example, Congressman Geoff Davis, that would require all major rules to be approved by a Congressional Joint Resolution before they became operative. In a completely different example Nebraska Democratic Senator Ben Nelson, concerned about money withheld from the University of Nebraska, described it as “unaccountable Federal bureaucrats diverting millions of dollars into agency ‘slush funds.’”

In the discussions leading to the enactment of the IRS Restructuring and Reform Act of 1998, Senator Frank Murkowski from Alaska said that “[f]ederal agencies tend to act as if they are a law unto themselves, believing they are accountable to no one. . . . [T]he system was designed to avoid accountability.”

The studies of agencies as accountability villains capture an


important part of the truth; abuses occur in the administrative state—some that clearly cannot be justified, such as the behavior of the Minerals Management Service—but others which might be looked upon as making the best of a bad situation. There are times when the only way to get a job done is to bend some rules, given the complex situations that sometimes face agencies. Consider, for example, the FDA’s behavior described by Noah, which could be seen as the only way for the agency to actually get its job done in the face of severe financial and staffing constraints.\(^{39}\) But this is not the whole picture.

**B. Agencies as Victims**

A completely different view of agency accountability emphasizes the problematic nature of accountability in the American administrative state, focusing on an alleged excess of accountability mechanisms. Scholars who take this view do not deny that there are abuses by agencies.\(^{40}\) However, they suggest that it is more common to find a well meaning, hard working administration assailed from all sides by demands and accusations, so that simply doing its assigned job becomes extremely difficult.

The more neutral of these studies describe the complexity of the administrative state and address how agencies deal with that complexity. For example, in their study of public administration, Romzek and Dubnick offer a classic typology of accountability—bureaucratic, professional, political, and legal.\(^{41}\) This well-accepted classification\(^ {42}\) distinguishes between types of accountability on

---

\(^{39}\) Noah, *supra* note 22, at 902–903.


\(^{42}\) For examples of use, see Mark Bovens, *Public Accountability*, in THE OXFORD HANDBOOK OF PUBLIC MANAGEMENT 182, 185 (Edwin Ferlie,
two dimensions: whether the degree of control exercised by the accountability holder is high or low, and whether the source of control is external or internal.

Based on these two dimensions Dubnick and Romzek identify four possible types of accountability. Bureaucratic (also known as hierarchical) accountability refers to a high degree of control exercised within the agency or within the executive branch—by other agencies, the White House, and the President. It is hierarchical in nature. It includes, but is not limited to, relations between lower agency officials and higher agency officials. Professional accountability is internal but involves a low degree of control—it emphasizes professional norms and reputational mechanisms to control experts who require discretion to do their job. Legal accountability involves a high level of control exercised by an external actor; for agencies, this includes control by courts and Congress, through legislation or the budget. Finally, political accountability refers to a low level of control exercised by an external actor—for example, influence or pressure exerted by Congress-members, the media, and interest groups. Agencies face all these forms of accountability simultaneously and have to respond to them. The science of public administration, say Dubnick and Romzek, is the science of managing conflicting expectations.

Similarly, Radin discusses the challenges facing agencies when they try to deal with the accountability apparatus by examining a hypothetical new head of the Department of Health and Human Service and his ability to juggle the conflicting accountability demands he faces in his new job.

Hargrove and Glidewell ask how officials deal with “Impossible Jobs,” where the “clients” (e.g., welfare recipients,
prisoners) are not considered very sympathetic, and where there is high conflict among interested constituencies and little confidence in the profession and the agency. There are many other examples.

Other studies go further and argue that the intense accountability pressures agencies face have severe negative repercussions for the administrative state and the public interest. In his study of accountability, Behn suggests that the multiplicity of accountability mechanisms leads agencies to be blamed regardless of what they do, and that this excess blame can lead to a range of negative results—from defensive behavior to despair.

Similarly, Kagan, in his book *Adversarial Legalism*, tracks the problematic effects of the decentralized, multilayered system of government in the United States on making public policy. For example, he describes how the involvement of multiple actors—several federal and state agencies, as well as federal and state courts—made the dredging of the Port of Oakland very slow, much more costly than anticipated, thus costing the city of Oakland jobs and revenues. He acknowledges that adversarial legalism has benefits—making the system more open to new claims and more


47 Behn, supra note 14, at 1–6 (discussing the phenomenon); id. at 69–72 (discussing some of the negative impacts).


49 To such an extent it became uncertain whether the port would actually be dredged, costing Oakland’s port further business. Id. at 27–30.
However, at least when it comes to regulation, Kagan strongly suggests that the costs of adversarial legalism outweigh the potential benefits.51

The debate about “ossification” of the rulemaking process is another example of studies warning against the negative effects of excess accountability. In a famous and very strongly written article, Thomas McGarity criticized the complexities added to the rulemaking process as harmful to the functionality of the administrative state.52 In a subsequent article, he emphasized that it is unrealistic to tie agencies’ hands so thoroughly while expecting them to deliver and be effective.53 He also emphasized that the extensive accountability used undermines the agencies’ efforts to protect the public from the harms they were designed to combat.54 Other scholars expressed similar concerns,55 though recent empirical studies have cast doubts on the extent to which agency rulemaking is indeed ossified by having too much accountability forced upon them.56

50 Id. at 31–32.
51 Id. at 196–204.
52 See McGarity, Some Thoughts, supra note 40, at 1448–59.
54 Id. at 530–33.
The common thread that runs through all of these studies is the idea that agencies, subject to extensive accountability mechanisms, are often unfairly blamed for problems not of their own making, serve as politicians and scholars’ whipping boys, and have trouble doing their jobs. As with the “agencies as villains” narratives, these studies capture a part of the picture, but ignore another. It is to this missing link I turn now.

C. Agencies as Accountability Initiators

The part of the picture that current literature underemphasizes is the role of agencies as sophisticated actors managing their own accountability environment by creating and adding accountability mechanisms.

Public administration scholars acknowledge agencies acting autonomously in other contexts. For example, Carpenter describes in detail how several agencies created their own autonomous policies and managed to get the legislation they wanted from Congress by building a reputation for competence and for supporting the public interest. A focus on agency action or agency discretion, implicitly acknowledging that agencies have freedom to act, is at the core of most studies of public administration and the problem the “agencies as villains” narrative confronts.

However, these insights have not been applied to the study of accountability—i.e., so far agencies have not been treated as autonomous actors that can create and contribute to their accountability environment.

One paper that stands out as an exception is Elizabeth Magill’s

---


recent study of self-regulation. Magill examines agencies’ voluntary adoption of regulations that limit their discretion. She makes an important contribution to the literature by taking a clear, unbiased view of agencies’ activities and by emphasizing the agencies’ role in creating the regulatory environment they operate in.

Many of the limits agencies adopt through self-regulation may be seen as reforms that increase accountability—but not all. And since Magill emphasizes, for most of her discussion, self-regulation activities that were actually embedded in agency rules—which will therefore be enforced by the courts under the Accardi principle—some of the efforts agencies make to increase their accountability are not captured by her discussion, such as the IRIS system mentioned in the introduction to this Article. More importantly, her methodology—an analytical discussion—is dramatically different from the qualitative empirical description based on case studies used in this Article. Also, her explanations for why agencies adopt self-regulation focus mostly on what I describe as “rational choice” explanations, and are therefore more limited than those used here.

This Article adds to the literature by suggesting that agencies also act voluntarily to increase their accountability; moreover, it provides detailed case studies of such behavior. Agencies do so for a number of reasons, including their self-interest, but also due to the power of ideas brought in by appointees from outside the civil service and to the role conception of the civil servants.

One challenge a claim like this faces is how to define accountability. The word “accountability” suffers from overuse and misuse, and scholars have expressed concerns about the term

60 Id. at 861.
61 See infra text accompanying notes 125–32 for examples of reforms that increase accountability.
62 According to which agencies are bound by their own regulations. Magill, supra note 59, at 877–81.
63 Id. at 884–91.
Account Me In

losing its meaning. It has been defined as covering electoral accountability, punishment, or control of one party by another. Very broadly, cognitive psychology scholars see accountability as:

[an] implicit or explicit expectation that one may be called on to justify one’s beliefs, feelings, and actions to others . . . [It] also usually implies that people who do not provide a satisfactory justification for their actions will suffer negative consequences ranging from disdainful looks to loss of one’s livelihood, liberty, or even life . . . . Conversely, people who do provide compelling justifications will experience positive consequences . . . .

To solve the problem of defining accountability, I am focusing my discussion on two types of reforms that agencies often adopt: increasing their transparency and increasing their responsiveness to external actors. Increasing transparency is often suggested as a means of increasing accountability and allowing external actors

---


66 Id. at 14–16.

67 Id. at 14–16.

68 Jennifer S. Lerner & Philip E. Tetlock, Accounting for the Effects of Accountability, 125 PSYCHOL. BULL. 255, 255 (1999). Although developed in relation to studies of individuals, the definition has been used in relation to agencies before. Reiss, Agency Accountability, supra note 46, at 114 n.1; Mark Seidenfeld, The Psychology of Accountability and Political Review of Agency Rules, 51 DUKE L.J. 1059, 1064 n.26 (2001).

more of a say increases control, as well as the threat of sanctions.\textsuperscript{70} Therefore, reforms in which agencies increase their transparency or the influence of other actors can be taken with some confidence as examples of efforts to increase accountability.

II. AGENCIES WANT TO BE ACCOUNTABLE

If, as I argue, some agencies want to be accountable, the question remains, why? This part of the Article suggests some explanations, which are supported by the case studies detailed in Part III. I will draw on several strands of literature about the administrative state, as well as on general features of agencies, and examine three strong motivators that can create a quest for accountability within an agency. The first focuses on what we may term the “rational choice model”—i.e., efforts by an agency to minimize costs arising out of a successful accusation of lack of accountability, and to maximize the benefits to itself that derive from accountability.\textsuperscript{71} However, other explanations are just as persuasive; this Article examines two of those alternative explanations.

One explanation is based on the power of ideas, and specifically upon agencies’ acceptance of new ideas about the importance of transparency and participation—ideas drawn from practitioners, consultants, and scholars which have become prevalent in the world of governance. The second is based on agencies’ role conception. This explanation claims that in today’s world, bureaucrats have internalized the need to be accountable as part of their mission and role conception, and invest in accountability as an integral part of doing their job. In a sense, these explanations overlap, but there is an important difference in

\textsuperscript{96} (characterizing transparency as a dimension of accountability); May, \textit{supra} note 14, at 11–12 (characterizing transparency as increasing accountability).


\textsuperscript{71} This, for example, is the argument suggested to explain the EPA’s behavior in Part III.B.
the “target audience” on which they focus. As the case studies suggest, ideas are frequently brought into an agency through the appointment of outsiders. Peter Hutt came to the FDA from private practice, Richard Merrill from academia, and Charles Rossotti was a businessman before becoming commissioner of the IRS. All three were strongly committed to transparency in a way that the staff may not have been at the time they came in.

However, sometimes separately and sometimes as a response to the reform, ideas can also be internalized by the agency’s civil servants, those who run day-to-day operations of the agency. In that case we are talking about role conception. It’s not just about the power of the idea itself; the issue becomes the way the civil servant sees her job, a matter of duty rather than ideology. Even before the 1998 reform, several of the IRS staff members were promoting increased transparency; they had, apparently, internalized the ideas of transparency as part of their role. Even more striking, after the reforms introduced by Hutt were implemented, there were some FDA officials who wanted to go further in transparency than he did. Such occurrences seem to reflect a redefinition of the officials’ roles in their own minds. All three of these explanations have some applicability to the case studies in Part III.

A. Rational Choice: Agencies Want to be Accountable to Maximize Benefits and Minimize Harm

The rational choice paradigm as it applies to agencies sees bureaucrats as self-interested utility maximizers. The classical approach posits that bureaucrats seek to maximize their budget. However, more recent approaches add in bureaucrats’ desire to maximize preferred policy outcomes.


73 NISKANEN, supra note 72, at 40–41.

74 Sean Gailmard & John W. Patty, Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise, 51 AM. J. POL. SCI. 873, 874
While there has been substantial criticism of rational choice theory as not empirically grounded and in tension with the realities of the administrative state, in the context of accountability, there is some indication in the case studies and in current literature that agencies do act to increase their accountability because of a cost-benefit analysis.

In the last decades governments have suffered from a legitimacy crisis. Trust in government has been dropping substantially. The perception of a legitimacy crisis easily leads to increased pressure on agencies to be accountable and more and more efforts are put into holding them accountable. The large amount of literature on accountability in the last decades demonstrates how important the issue has become. In a sense, the administrative state today is the administrative state under attack. In this environment, agencies pay a very high price for an accusation of lack of accountability that sticks. All the agencies


78 See supra note 14.

79 Rubin, supra note 14, at 74–75.
described in Part III were agencies under attack, and there are indications that they tried to increase accountability to reduce pressure and prevent negative consequences. Claims of lack of accountability are often raised to justify demands to reform agencies, cut their budget, change the governing legislation, or other adverse consequences. For example, the IRS’s reorganization of 1998 was motivated at least in part by complaints that the IRS was not sufficiently accountable to Congress. Agencies naturally want to avoid such consequences. At the very least, not being accountable can mean another actor will add accountability mechanisms, and as demonstrated in Part I.C, agencies already face a plethora of them; these demands add work and take up resources that can be used elsewhere; what sane bureaucrat would want more of them imposed from the outside? No wonder, then, that administrators and agencies want to demonstrate that they are accountable and do not fall into the category of “evil,” unaccountable bureaucrats.

This is the main argument Magill uses in her article; she describes agency self-regulation as motivated by a rational desire to increase the benefits to the agency. The reasons she suggests include giving agency heads the ability to control lower officials to whom they delegate authority; clarifying the problem internally and helping bureaucrats explain their decisions; publicizing an agency’s policy and offering stronger commitment; limiting future changes of policy following a change of administration; protecting agency autonomy against intervention from political

---

80 See generally NAT’L COMM’N ON RESTRUCTURING THE INTERNAL REVENUE SERV., A VISION FOR A NEW IRS (1997).
81 See BEHN, supra note 14, at 14–15. Behn gives the example of the rules adopted by government procurement officials, observed by scholar Steven Kelman to add complexity to the process to avoid legal protest or political challenges. See id. at 15.
82 Magill, supra note 59, at 884–91.
83 Id. at 884–86.
84 Id. at 887.
85 Id.
86 Id. at 888.
actors; and increasing production of collective goods as information and reputation.

Some of these reasons apply to agency measures increasing accountability. Increased accountability can limit the illicit power of any one actor by forcing the agency to make the basis of its decisions clear, and thus protect agencies against political interference. In addition, it may make it harder to change existing policy; may provide the higher ranks of the agency with more information about what the rest of the agency is doing; and may increase its reputation.

B. Pantouflage: The Power of Ideas and the Role of Political Appointees

However, on its own, the cost/benefit argument is insufficient, for two reasons. First, a strong cost/benefit argument can also be made that suggests agencies should want to avoid accountability and not add to the already existing complicated system they face, an argument commonly made. That’s one of the arguments the “agencies as villains” story draws on. For example, one of Paul Light’s interviewees said about the desire of Presidential administrations for a strong inspector general: “Everybody wants a strong IG operation until it starts investigating them. The administration may start out thinking they want junkyard dogs and

---

87 Id. at 889.
88 Id. at 890–91.
89 The term Pantouflage is taken from studies of French public administration and refers to the prevalent practice of the senior elite trained in the most prestigious schools moving back and forth between the public and private sectors. Luc Rouban, The Administrative and Political Elites, in PASCAL PERRINEAU & LUC ROUBAN, POLITICS IN FRANCE AND EUROPE 121, 132 (2009). The United States parallel is that one way to maintain connections and flow of ideas between the private and public sector is through personnel exchange, appointing people with previous private sector background to senior positions in the public service. This is somewhat similar to the United States concept of “revolving door.”
90 See McCubbins, Administrative Procedures, supra note 16, at 248–49; see also BEHN, supra note 14, at 15; Dobkin, supra note 29, at 379–82; Hood, supra note 69, at 192.
what they end up getting is French poodles."\(^{91}\)

Second, while a rational choice explanation may explain why an agency chooses to work toward increased accountability, it says nothing about the choice of method. Nor does it explain the level of commitment some agencies show to the accountability endeavor. In a world in which government agencies have limited resources—what has been referred to by some as a period of austerity\(^{92}\)—some agencies devote substantial portions of their scarce resources to accountability. There is clearly more going on here than mere protection of self-interest; a more nuanced explanation would appear to be needed.

One such explanation is the power of ideas. As demonstrated by B. Guy Peters, current ideas about governance draw on several extremely important traditions, many of which are connected to accountability.\(^{93}\) Scholars have demonstrated that ideas can have strong influence on the behavior of organizations.\(^{94}\) In the case of


\(^{93}\) See B. GUY PETERS, THE FUTURE OF GOVERNING 16–22 (2001) (demonstrating that among the ideas that shape governance in today’s world are participation and market-based ideas, including ideas of transparency).

agencies, ideas can also be translated into actual pressures and changes through politicians, interest groups, and the administrators’ epistemic community, including those that write about agencies.

One set of ideas which greatly influences the behavior of agencies draws upon market ideology and private sector reforms. An argument made by supporters of these ideas is that traditional hierarchical mechanisms of accountability do not work very well; they argue that market style mechanisms can provide better accountability and achieve better results. Whether or not, as an empirical matter, market-style reforms do in fact improve accountability—and there is doubt about that claim—the idea


95 For example, reforms attempting to introduce ideas current in the private sector—such as competition, privatization, incentive-based approaches—into the public sector. See, e.g., JOHN E. CHUBB & TERRY M. MOE, POLITICS, MARKETS, AND AMERICA’S SCHOOLS (1990); PETERS, supra note 93, at 22–30; Isabel M. Bjork & Catherine R. Connors, Free Markets and Their Umpires: The Appeal of the U.S. Regulatory Model, 22 WORLD POL’Y J. 51, 51 (2005); David Levi-Faur, The Global Diffusion of Regulatory Capitalism, 598 ANNALS AM. ACAD. POL. & SOC. SCI. 12 (2005); David Levi-Faur, Regulatory Capitalism: The Dynamics of Change beyond Telecoms and Electricity, 19 GOVERNANCE 497 (2006).


97 For those raising concerns about the effect of market style reforms on accountability, see Carol Harlow, Public Service, Market Ideology, and Citizenship, in PUBLIC SERVICES AND CITIZENSHIP IN EUROPEAN LAW (Mark Freedland & Silvana Sciarra eds., 1998), GREG PALAST ET AL., DEMOCRACY AND REGULATION: HOW THE PUBLIC CAN GOVERN ESSENTIAL SERVICES 20–22 (2003) (describing British and Indian transactions in which review and comment or other public review processes were insufficient and thus unnecessarily costly), and Martha Minow, Public and Private Partnerships: Accounting for the New Religion, 116 HARV. L. REV. 1229, 1260–63 (2003). On the other hand, there is also a substantial literature supporting the claim. See e.g., STEVEN K. VOGEL, FREER MARKETS, MORE RULES 17–21 (1996) (incorporating market forces considerations into a state institutions framework); Scott Furlong, Political Influences on the Bureaucracy—the Bureaucracy Speaks, 8 J. PUB.
exists and is powerful, and this belief in the efficacy of the market can easily drive public servants to try to increase their own accountability through methods that aim at exploiting the reputed advantages of the market. Agencies whose work is not easily privatized or which cannot really compete with the market’s reputed price discipline may emphasize, instead, more achievable benefits such as increased information and transparency.98

Many times people appointed to lead agencies have come in with a strong belief in transparency.99 In the cases described in this Article, the appointment of Peter Hutt as Chief Counsel of the FDA, with his belief in transparency, directly influenced the reforms adopted. In that case, the politically appointed commissioner was also onboard. In the case of the IRS, Commissioner Rossotti’s belief in the need for reform and his continued belief in transparency also advanced the reform.

Yet another set of important ideas that influence modern agencies relate to increasing public participation and the role of citizens in government.100 Substantial amounts of scholarship have promoted the idea of giving citizens more opportunities to participate, often suggesting new and original modes of doing so.101 Practical experiments in participatory government have been

---

98 See COSMO GRAHAM, REGULATING PUBLIC UTILITIES: A CONSTITUTIONAL APPROACH (2000), for an example of the regulatory agencies in England trying to increase their legitimacy through transparency.

99 See generally WOOD & WATERMAN, supra note 18; see also Furlong, supra note 97, at 39.

100 PETERS, supra note 93, at 50–64.

conducted. The effect of this may have been much greater on agencies than on Congress or the President. Much attention has been given to efforts to increase participation in agency proceedings. Just as with market ideology, a good deal of

---


criticism has been directed towards these efforts. Quite a lot of this criticism has been directed at the methods, mostly in relation to inherent inequalities in the ability to participate and influence, but there has also been discussion of the inappropriateness of participation to certain administrative decisions. Even so, the influence on bureaucrats has been substantial. These ideas enter the bureaucratic consciousness through training and scholarship, as well as pressure from political appointees to the agency and from the White House.

In the cases discussed below in Part III, Hutt’s belief in input led him to raise the importance of comments by requiring the FDA to respond to each comment it received. The desire to increase input evidenced by both the EPA staff and commissioners directly relates to the style of reforms that were adopted there. At the IRS, the Taxpayer Advocacy Panel directly increased the role of private citizens—hopefully representative of “the public”—in decision-making. In all three cases, the reforms aimed at increasing stakeholder participation.


C. Role Definition: Bureaucrats Expand Their Commitment to Mission to Include a Commitment to Accountability

An important theme that emerges from the public administration literature is bureaucrats’ involvement in policy making\(^\text{106}\) and bureaucrats’ strong commitment to the mission of their particular agency.\(^\text{107}\) Challenging the rational choice view of the self-interested bureaucrat, a whole line of public administration studies have suggested that many of those going into the public service do so because they are motivated to participate in the making of policy and in achieving policy goals, creating a better world.\(^\text{108}\) The most recent line of public administration studies addressing this, starting in the 1990s, coined the term “Public Service Motivation.”\(^\text{109}\) This line of literature used empirical

---

\(^{106}\) ABERBACH & ROCKMAN, supra note 58, at 8.


Not all civil servants are devoted to the mission, even under theories that acknowledge that some are, see DOWNS, supra, at 83, and such devotion does not always have positive consequences, for example strong bureaucratic loyalty to one mission may motivate administrators to resist certain changes and tasks. See Brewer et al, Individual Conceptions, supra note 107, at 261.


survey data in an attempt to compare the attitudes and motives of public servants to those of business executives. It consistently showed that high level civil servants were more likely to be motivated by the mission and the public interest than were their private sector counterparts.110 Not only that, but the studies provide evidence that suggests that public servants are more motivated by intrinsic job satisfaction and the opportunity to provide service and less motivated by financial rewards than their private sector counterparts;111 that they are more strongly motivated when they feel their mission is important;112 and their devotion to the public interest is not limited to their jobs, as they also volunteer more outside their professional life in terms of both money and time.113 Needless to say, this is a general description and does not describe all civil servants; at least one study classified public servants according to their motivation and found some variety.114 But the trend is clear, and perhaps not surprising—after all, in the United States, high level public servants usually have a graduate degree,115


110 See Brewer & Selden, supra note 109, at 429–33; Philip E. Crewson, Public-service Motivation: Building Empirical Evidence of Incidence and Effect, 7 J. PUB. ADMIN. RES. & THEORY, 499, 500, 512 (1997); Perry, Measuring Public Service Motivation, supra note 107, at 6–9, 20–21; Perry & Wise, supra note 109, at 369-70. One limitation of this literature is that it focuses almost completely on the upper levels of the civil service, i.e. people in management positions, and therefore will not tell you much about the motivation of your mail carrier or customs official; however, the kind of accountability mechanisms discussed here are usually created at the policy-making level, as the case studies demonstrate. Thus, the high-level population that this literature describes is exactly the right one to study for the purposes of the present Article.


112 See Wright, Public Service and Motivation, supra note 111, at 60.


114 See Brewer et al., Individual Conceptions, supra note 107, at 255.

but compared to private sector employees they are paid less and certainly enjoy less prestige than managers in private businesses or equivalent level civil servants in other countries. They face jobs that are typically very challenging and have relatively less power than elected officials. Why would anyone take on such work unless he/she cared deeply about it and the interests it serves?

In today’s world, where accountability is such a strong word, it is no wonder that certain civil servants accept commitment to accountability as part of their mission, as the case studies in Part III reflect. Agencies believe they should be accountable, not just because they buy into ideas of transparency and participation, but because it’s part of their role definition: as they view it, one aspect of doing a good job is to be accountable. Accordingly, they are willing to make efforts and act in ways that will promote accountability. Furthermore, at least one study demonstrated that reforms in the public sector could increase the level of public service motivation, including reforms aimed at increasing accountability.


120 See generally Moynihan & Pandey, supra note 109.
In other words, civil servants—bureaucrats—are strongly motivated to serve the public interest; they will begin to actively seek increases in transparency and/or responsiveness if and when they perceive that the public interest calls for that kind of reform. In a reality that emphasizes accountability, we can expect civil servants to internalize the idea that accountability is an integral part of their mission, one that is completely necessary if their job is to be done well.

Some support for the idea that bureaucrats internalize the need for increased transparency and responsiveness comes from a recent survey of bureaucrats’ attitudes to e-rulemaking conducted by Jeffrey Lubbers, a renowned expert on rulemaking. Aside from his empirical findings, Lubbers reports on the comments made about e-rulemaking, many of which were positive. Among the positive aspects bureaucrats emphasized were the improvement rulemaking creates in the ability of the public to participate and the transparency of the process.121 Here are some examples from his responses:

_E-rulemaking is the obvious choice for encouraging public comment and allowing easy access to records from anywhere and without risking the loss of original hard copies._

_With more people using the Internet, it seems the right way to conduct rulemaking and promises to reach more folks who don’t read the Federal Register._

_In addition to reaching older members of society, making the process available online makes it more likely we will reach members of Generation X and the Millennium Generation._

_E-rulemaking is better at letting the public know what the agencies are doing than it is at providing thoughtful input_

into the decisions themselves.

[M]aking agency rulemaking more accessible to the public...

Makes it much easier for the public to see the comments...\textsuperscript{122}

There were, of course, also negative comments,\textsuperscript{123} but the reason given for the positive comments—the increase of access—supports the claim that many bureaucrats had internalized the need for public access and accountability.

The level of resources and efforts devoted to accountability that I have described in the case studies below may or may not truly indicate an internalization of a desire for accountability by agency staffs. However, officials of the agencies repeatedly express the view that the observed actions demonstrate exactly that. Hutt strongly emphasized how the FDA officials internalized the need for transparency, to a degree where he felt obliged to say “stop,” refusing the publication of trade secrets related to new drug applications because of the negative policy consequences that would have resulted. In speaking to the IRS’s officials, they too say many of the staff, especially at the higher levels, internalized the need to increase transparency.\textsuperscript{124} In neither case did I conduct a thorough investigation of the staff to see whether they have, indeed, internalized the norms; but that is the impression of important, well informed observers and it is supported by the efforts devoted to increasing accountability.

III. CASE STUDIES OF AGENCIES SEEKING ACCOUNTABILITY

Through several case studies, this section develops an argument that agencies seek out ways to be accountable. Showing

\textsuperscript{122} Id. at 471–72.
\textsuperscript{123} See id. at 472–74.
\textsuperscript{124} Telephone Interview with IRS Official (under promise of confidentiality) (June 24, 2009).
Account Me In

that agencies are willing to add accountability mechanisms is not hard. Examples are legion. To use a major one, the APA exempts matters “relating to agency management or personnel or to public property, loans, grants, benefits, or contracts” from rulemaking procedures.125 On its face, this would mean agencies providing government benefits are not required to go through notice and comment. However, again and again, government agencies dispersing benefits adopted the notice and comment requirements in their own regulations, without any legal obligation—even though such adoption subjects them to possible criticism and facilitates judicial review.126 The Department of Labor adopted notice and comment procedures when implementing the Comprehensive Employment and Training Act, saying:

It is the policy of the Secretary of Labor, that in applying the rulemaking provisions of the Administrative Procedure Act (5 U.S.C. 553), the exemption therein for matters relating to public property, loans, grants, benefits or contracts shall not be relied upon as a reason for not complying with the notice and public participation requirements thereof . . . .127

Other examples cited by Lubbers in his book on rulemaking include the Department of Transportation, the Department of Housing and Urban Development, the Department of Defense, the Department of Health and Human Services, the Department of Agriculture, and the Small Business Administration.128 As Part III.A documents, a self-initiated elaboration and deepening of the elements of notice and comment was a substantial part of the reform of the FDA procedures in the 1970s. Courts treat deviations from these regulations as if they were violations of the notice and comment procedures in the APA;129 accordingly, an agency

127 29 C.F.R. § 2.7 (1979).
128 See LUBBERS, A GUIDE, supra note 126, at 63 n.60.
129 See, e.g., Rodway v. U.S. Dep’t. of Agric., 514 F.2d 809 (D.C. Cir. 1975); see also LUBBERS, A GUIDE, supra note 128, at 63 n.61.
adopting them is at risk of having its work overturned on procedural grounds or because of not responding to comments—incurring an increased risk of sanctions.

In another example, before engaging in the rulemaking that was the subject of the famous Vermont Yankee case, the Atomic Energy Commission (as it was then called) voluntarily adopted adversarial proceedings, including cross-examinations, even though those added to the length of proceedings.131

Taking this further, Elizabeth Magill demonstrated that agencies often engage in self-regulation, regulation limiting their discretion, providing many examples.132

A few of these were:

[T]he Social Security Administration’s “grid” regulations, which succeeded in turning the question of whether a party is disabled into a series of (more) objective questions.133

[T]he Food and Drug Administration’s decision to provide notice and invite comment on its “guidance documents” even though the APA would not have required it.134

In the following case studies, I track in more detail some examples of increased accountability, focusing on three agencies, in chronological order of the efforts described. The FDA reformed its procedures to increase transparency and participation in the 1970s, and provides an early case study of such behavior. Since at least the 1980s, the EPA has been experimenting with new ideas to increase its accountability. Finally, the IRS, widely held up for years as an example of complete non-accountability, has been working for over ten years on increasing its transparency and

---

132 Magill, supra note 59, at 866–69.
133 Id. at 867.
134 Id. at 868.
responsiveness to the general public.

The history of these agencies shows the persistent phenomenon of agencies seeking to increase their accountability, covering as it does a span of time that begins in the 1970s (for the FDA cases) and continues through the 1980s (with the EPA’s efforts) up to the present (with the FDA and the EPA’s later efforts and the IRS’s actions). When considering the significance of the following three case studies, we should note that the subjects of the studies are all large and important agencies. The IRS is a mammoth agency with over 100,000 employees spread throughout the country, and its actions affect the lives of almost every citizen and resident. The FDA and the EPA likewise work in areas that directly affect the quality of life of most citizens of the United States, regulating, between them, food, air quality, water quality, medicine, and other areas. These are bodies whose actions have great effect upon United States public policy.

A. The Efforts of the FDA to Increase its Accountability

In the 1960s and 1970s dramatic procedural changes were imposed on agencies. In 1966 Congress adopted the Freedom of Information Act (“FOIA”) requiring that governmental information be made public unless it fit into one of the exceptions in the Act.136 In the 1970s, the D.C. Circuit required that agencies respond to the major issues raised in comments137 and emphasized that the basis of a decision must be clearly explained.138 But even as the D.C.

---

136 Today it is codified as 5 U.S.C. § 552.
An agency need not respond to frivolous or repetitive comment it receives. However, where apparently significant information has been brought to its attention, or substantial issues of policy or gaps in its reasoning raised, the statement of basis and purpose must indicate why the agency decided the criticisms were invalid.
See also Portland Cement Ass’n v. Rucklehouse, 486 F.2d 375, 394 (D.C. Cir. 1973).
Circuit started its quest to control rulemaking,139 certain agencies started adopting similar requirements on their own, increasing their accountability voluntarily. This section describes one such story, that of the FDA, and examines two sets of efforts: the FDA’s work to increase its transparency by internalizing the values in the FOIA140 and the FDA’s procedural reform, both of which occurred during the same time period (early to mid 1970s) and reflected the same spirit of increasing transparency and participation.

1. Efforts to Increase Implementation of FOIA

Criticisms of the weakness of FOIA implementation were common in the early 1970s for agencies generally, and the FDA was no exception; several specific criticisms of it were made.141 Historically, the FDA, like other agencies, produced minimal compliance, releasing information only sparsely and reluctantly. This was a defense mechanism against criticism: “[I]f the public doesn’t get information, they have a difficult time objecting to what FDA does.”142 That changed in 1972.

When Peter Barton Hutt was appointed as Chief Counsel for the FDA in 1971, the importance of improving the agency’s transparency was impressed on him both by his predecessor, William H. Goodrich, and by the then Commissioner, Charles C. Edwards. The mandate fit in with Hutt’s own philosophy. Hutt


140 Administrative Procedure Act, 5 U.S.C.A. § 552 (West 2010).


142 Telephone Interview with Peter Barton Hutt, formerly Chief Counsel, FDA, currently Senior Counsel, Covington & Burling LLP (June 20, 2009).
viewed transparency as an important accountability mechanism, one that should be put in place:

My personal philosophy is that full disclosure to the public is the essence of democracy.... if the public doesn’t understand what’s going on in government, government can hide all forms of mischief and injustice. Whereas if everything is made public government has someone looking over their shoulder—sunlight kills a lot of issues.\(^{143}\)

His first weeks on the job were devoted to determining how the FDA should comply with FOIA in a way that would increase its transparency.\(^{144}\) He used his substantial discretion in implementing the mandate he received to create a solid framework, categorizing each document and creating a set of rules regarding its disclosure. He met every Friday with a team consisting of the FDA Commissioner and other senior officials in the FDA, keeping them informed and on board, and providing them a chance for input into the process. While the work was Hutt’s, the policy was endorsed and supported by the agency’s heads, all of whom at least accepted, if not actually endorsed, the need for increased transparency. Richard Merrill, Hutt’s successor as Chief Counsel, explained that, “Hutt did the major job of convincing the main people in the agency—the ones I refer to as our client—people responsible for substantive programs.”\(^{145}\)

The result was a proposed rule published in the *Federal Register* on May 5, 1972.\(^{146}\) The rule spread over ten of the Register’s small-print, three-column pages, and included substantial details. In the preamble to the rule, the Commissioner, signing the proposed rule, expressed his commitment to FOIA’s basic premise that “public disclosure should be the rule rather than the exception.”\(^{147}\)

\(^{143}\) *Id.*

\(^{144}\) *Id.*

\(^{145}\) Telephone Interview with Richard A. Merrill, formerly Chief Counsel, FDA, currently Professor of Law, Emeritus, Virginia Law School (July 7, 2009).


\(^{147}\) *Id.*
That does not mean every piece of information the agency had was disclosed. The most notable exception was in relation to New Drug Applications (“NDAs”), and information containing trade secrets submitted with them. When submitting an application to license a new drug, a company has to submit substantial amounts of information to demonstrate the drug’s safety and efficacy, much of which it would not want its competitors to know. Trade secret confidentiality is preserved under at least three legal sources: Section 552(b)(4) of FOIA (permitting confidentiality), the Trade Secrets Act (prohibiting disclosure of trade secrets by federal officials), and section 331(j) of the Food, Drug, and Cosmetic Act (prohibiting the revealing of trade secrets to anyone outside the department other than the courts or Congress). Hutt decided that the strong prohibitions on disclosing trade secrets, as well as longstanding agency precedent, supported a restrictive approach to disclosing materials attached to NDAs, under which the supporting materials would remain confidential. However, to


149 Administrative Procedure Act, 5 U.S.C.A. § 552(b)(4) (West 2010) (“This section does not apply to matters that are . . . trade secrets and commercial or financial information obtained from a person and privileged or confidential.”).


Whoever, being an officer or employee of the United States or of any department or agency thereof, . . . publishes, divulges, discloses, or makes known in any manner or to any extent not authorized by law any information coming to him in the course of his employment or official duties . . . which information concerns or relates to the trade secrets . . . shall be fined under this title, or imprisoned not more than one year, or both; and shall be removed from office or employment.


The following acts and the causing thereof are prohibited; . . . revealing, other than to the Secretary or officers or employees of the Department, or to the courts when relevant in any judicial proceeding under this chapter, any information acquired . . . concerning any method or process which as a trade secret is entitled to protection . . .

152 Telephone Interview with Peter Barton Hutt, supra note 142.
Account Me In

preserve the public’s right to know and in the interest of transparency, the agency published summaries of the reasoning and materials behind a decision—a Summary of Basis of Approval (“SBA”). This decision was arrived at through a consultation between Hutt and Richard Crout, head of the Bureau of Drugs, in which Crout agreed that the SBA was a reasonable compromise between protecting trade secrets and assuring transparency. It was criticized by industry members who saw the SBA as disclosing too much information and by consumer interest groups for not providing enough information. But it was clearly an effort by the agency to go beyond its previous practices and increase its transparency beyond what was mandated under FOIA, as the agency interpreted it.

It took over two years to finalize the FDA’s public information rule (which is not unusual—a recent study found that the average time for rules from notice to final rule is 2.2 years), and during that time the agency received 667 letters, including 68 substantive comments to the rule (which the final regulation answered in

155 Judith Axler Turner, Consumer, Report/FDA Pursues Historic Role Amid Public, Industry Pressures, 1975 NAT’L J. REPS 250, Feb. 15, 1975, at 250, 254 (on file with author). For a more recent example of the same criticism, a known consumer rights advocate on health matters, Dr. Sidney Wolfe, a member of the consumer watchdog group Public Citizen, said in response to the FDA’s creation of a task force to increase transparency in 2009:

For something like 36 years, through litigation and every other means, we have been trying to expand access to data on drug safety and efficacy . . . To make access to clinical trial data [happen] much sooner is a great idea for the public, for everyone that’s involved. . . . It’s anti-scientific and anti-intellectual to have these important data secret.

Steven Reinberg, FDA to Study Ways to Be More Open with Public, BLOOMBERG BUSINESS WEEK (June 2, 2009, 4:00 PM), http://www.businessweek.com/lifestyle/content/healthday/627708.html; see also Halperin, supra note 153, at 287.
156 West, supra note 14, at 69.
However, the change in policy was implemented immediately, without waiting for the regulations to be finalized, although it was changed somewhat as a result of the comments received.158

The regulations that the FDA promulgated between 1972 and 1974 had a dramatic impact on the FDA’s FOIA practice: before the regulations, the FDA granted only about 10 percent of FOIA requests submitted to it; after them, it granted about 98 percent.159 The agency seemed to have internalized and bought into the new and expansive approach to transparency. In fact, Hutt describes how agency staff wanted to go further than he thought appropriate on certain issues:

[Agency staff] . . . feel conflicted on safety and effectiveness data. They get so much criticism for not releasing it from people who don’t understand that that’s confidential trade secret, and they want to just release it— and get rid of the criticism. But you have to understand the consequences—releasing all trade secrets will destroy the American pharmaceutical industry. 160

The goal of this increased transparency, as stated by Peter Hutt, was to increase public scrutiny of the FDA’s actions with a view to allowing the public to prevent abuses. Critics continued to challenge the FDA on grounds that it still has not done enough;161 but the agency, in this case, acted with the goal of increasing accountability. Nor was this the last time the FDA acted to increase its transparency, though the later examples are beyond the scope of this project.162

158 Telephone Interview with Peter Barton Hutt, supra note 142.
159 Halperin, supra note 153, at 286.
160 Telephone Interview with Peter Barton Hutt, supra note 142.
162 More recently, the agency created a task force dedicated to increasing its transparency and a “transparency blog” dedicated to following the agency’s efforts to increase its transparency and allowing the public to comment. See About This Blog, FDA TRANSPARENCY BLOG (Nov. 10, 2008), http://fda
2. Procedural Reform

Today’s administrative law is marked by complex procedures and substantial demands on agencies. To avoid having their decisions labeled “arbitrary and capricious,” agencies must explain their actions in the “concise general statement” demanded by the APA and must respond to the comments they received. In fact, in today’s world, the multitude of requirements for explanations by agencies are criticized more often than not. But there was a time when the requirement of explaining an agency’s action was new and exciting. So when the FDA adopted a set of procedural regulations that required extensive information to be provided in a rule’s preamble, it was ahead of its time in imposing these requirements on itself to increase its accountability.

The background to the procedural reform was a period of divided government, with President Richard Nixon in the White House until 1974, followed by Gerald Ford, both Republicans facing a Democratic Congress. The democrats in Congress were wary of the FDA, and kept a close eye on it, repeatedly criticizing

transparencyblog.fda.gov/about-this-blog.html.

Administrative Procedure Act, 5 U.S.C.A. § 553(c) (West 2010).


SHAPIRO, supra note 165, at 47–48; Melnick, supra note 165, at 246–47; Pierce, Seven Ways to Deossify Agency Rulemaking, supra note 55, at 68–69.

its actions and demanding explanations from officials.\textsuperscript{168} In an effort to increase the agency’s legitimacy and reduce Congress’ concerns, Hutt undertook to reform the agency’s procedures,\textsuperscript{169} implementing what were innovations at the time, though these rules would later become a staple of administrative law.

The change was perhaps most extreme in relation to rulemaking. Like many other agencies, the FDA initially used adjudications as its main mode of decision-making. However, as the agency’s authority was expanded by Congress, mounting pressures of workload made this inefficient, and the agency increased its use of rulemaking in the early 1960s.\textsuperscript{170} One of Hutt’s major projects was to guide the implementation of the change to an agency that works primarily through rulemaking.\textsuperscript{171} This change fit with the general trend towards increased rulemaking in the 1960s–1970s,\textsuperscript{172} fueled in part by concerns about the administrative state’s accountability—rulemaking was seen as more “sleek, efficient and fair” compared to the slowness, uncertainty, and potential arbitrariness inherent to agency adjudication.\textsuperscript{173}

However, rulemaking as the FDA implemented it did not look like rulemaking as today’s administrative scholars describe it.\textsuperscript{174} Instead, rulemaking included a short notice including the proposed regulation (and the other minimal information required by the APA),\textsuperscript{175} and after the comments had been submitted, the agency would publish the final rule with a statement that “having considered the comments, I [i.e., the commissioner] hereby

\begin{footnote}
\textsuperscript{168} Telephone Interview with Peter Barton Hutt, supra note 142.
\textsuperscript{169} \textit{Id}.
\textsuperscript{170} \textit{See} Schiller, \textit{supra} note 139, at 1148–49.
\textsuperscript{171} Telephone Interview with Peter Barton Hutt, \textit{supra} note 142.
\textsuperscript{172} \textit{See} Schiller, \textit{supra} note 139, at 1145–52.
\textsuperscript{173} \textit{Id.} at 1140–41; \textit{see also} Pierce, \textit{Seven Ways to Deossify Agency Rulemaking}, \textit{supra} note 55, at 59 (elaborating on the benefits of agency rulemaking).
\textsuperscript{174} Though it probably looked more like the process anticipated and designed by the drafters of the APA. \textit{See} Schiller, \textit{supra} note 139, at 1159.
\textsuperscript{175} \textit{See} Administrative Procedure Act, 5 U.S.C.A. § 553(b) (West 2010) (requiring time, place and nature of public rulemaking proceedings, if any, reference to legal authority for the rule, either terms of rule or description of subjects and issues involved).
promulgate the final regulation.”

Under the same rationale of increasing transparency to increase agency accountability and reduce abuses, Hutt implemented two related innovations. First, he required the agency to have a detailed preamble both in the proposed rule and in the final rule. The preamble would have to include:

[A] summary first paragraph describing the substance of the document in easily understandable terms, . . . (vii) supplementary information about the regulation in the body of the preamble that contains references to prior notices relating to the same matter and a summary of each type of comment submitted on the proposal and the Commissioner’s conclusions with respect to each. The preamble is to contain a thorough and comprehensible explanation of the reasons for the Commissioner’s decision on each issue.

In short, it called for a great deal of information to be produced by the agency.

The reform was not universally welcomed by agency staff: “People at FDA were at first blush horrified,” but Hutt was not deterred. “[This] was a form of transparency. We told the American public, ‘this is why we are doing it.’” The courts’ first steps into requiring explanation, occurring at the time, assisted Hutt, and his successor Merrill, to convince the agency that the steps were necessary.

In addition, the preamble was to include the commissioner’s response to the comments submitted, as described in the regulation above, preempting the demands the courts would later apply to the FDA. That is not to say that the agency always responded to comments to the satisfaction of commentators and/or the courts. In

---

176 Telephone Interview with Peter Barton Hutt, supra note 142.
177 21 C.F.R. § 10.40(b)(vii).
178 Id. § 10.40(c)(3).
179 Id.
180 Telephone Interview with Peter Barton Hutt, supra note 142.
181 Telephone Interview with Richard A. Merrill, supra note 145.
the *Nova Scotia* case, for example, the court criticized the FDA’s commissioner for not responding to certain issues raised directly by the Bureau of Commercial Fisheries of the Department of the Interior. But the agency took steps to increase its accountability, even though that added substantial work and at the same time increased the risk that its actions would be challenged.

**B. The EPA Works at Being Accountable**

Like the FDA, the EPA engaged in many efforts to increase its accountability. And just as with the FDA, this Article will focus on only a small sample of these behaviors. The laws governing the EPA were designed to promote accountability through the “fire alarm” approach discussed by the trio of scholars collectively known as “McNollgast.” When using a fire alarm approach Congress creates accountability mechanisms that allow private citizens to take action to challenge or block agency deviation from Congressional mandates, rather than having only “top-down” oversight. Many of the statutes delegating power to the EPA include provisions for citizen lawsuits. One source estimated that in the 1980s, about 80 percent of the EPA’s rules were subject to litigation, and described the EPA as “embattled and embroiled in litigation, threats of litigation and expressions of general dissatisfaction on the part of all of its outside constituencies—industry, environmentalists, and state government.”

In an effort to reduce the dissatisfaction with its programs and

---

183 Id. at 248.
185 For example, the Clean Water Act provision allowing anyone who is or might be affected by violation of the act’s provision to sue a violator. See Clean Water Act, 33 U.S.C.A. § 1365 (West 2010). The Clean Air Act has a similar citizen suit provision. See Clean Water Act, 42 U.S.C.A. § 7604 (West 2010).
to increase its legitimacy, the EPA has engaged in many efforts to increase its accountability by increasing its transparency and the opportunities for public participation, sometimes through its own initiatives and sometimes following political prodding. Since at least the 1980s, the EPA has made substantial efforts to engage the public in dialogue and increase the input of stakeholders. It expresses commitment to accountability in the reports describing its performance and measuring goal achievements. For example, the agency says in its 2008 “Performance and Accountability Report”\(^\text{187}\) that the report “demonstrates EPA’s commitment to be held accountable for results.”\(^\text{188}\) In its “Framework for Implementing EPA’s Public Involvement Program,”\(^\text{189}\) the EPA explained that its goal is “to have excellent public involvement become an integral part of EPA’s culture, thus improving all of the Agency’s decision making.”\(^\text{190}\)

Many examples can be provided, but this Article will describe just three in chronological order: adoption of negotiated rulemaking, the development of the IRIS system, and the 2001 online dialogue.

1. Negotiated Rulemaking

Starting in 1983 the EPA voluntarily conducted a pilot program of negotiated rulemaking procedures following a recommendation by the Administrative Conference of the United States; it was one of the first agencies to do so.\(^\text{191}\) The Negotiated Rulemaking Act\(^\text{192}\)


\(^{188}\) Id. at 2.


\(^{190}\) Id. at 1.

\(^{191}\) Siobhan Mee, Negotiated Rulemaking and Combined Sewer Overflows (CSOs): Consensus Saves Ossification?, 25 B.C. ENVTL. AFF. L. REV. 213, 232 (1997) (suggesting EPA was the first). But see Cary Coglianese, Assessing Consensus: The Promise and Performance of Negotiated Rulemaking, 46 DUKE L.J. 1255, 1263 (1997) [hereinafter Coglianese, Assessing Consensus] (demonstrating the FAA was the first, with the EPA and a few other agencies
was only enacted in 1990, at least partly drawing on the EPA’s experience.\textsuperscript{193} The EPA chose negotiated rulemaking in an effort to reduce the adversarial nature of the regular rulemaking process and especially the litigation that accompanies it, but the choice also reflects the increasing acceptance of negotiation as a form of decision making in the environmental context.\textsuperscript{194}

Two initial negotiations on nonconformance penalties under the Clean Air Act and the criteria for emergency pesticide handling successfully ended in a consensus that was used in issuing the Notice of Proposed Rulemaking.\textsuperscript{195} Participants expressed satisfaction with the process.\textsuperscript{196} This initial success spurred the agency to engage in further negotiated rulemaking.

The EPA engaged in a very systematic effort, creating a project staff that was devoted to identifying appropriate rules for negotiated rulemaking, monitoring and evaluating facilitators, and generally improving the process.\textsuperscript{197} Since 1990, the project’s role was redefined to include general conflict resolution mechanisms.\textsuperscript{198}

By 2000, the EPA had conducted twenty-one negotiated rulemakings, more than other agencies, on a wide range of topics.\textsuperscript{199} It engaged in many more evaluations to see whether certain rulemakings were appropriate for negotiated rulemaking.\textsuperscript{200} While the evaluation of negotiated rulemaking is mixed,\textsuperscript{201} the

\begin{thebibliography}{99}
\bibitem{note192} Negotiated Rulemaking Act, 5 U.S.C. § 561 (West 2010).
\bibitem{note193} Mee, supra note 191, at 216.
\bibitem{note195} Id.
\bibitem{note196} Id. at 30.
\bibitem{note197} Dalton, supra note 186, at 135, 146–49.
\bibitem{note198} Id. at 151–52.
\bibitem{note199} Id. at 135. For some descriptions, see id. at 135–46.
\bibitem{note200} Id. at 151.
\end{thebibliography}
EPA made a clear effort to engage stakeholders in its decision making.

2. The Integrated Risk Information System Database

The EPA’s Integrated Risk Information System database (IRIS) is, as described at the beginning of this paper, a database including assessment of the risks involved with various chemicals. IRIS was adopted by the EPA in 1985, but the major efforts at reforming its procedures started in the mid 1990s.202

After a 1997 review process, the EPA introduced several changes which included, for example, creating a hotline for users, publishing an annual agenda specific to IRIS assessments in the Federal Register, publishing external peer review drafts of IRIS assessments on its website and considering public comments on the drafts.203 In 2004 the agency also added, at the request of the Office of Management and Budget in the White House (“OMB”), a process that allows OMB and other federal agencies to review and comment on assessments; OMB involvement increased over the years.204 Adding the OMB process was controversial and critics attacked it on several grounds. It was seen as adding a political element to what should be a professional endeavor.205 It was seen as an effort by the Bush administration to add delays to the assessment process, as part of a pro-business agenda, since the assessment process is a first step in regulating a given chemical. However, allowing the OMB to review the EPA documents

---

113, 122–24 (1992) (giving positive evaluations), with Coglianese, Assessing Consensus, supra note 191, at 1309–10 (asserting that negotiated rulemaking is not shorter and has no less litigation than regular rulemaking); id. at 1281–1304 (analyzing the data). See also Cary Coglianese, Assessing the Advocacy of Negotiated Rulemaking: A Response to Philip Harter, 9 N.Y.U. ENVTL. L.J. 386, 405–27 (2000-2001) (offering a much more pessimistic assessment) [hereinafter Coglianese, Assessing the Advocacy].

202 U.S. Gov’t Accountability Office, supra note 2, at 6–10.

203 Id. (taken almost, though not completely, verbatim from the report).

204 Id. at 12, 22–23.

increases accountability in two ways: it provides another layer of inter-executive branch review, that is, another layer of bureaucratic accountability, and since OMB is subject to presidential control, this procedure tends to strengthen the President’s control over other agencies. Since the President is the only official in the executive branch directly elected, strengthening his control over the administrative state strengthens political control over the professional civil service, thus strengthening at least one type of accountability.

By 2007, the IRIS assessment process involved the following stages:

1. Before assessing a substance, the EPA would ask the public and other federal agencies or interested parties for nominations.
2. The EPA would list which of the nominated substances would be assessed, and publish that list in their annual agenda, at the same time soliciting scientific information about the listed substances, both from the public and other federal agencies.
3. The EPA would also do its own literature search and create a literature review.
4. The EPA would do a quantitative toxological review.
5. OMB would review the toxological report and distribute it to other agencies for comment. According to the General Accountability Office (“GAO”) report, “OMB informs EPA when EPA has adequately addressed interagency comments.”
6. The EPA would publish the results of the toxicological review and convene a public meeting of external peer reviewers.

206 This description is based on a combination of the GAO report 08-440, see U.S. GOV’T ACCOUNTABILITY OFFICE, supra note 2, at 13, and a descriptive article from the non profit OMB Watch that makes a clear distinction between the process as it was before the 2008 reform described below and the post-2008 process. See White House Gains Influence in Toxic Chemical Assessments, OMB WATCH (April 15, 2008), http://www.ombwatch.org/node/3642.
207 U.S. GOV’T ACCOUNTABILITY OFFICE, supra note 2, at 13.
7. After external review, the EPA would revise the assessment as necessary.
8. A second OMB review would then be conducted, with OMB again disseminating the information to other agencies.
9. Finally, the EPA would post the assessment on its website.

In 2008 this process was strongly criticized by the GAO as too long and inefficient, but it certainly has substantial accountability built in, most of which was voluntarily taken on by the agency (though adding the OMB review seemed to be due to political pressure).²⁰⁸

On April 10, 2008, the EPA changed its process, adding several steps. The new steps substantially increased OMB’s role in the process as well as the opportunity for public comment. For example, at the selection phase, in addition to asking the public for nominations of chemicals, the EPA was required to consult with OMB after receiving nominations to determine which of the substances nominated it would evaluate. Before creating its toxological review, the EPA would prepare a qualitative assessment of the chemical, including potential health risk, susceptible populations, and potential uncertainties. This assessment would then be open to comments from the public and OMB (which would provide the assessment to other agencies). In addition, if another agency demonstrated the chemical to be critical to its mission, that agency could require further study of the substance.²⁰⁹ In other words, the process, while adding to the delay criticized by the GAO, allowed for increased accountability of the process towards the OMB and other agencies as well as increased opportunities for public input.

In May 2009, the EPA’s administrator revised the process once again, announcing the change in a memo to top EPA officials that was also published on the agency’s website.²¹⁰ The goal of the new process was to streamline and simplify the review process. OMB

---

²⁰⁸ Id. at 3–5.
²⁰⁹ White House Gains Influence in Toxic Chemical Assessments, supra note 206.
²¹⁰ Memorandum from Lisa P. Jackson, supra note 6.
would now review only at two stages, before and after the input of the expert peer review. The agency announced that it would lead the new process (previously, the process was coordinated and managed by OMB)—the EPA would give the other agencies opportunity to comment and will meet with them, but it intends to have the final say.\(^{211}\) Other agencies will no longer be able to delay the process to conduct research on “mission critical” chemicals.\(^{212}\) In addition, all written comments from other agencies and the White House were to be made public.\(^{213}\)

On the one hand, the agency made substantial efforts to reduce delays, but on the other hand it increased accountability to the public by providing more information on its decision-making process. Generally speaking, throughout all the reforms, the EPA struggled to balance thorough review of the assessment process with efficiency, sometimes leaning more one way, sometimes more the other; but in all cases, it worked hard to increase accountability, primarily by increasing the input of external actors and thus giving them more opportunities to impact the final decision.

3. The Online Dialogue

In 2001, the EPA engaged in an online dialogue to supplement traditional hearings for comment on its “Draft Public Involvement Policy” (PIP) and on ways that it could be implemented.\(^{214}\) The fact that the EPA conducted such a dialogue at all evidences their search for ways to become more accountable; the content of the dialogue provides additional evidence that the agency was actively

\(^{211}\) Id. at 4.

\(^{212}\) EPA Announces New IRIS Assessment Development Process, ENVTL. PROT. AGENCY, (May 21, 2009), http://yosemite.epa.gov/opa/admpress.nsf/48f0fa7dd51f9e9885257359003f5342/065e2c61afea0917852575bd0064e9db/OpenDocument.

\(^{213}\) Memorandum from Lisa P. Jackson, supra note 6.

Account Me In

trying to improve its accountability.

The 2001 dialogue involved 1,166 members of the public and substantial numbers of the EPA staff. 215 The process started with the circulation of a draft of the PIP for comments in December 2000. 216 The process was open to anyone, and the EPA engaged in substantial efforts to advertise it. Quoting from Beirle, the author of a report on the dialogue:

EPA staff sent announcements via EPA mailing lists and listservs and spread the word through personal contacts and mailings to a wide variety of institutions involved in environmental policy, including environmental organizations, state and local governments, small businesses, and tribal groups. Internally, they distributed information to 1,500 EPA staff—including all coordinators of environmental justice, tribal, communications, and community-based environmental protection programs—with a request to pass on information about the Dialogue to their regular contacts. Information Renaissance publicized the Dialogue through information channels it had developed through previous on-line dialogues. Some people who received announcements about the Dialogue forwarded them through their own networks. 217

About a month before the dialogue, people could register, either as active or passive participants. ("Active" meant one was allowed to both read and post messages, "passive" meant reading privileges only.) 218 There was a dialogue website where the EPA and Information Renaissance posted an electronic briefing book with substantial amounts of materials. 219

Ten of the EPA’s offices held a day of discussion each. Participants could post messages or answer previous messages in a

215 Id. at 8; Patricia A. Bonner et al., Bringing the Public and the Government Together Through On-Line Dialogues, in THE DELIBERATIVE DEMOCRACY HANDBOOK: STRATEGIES FOR EFFECTIVE CIVIC ENGAGEMENT IN THE 21ST CENTURY, supra note 102, at 146–47.
216 BEIERLE, supra note 214, at 15.
217 Id. at 16.
218 Id. at 17.
219 Id. at 18.
Officials responded to “something [sic] that was relevant to their programs.” The material collected was included in the EPA brochures and in a Public Involvement Policy issued in June 2003. A report was prepared after the fact describing the process in detail.

In these three very different examples, the EPA constantly strove to increase input from stakeholders into the process and to increase the transparency of the process. In the IRIS case, at least, the costs of accountability may have outweighed the benefits; but all cases demonstrate the agency’s strong commitment to accountability.

C. The IRS Seeks Accountability

In 1998 Congress passed the Internal Revenue Service Restructuring and Reform Act of 1998, substantially changing the regulatory environment of the IRS. One of the goals of the act was to increase the agency’s accountability. The reform was passed amid substantial accusations of lack of accountability. In a typical example, Senator Grassley said the IRS:

“[S]eems to be squeezing the little guy to get the money while this set of four witnesses are telling us that the big tax liability is often forgiven. And the cause of this, of course, I think, and it is the basic unfairness is the lack of accountability.”

---

220 Id. at 17.
221 Bonner, supra note 215, at 147.
222 Id. at n.21; Online Publication Title List, U.S. ENVTL. PROT. AGENCY, http://nepis.epa.gov/EPAPub/EPAPubTitleOther.htm (last updated Sept. 14, 2010) (providing brochures on public involvement, numbered 233F03005-12, 233F03014).
224 BEIERLE, supra note 214, at 8.
226 E.g., Henning, supra note 135.
227 IRS Oversight, supra note 38 (statement of Sen. Charles E. Grassley,
Similarly, Senator Frank Murkowski from Alaska said: “[W]e agreed that there was no accountability in the IRS. We agreed that the system was designed to avoid accountability.” To correct that, among other things, the Act made customer service one of the main goals the IRS should aspire to, and added a variety of accountability mechanisms.

It is of course true that this reform was imposed from the outside by Congress, and had input from some members who were actively hostile to the IRS. However, the IRS Commissioner, Charles O. Rossotti strongly supported reforming the agency and had substantial input into some of the provisions of the Act. Rossotti continued to express commitment to the reform, and under his direction the IRS engaged in substantial efforts to increase its accountability, efforts that continued under Commissioner Shulman. These reforms paralleled and reinforced ongoing efforts of officials inside the IRS, which had begun before the reform, to increase its transparency and responsiveness.

This section will mention two examples: the Taxpayer Advocacy Panel, which allowed a panel of citizens substantial input into the IRS’s operations, and the efforts to increase the accuracy and transparency of the IRS policies. This is a shortened description; both efforts deserve a much more detailed treatment, but that is a matter for another article.

---

228 Id. at 9 (statement of Sen. Frank S. Murkowski, Alaska).
229 See Thorndike, supra note 135, at 768. Some of these were quite draconian; for example, the “ten deadly sins” provision decreed that IRS employees will be fired if they violated one of its ten vague provisions (e.g. violating any of the internal Revenue Manual provisions). See Rainey & Thompson, supra note 9, at 599.
230 See Thorndike, supra note 135.
231 Id. at 775–77; Rainey & Thompson, supra note 9, at 597.
232 Telephone Interview with IRS official (under promise of confidentiality) (June 24, 2009).
233 Id.; Telephone Interview with IRS official (under promise of confidentiality) (July 1, 2009); Telephone Interview with Deborah Gascard Wolf, Director of the Office of Privacy, Information Protection and Data Security, IRS (July 14, 2009).
1. Taxpayer Advocacy Panel

In 2002 the Department of the Treasury, together with high IRS officials and the Taxpayer Advocate, Nina Olson, created the Taxpayer Advocacy Panel.234 The panel is a collection of one hundred volunteer citizens, organized into a number of issue committees, who contribute three hundred hours each to reviewing the IRS activity and offer recommendations for improvements. The reports that come out of this activity include very detailed recommendations by the panel and a response—often detailed—from the relevant IRS official. The response can be adoption of the panel’s recommendations,235 a promise to consider them,236 or

---

234 See What We Do, TAXPAYER ADVOCACY PANEL, http://www.improveirs.org/about-us/ (last visited Jan. 28, 2011) (describing the panel). According to the TAP’s first annual report from 2003, it was created when [t]he Department of the Treasury, in response to a review of Federal Advisory Committee Act Boards, recommended nation-wide expansion of the Citizen Advocacy Panel established in June 1998, to be renamed the Taxpayer Advocacy Panel (TAP). The Internal Revenue Service (IRS) established a Design Steering Committee comprised of the National Taxpayer Advocate, Executives from Wage & Investment (W&I), Small Business/Self Employed (SB/SE), the Communications and Liaison Office, and National Treasury Employees Union Representatives to design the new Panel.


The Committee recommended and the OTBR [Office of Taxpayer Burden Reduction. D.R.] accepted the recommendations of removing, modifying and/or consolidating lines on the Form 2678 and on the Schedule R (Form 941). The OTBR also accepted the recommendation to make some minor changes to the verbiage on both forms. A major change to the Form 2678 was to have both employer and agent’s signature on the Form as opposed to only having the employer’s signature as in the past.

236 See, e.g., id. at 7:

Form 12153 is a critical part of Collection Due Process that begins the
rejection with an explanation.\(^{237}\)

To give one example, as part of the effort to protect citizens’ privacy by reducing the availability of social security numbers, the IRS is working on regulations that will allow issuers of form 1099\(^{238}\) to only provide some digits of the social security number rather than the entire number. That decision, explained Deborah Wolf, Director of Privacy, Information Protection and Data Security in the IRS, was the result of a recommendation by the Information Reporting Advisory Committee, a committee of the Taxpayer Advocacy Panel.\(^{239}\) The committee recommended treating the W2 forms provided by employers in a similar fashion, but that requires a change in legislation, which will take longer.\(^{240}\)

The creation and continuing existence of the Taxpayer Advocacy Panel shows in two ways how the two agencies (the IRS and the Treasury) are actively working to increase their accountability. First, the Treasury Department acted voluntarily to create the panel (with collaboration by the IRS). Second, the IRS regularly engages in dialogue with the panel, with top officials responding to panel recommendations in ways that can include changes in policies in accordance with what the panel has recommended—again, opening themselves to criticisms of their response and to input from stakeholders.

\(^{237}\) As was done for one of the recommendations in TAP A07-4066. See id. at 8–9.


\(^{239}\) Telephone Interview with Deborah Gascard Wolf, supra note 233.

\(^{240}\) Id.
2. Increasing Transparency and Accuracy of the Internal Revenue Manual

The core of this discussion is the IRS’s effort to increase its transparency through making its policies more easily accessible. One of the criticisms raised against the agency in the discussions leading to the 1998 reform was the lack of transparency of its policies, which sometimes had the effect of making life very difficult for citizens who had been accused of non-compliance due to errors. Even before that, the IRS had been working on improving the transparency in its policies and procedures. However, it seems—though the causality is hard to trace—that the criticisms raised during the reform gave the project of increasing transparency an extra push. The IRS expressed a commitment to making the Internal Revenue Manual (“IRM”) reflect current policies (not easy for an extremely large agency with an annual deadline) and has taken a series of steps in that direction. The Office of Servicewide Policy, Directives and Electronic Research (“SPDER”), created in 1999, engaged in a series of initiatives to oversee and coordinate transparency, including sending out memoranda reminding staff of the need to make interim guidance available to the public and to train staff accordingly. Guided by SPDER, the IRS units created and implemented internal procedures to post interim guidance memoranda electronically, a process monitored by SPDER.

As a first step, SPDER worked to change the format of the IRM from paper to electronic and to restructure it so that it was organized by processes rather than by the IRS organizational units. The change to an electronic format was crucial to achieve accuracy and accessibility. An electronic format was easier to search, and thus made it easier to find

---

241 See generally Henning, supra note 135.

242 The description is taken from the comments of IRS officials to the Taxpayer Advocate’s criticisms of their lack of transparency. Nat’l Taxpayer Advocate Serv., 2006 Annual Report to Congress 24–26 (2006), http://www.irs.gov/pub/irs-utl/2006_arc_vol_1_cover_section_1.pdf. The substance of the descriptions has been acknowledged by the National Taxpayer Advocate. Id. at 27.
inconsistencies when searches came up with a variety of results.\textsuperscript{243}

Prior to the changes, the IRM was organized by unit, by the title of the office, not by the process in question—which made it easy for office members to update, but hard for people outside the IRS to know their way around. It also made it very easy for inconsistencies to be generated in relation to specific processes and for inconsistencies to persist. This first step presented both an administrative challenge—it was a large reorganization—and an educational challenge: members of the staff had to be convinced that the change was necessary, or at least inevitable. Some resistance was encountered because many employees strongly identified with their particular units.\textsuperscript{244}

SPDER also invests substantially in training IRM authors in how to write the manual in a transparent and easy to use way, offering aggressive (though voluntary) training programs.\textsuperscript{245} The goal is to write the IRM following information mapping principles, active voice, and plain language.

Another change was a revision of the policy governing disclosure of materials which are intended only for internal use.\textsuperscript{246} Prior to the mid 2000s, the agency’s policy was not to publish documents if any part of that document was for official use only. SPDER initiated a redaction process whereby paragraphs that are for internal use only are removed from documents, and what remains is published, substantially increasing the amount of publicly available material.\textsuperscript{247}

As can be seen from this very short description, substantial efforts were made by the IRS to increase the transparency and accessibility of its policy.

\textsuperscript{243} Telephone Interview with IRS official (under promise of confidentiality) (July 1, 2009); Telephone Interview with IRS official (under promise of confidentiality) (June 24, 2009).

\textsuperscript{244} Telephone Interview with IRS official (under promise of confidentiality) (July 1, 2009).

\textsuperscript{245} Id.

\textsuperscript{246} Which may fall under any of a number of exemptions to FOIA. See Freedom of Information Act, 5 U.S.C.A. §§ 552(b)(2), (5), (7) (West 2010).

\textsuperscript{247} Telephone Interview with IRS official (under promise of confidentiality) (July 1, 2009). A more detailed description of the IRS reforms is in preparation.
IV. IMPLICATIONS AND DISCUSSION

A. What is NOT Implied?

Let’s start with what this Article is not intended to imply. I am not saying that the need for external accountability mechanisms is obviated by the discovery that agencies often will be accountable on their own. Most of the efforts to increase accountability described in this Article were undertaken by agencies facing the vast array of accountability mechanisms described in Part I, and their actions were taken in the context of those mechanisms. Further, as the case studies demonstrate, the agencies making these efforts were agencies “under fire”—they were already being subjected to substantial criticism for, among other things, lack of accountability. The IRS already faced the experience of accountability mechanisms added, and massive reorganization imposed, because of accusations of lack of accountability that stuck.248 It is unclear whether the agencies would have made the same kinds of efforts to be accountable without external accountability pressures, but a strong argument can be made that at least in part their efforts to increase accountability were motivated by the desire to head off more external attempts to do it for them. It does not require evil intent for agencies not to undertake such efforts: agencies have many other tasks besides being accountable, and without strong motivators to invest in accountability, these can easily (and possibly with good reason) take precedence.

Beyond that, even an agency that is strongly committed to accountability because of its sense of mission may not construct accountability mechanisms in ways that are valued by either its political masters or the public in general. Those outside the agency may have different preferences than the agency as to what form the accountability should take, and how it should be structured, and without external mechanisms they may not be able to influence an agency’s choice. For example, the EPA’s design of the accountability framework surrounding IRIS was strongly criticized by the GAO as damaging to its efficiency, and by OMB Watch as

248 See supra Part III.C.
inserting a political component into a professional decision.\textsuperscript{249} The FDA’s decision to summarize and publish information about its decisions on new drug applications was criticized by industry members for providing too much information and harming the industry, and by consumer protection organizations for not providing enough data.\textsuperscript{250}

Even if the motivations of the agency are accepted as valid, and an agency strongly buys into the accountability language, it may not be especially well versed in the tools and mechanisms available. An agency may lack expertise and not do a particularly good job of increasing its accountability even if it is extremely professional in other areas. For example, while the IRS has been making efforts to increase the transparency of its policies for several years, the 2006 Taxpayer Advocate Service Annual Report highlighted certain problems, such as internal memos not published.\textsuperscript{251} Hutt sees the FDA as working to increase its transparency, but as demonstrated in Part II.A, external observers criticize it for lack of transparency. External direction may be required to help agencies steer their accountability choices.

Finally, as demonstrated by the example of the Minerals Management Service,\textsuperscript{252} not all agencies seek accountability—or are accountable—all the time, and to prevent extreme cases of abuses, external mechanisms are crucial.

The other thing I am not asserting in this Article is that efforts by agencies to increase accountability are generally successful. Assessing success of accountability mechanisms requires two complex, challenging inquiries, neither of which I undertake here. The first is a value judgment as to just what a successful accountability mechanism would look like. For example, suppose it is a mechanism that increases input from outside parties; if that proves so effective that it gives stakeholders or citizens full control, is that outcome desirable, or not? Substantial input from

\textsuperscript{249} See supra Part III.B.

\textsuperscript{250} See supra Part III.A.


\textsuperscript{252} See supra Part I.A.
stakeholders can be seen as an agency being responsive or as an agency captured; in the matter of public participation, some scholars strongly support extensive participation, others are concerned that it may harm expertise and decision making. As for publication, the benefits and costs of transparency are also debated. If increasing transparency is a good, should agencies provide more information or provide information in a more simplified, easy to access way?

Agreeing on some standardized benchmarks is one challenge to making a scholarly assessment of success at improving accountability; another challenge is presented by the wide range of empirical measurement problems that exist—experimental design and data collection present serious difficulties in real-world situations. For instance, if we agree that public participation and input of citizens into the process is a good, how do we measure whether there actually was input? Do we look at the number of participants and the number of comments they made, as was done in studies evaluating participation in rulemaking? This approach can be challenged on the grounds that “comments” can be nothing but answers checked off on standardized forms, or that merely submitting comments does not mean that any effective input to the process occurred. Or, alternatively, do we look at ways the agency may have changed its views after receipt of the comments? This is another measure some of the same scholars used. Here, we can ask whether “change” is for the better or not—change is not

253 See generally James S. Fishkin, DEMOCRACY AND DELIBERATION: NEW DIRECTIONS FOR DEMOCRATIC REFORM (1991); Leib, supra note 101; Arnstein, supra note 101; Innes & Booher, supra note 101.
257 See Reiss, Tailored Participation, supra note 20, at 331–35; Schlosberg, supra note 103.
always justified. A more general reason this Article does not address the question of assessment is that while there is room and need for articles examining the success of efforts at increasing accountability, this article has a different focus. Accordingly, this Article is not being presented as evidence that agencies do a good job at being accountable—just that they try.

B. What is Implied?

In this Article I intend to show that agencies are not always the enemy in the “accountability game” and are never just a pawn either. First, as to agencies not being the enemy—as this Article demonstrates, important agencies make substantial efforts to increase their accountability and agency officials often buy into reforms aimed at increasing accountability. Several of the mechanisms later adopted by Congress or the courts were initially tested or adopted by agencies. For example, the FDA adopted a preamble requirement and a requirement of answering comments independent from judicial review; in addition, various agencies experimented with negotiated rulemaking before Congress passed an act that mandated it.

The case studies suggest that a variety of motivations influence agencies to make these efforts. For example, a typical case is that of an agency under attack trying to improve its accountability so as to reduce the severity of the attacks and simultaneously make real improvements in performance. All three agencies that I report on in this Article acted to increase their accountability after being strongly attacked. This would support the view that agencies are

258 Several studies have already attempted to analyze the actual impact of comments. See, e.g., Cuéllar, supra note 256, at 433–34; Golden, supra note 104, at 261; West, supra note 14, at 68; Jason Webb Yackee & Susan Webb Yackee, A Bias Towards Business? Assessing Interest Group Influence on the U.S. Bureaucracy, 68 J. Pol. 128 (2006).

unwilling participants in the accountability game. However, the extent of agencies’ efforts and their level of innovation suggest that a desire to avoid punishment was not the only thing driving the agencies—a level of commitment to the idea of accountability, and some effect of the ideas of participation and transparency, also exists.

My second major point is that viewing agencies as the “subject” is a mistake. Much of the accountability literature sees agencies as being “acted upon,” as primarily responding to accountability mechanisms put in place by others. The famous McNollgast studies address how Congress can shape agencies’ environment to assure compliance with Congressional preferences.\(^{260}\) Studies of judicial review focus on behavior and how courts evaluate agencies,\(^{261}\) rather than seeing the agency and courts as co-participants in the game, with the agency having substantial power to affect how a court will review it.\(^{262}\) Agencies are typically sophisticated political actors.\(^{263}\) They can and do

---

\(^{260}\) McCubbins, Administrative Procedures, supra note 16, at 243–45; McCubbins, Structure and Process, supra note 16.


\(^{262}\) With some exceptions—one possible exception is O’Reilly, Losing Deference, supra note 17, at 949–50, 977–78 (loss of deference to FDA is largely because it became captured by political actors). But see David C. Vladeck, The FDA and Deference Lost: A Self-inflicted Wound or the Product of a Wounded Agency? A Response to Professor O’Reilly, 93 CORNELL L. REV. 981, 983–85 (2008) (offering an opposing view).

\(^{263}\) See, e.g., CARPENTER, supra note 57, at 19–25; Reiss, Agency Accountability, supra note 46, at 114–15; Dorit Rubinstein Reiss, Administrative Agencies as Creators of Administrative Law Norms: Evidence from the UK, France and Sweden, in COMPARATIVE ADMINISTRATIVE LAW 373–74 (Susan Rose-Ackerman & Peter Lindseth eds. 2010) (unpublished copy on file with author); JAMES Q. WILSON, BUREAUCRACY—WHAT GOVERNMENT AGENCIES
Account Me In

influence their accountability environment. Agencies do respond to other actors, but they also tailor their behaviors in ways that will help achieve the best accountability environment they can, given the political and institutional constraints they face. 264 One way agencies can influence their environment is by making efforts of their own to increase their accountability, thus preempting external efforts. This suggests a number of important policy implications for current administrative law and practice.

An important question is how to incentivize agencies to undertake more efforts to increase their accountability—assuming that increased accountability is considered necessary. 265 If accountability is a positive, more accountability is better than less, and we want agencies to be more accountable. But if that is the case, maybe agencies should be rewarded for their efforts at increasing accountability. Lack of incentives may lead agencies to work to improve accountability less often than is desirable. For example, a plausible argument is that the relative rarity of negotiated rulemaking undertaken by agencies can be explained by a lack of incentives. Several studies suggest that even if the participants are satisfied with the process, 266 from the agency’s point of view there may not be sufficient incentive, since it still has to go through the regular notice and comment procedures anyway. At any rate, in spite of efforts by the Clinton administration to increase the use of negotiated rulemaking, 267 the number of such

---

264 See Reiss, Agency Accountability, supra note 46, at 114–15.
265 Not at all obvious—as the “Agencies as Victims” literature described in Part I.B demonstrates, a strong claim may be made that there is enough, if not too much, accountability already in the administrative state. However, the claim for more accountability seems to be stronger—and even if we do not think MORE accountability is a positive, better accountability may be—more streamlined and efficient tools that achieve the same effects.
267 Harter, Assessing the Assessors, supra note 201, at 36–37.
rulemakings remains very small. 268 One incentive may be reducing some of the external controls for agencies which engage in their own processes—following the same logic as the EPA’s voluntary compliance plans that carry with them the suggestion of reduced enforcement against participants. 269 This will probably require legislation. Another incentive would be to allow agencies brought to court with claims of lack of participation or transparency to present evidence of efforts to increase their accountability and provide a holistic picture of accountability.

One way for a court to deal with a claim of insufficient participation or transparency is to charge the agency with producing a plan for increasing accountability, thus taking advantage of the ability of agencies to create their own accountability mechanisms. That way, the court does not directly dictate new procedures—forbidden under Vermont Yankee 270—but can still address a lack of sufficient procedures. Courts may be less reluctant to allow agencies to use guidelines 271 if the agencies


The phenomenon we see in this case is familiar. Congress passes a broadly worded statute. The agency follows with regulations containing broad language, open-ended phrases, ambiguous standards and the like. Then as years pass, the agency issues circulars or guidance or memoranda, explaining, interpreting, defining and often expanding the commands in the regulations. One guidance document may yield another and then another and so on. Several words in a regulation may spawn hundreds of pages of text as the agency offers more and more detail regarding what its regulations demand of regulated entities. Law is made, without notice and comment, without public participation, and without publication in the Federal Register or the Code of Federal


present a scheme for making sure those guidelines are transparent and allow for stakeholder input into their making. Guidelines have a number of advantages over rules: they are more flexible, and therefore may be more suitable to fast-changing environments, they do not suffer from the same degree of ossification, and they allow—in fact, mandate—agency deviation in appropriate specific cases. Allowing guidelines an official place under these narrow circumstances is especially useful since the jurisprudence about policy documents versus rules is vague and the criteria to distinguish them are unclear. If an agency created a document that could fall under either category through a process that either closely follows informal rulemaking procedures or goes beyond them and adds substantial accountability guarantees, why invalidate it?

This need for incentives is especially true since one of the challenges any agency head will face is that of competing demands on time; achieving accountability requires time and effort. For example, in the case of the FDA, when confronted with changes in the agency’s enabling act along with demands for investigation of alleged misrepresentations of the results of clinical trials, drafting the Medical Device Act of 1976, and a large range of other missions, promoting accountability was not always the first priority. Similarly, updating the Internal Revenue Manual to keep IRS policies transparent may not be the first priority for an official burdened with many other chores. One IRS official explained that it is seen as “record keeping,” and often with “so


274 Telephone Interview with Richard A. Merrill, supra note 145.
many other things to do, if they look at a list of priorities, IRM and updating it falls to the bottom of the list. According some deference to an agency engaged in accountability seeking can help balance those costs and provide an incentive for such behavior.

In addition to creating incentives, acknowledging that agencies engage in increasing accountability can open the door to more systematic study of such activity, as suggested by Magill in the context of self-regulation. Besides providing a promising source of information and thought for researchers trying to understand agency action, it can make a wealth of information more systematically available that will be useful for policy makers in all three branches. If one of the justifications for federalism is experimentation in different forms of democracy to allow innovation and the testing of ideas, the same can be said for experimentation with accountability mechanisms among agencies: a type of “administrative federalism.” If agencies are allowed, even encouraged, to experiment, all three branches can benefit.

The branch with the most natural access, and the one already benefitting from such experimentation, is the executive branch. Agencies learn from each other—for example, other agencies emulated the experience of the first agencies with negotiated rulemaking, and OMB incorporated it into an executive order, recommending it to all agencies.278

This learning can be made even more systematic if we learn from other countries’ experience. For example, in Australia, there is an “Office of Best Practice Regulation” to study how to make regulation better, including increasing its transparency. Another potential international source of inspiration is the Organisation for

---

275 Telephone Interview with IRS official (under promise of confidentiality) (July 1, 2009).
276 See Magill, supra note 59, at 860–61.
Economic Co-operation & Development’s (OECD) report comparing best practices across countries.280

Similarly, examining accountability mechanisms created by agencies could provide Congress with ideas for new accountability mechanisms and with real-world data as to what worked and what did not, thus facilitating the making of informed decisions when drafting legislation.

Finally, knowledge of agency practices could guide courts as to where to put—and where to avoid—emphasis. If a mechanism were found to have been widely adopted by agencies and seemed to be working, that could justify increased deference.281

Just as important, agencies engaging in increasing accountability require training and resources to do so. For example, in an age where transparency is increasing through use of online mechanisms, agencies need training in the use of information technology, and in many cases personnel with new kinds of abilities. Some agencies have already started hiring people with new skills:

[SPDER] hired a college professor to help with curriculum and the tax forms[,] . . . someone with background in lobbying . . . [, and] an IT person, working on the move to an electronic environment. . . . I would not have thought I would hire them, but as times change we need people with different skills.282

But for a variety of reasons, including budgetary constraints, this may not have been enough. These needs must be taken into consideration.


281 I suspect some of the requirements placed on agencies by courts over the years were actually ideas borrowed from existing agency practices, either in or across agencies. But that is a topic for another project and requires further research.

282 Telephone Interview with IRS official (under promise of confidentiality) (Aug. 3, 2009).
CONCLUSION

This Article suggests that agencies are accountable, and furthermore, that many agencies want to be accountable and make efforts in that direction. It suggests that agencies’ efforts to increase their accountability are not sufficiently noticed or acknowledged. In relation to their accountability behavior, agencies are either criticized or pitied in most of the current literature. Both approaches are too simplistic. In terms of criticism, being an administrator in the United States is an exceedingly difficult job. The United States is large and complex, which makes managing any problem difficult. Its political system is decentralized, which adds another layer of complexity—as if managing services or regulation for three hundred million people would not be enough, it needs to be done in a fragmented system with many veto points. And like most modern states, it engages in many different spheres of activities. This already makes an administrator’s job hard. In addition, administrators in the United States are everyone’s favorite whipping boy and routinely criticized, among other things for their lack of accountability.

This Article suggests that these claims are at least one-sided and ignore an important part of the picture. The accusation that bureaucrats are unaccountable and seek to avoid the public eye is, at best, of limited validity, and at any rate, requires proof before it can be made. In many situations, it is just not true.

The concern about agency victimization and too much accountability also focuses only on part of the picture, treating agencies as lacking control of their environment and ignoring their contribution to the accountability reality.

This Article is an attempt to give a fuller picture. Agencies also contribute to their accountability environment, by acting, at times, as “accountability entrepreneurs” and at other times as accountability brokers.

Does that mean we do not need to worry about agency

283 KAGAN, supra note 48 at 42.
284 See MEIER, supra note 13, at 2–5.
285 See examples supra Part I.A.
accountability? Unfortunately no. First, efforts by agencies to be accountable may not be effective, or may aim at the wrong problems. So, even when we do not need to worry about the motives of agencies in relation to accountability, we still need to worry about the design of accountability. Second, agencies still face the problem of multiple principals, and their priorities may differ from those of the Congress or the public. Assuming that legislative supremacy is ingrained and required under the American constitutional scheme, efforts may be necessary to align agency accountability behavior with congressional preferences. Finally, not all agencies will make these efforts, and those that do may not make them all the time.

Even so, taking all of the foregoing caveats into account and giving each its full weight, it still seems clear that we cannot ignore agencies’ efforts to increase their own accountability.

---

286 See Furlong, supra note 97, at 61.
ON THE ROAD TO CIVIL GIDEON:
FIVE LESSONS FROM THE ENACTMENT OF
A RIGHT TO COUNSEL FOR INDIGENT
HOMEOWNERS IN FEDERAL CIVIL
FORFEITURE PROCEEDINGS

Louis S. Rulli*

INTRODUCTION

Almost a half century ago, the Supreme Court unanimously held in *Gideon v. Wainwright* that a person accused of a crime could not be assured a fair trial unless counsel was provided to him. The following day, *New York Times* journalist Anthony Lewis reported that the Court had just handed down one of the most important decisions ever in the criminal law field. The *Gideon* ruling overturned long-standing precedent established in *Betts v. Brady* that the Sixth Amendment’s guarantee of counsel did not apply to the states, except in cases involving a death sentence or special circumstances.

The *Gideon* decision prompted national leaders to question whether the Court’s landmark ruling should apply to civil proceedings when the poor’s most vital interests were at stake. On Law Day in 1964, just one year after *Gideon*, Attorney General

---

* Practice Professor of Law and Director of Clinical Programs, University of Pennsylvania Law School. I would like to thank Benjamin Meltzer of the Biddle Law Library at Penn Law School for his valuable assistance. And, in a special note of thanks, I would like to express my deep appreciation to two wonderful and talented Penn Law students, Andrew Sokol and Brandon McCoy, who ably assisted me in completing this project.


Robert F. Kennedy delivered an inspiring challenge to the nation in an address at the University of Chicago Law School in which he stated: “We have secured the acquittal of an indigent person—but only to abandon him to eviction notices, wage attachments, repossession of goods and termination of welfare benefits.”\(^3\) With growing acknowledgment that the phrase “equal justice under law” inscribed on the Supreme Court was incomplete without a right to counsel in basic civil matters, the civil \textit{Gideon} movement was born.

It was also at this time that President Lyndon Johnson launched the War on Poverty, which for the first time provided federal funding to local legal services programs through the Office of Economic Opportunity. The purpose of the program was to have legal services lawyers for the poor “do no less for their clients than does the corporation lawyer checking the Federal Trade Commission for sloppy rulemaking, the union lawyer asking Congress for repeal of 14(b), or the civil rights lawyer seeking an end to segregation in bus stations.”\(^4\) In just a few short years, this new source of funding resulted in remarkable success for the poor in a succession of cases decided by the Supreme Court.\(^5\) These litigation victories demonstrated the profound impact lawyers could have when provided to represent the poor in civil matters. In turn, these favorable results generated substantial optimism that \textit{Gideon’s} core principle of fundamental fairness might soon be


extended to civil proceedings.

This optimism drew to an abrupt halt with the Supreme Court’s ruling in *Lassiter v. Department of Social Services of Durham County.* In *Lassiter*, the Court held that due process of law did not require the appointment of counsel for an indigent mother facing the government alone in civil court proceedings brought to involuntarily terminate her parental rights to her son. For many years after *Lassiter*, the civil *Gideon* movement was dormant in the face of increasingly conservative federal courts that appeared hostile to expanding fundamental rights.

Today, however, almost one-half century after *Gideon*, there is renewed optimism that a civil right to counsel to protect basic human needs is indeed possible and may even be relatively close at hand. With strong support from the organized bar and a coalition of diverse interests, there is a flurry of robust experimentation in the states reflecting intense determination to establish new standards in expanding access to counsel for the poor. Perhaps, most importantly, there is a growing acceptance across the land that fundamental fairness in our civil justice system requires much more than what the Supreme Court was willing to mandate in *Lassiter*.

The civil *Gideon* movement moves forward today largely in state and local legislatures, state courts, bar sponsored pilot programs, and in the court of public opinion. Over time, advocates hope that successful outcomes in state and local venues will effectively eliminate the *Lassiter* rule and create a favorable climate for the Supreme Court to reconsider and overrule its holding in *Lassiter*, just as it did in *Gideon* when the time was right to reconsider its prior holding in *Betts v. Brady*.

Part I of this Article briefly reviews the Supreme Court’s decisions in *Gideon* and *Lassiter*. Part II examines recent developments within the states and among numerous bar associations that have pumped new life into the civil right to counsel movement. Part III takes a close look at reform legislation enacted by Congress in 2000 which established a right to counsel

---

7 *Id.* at 32–33.
at public expense for indigent homeowners whose primary residences are the subject of federal civil asset forfeiture proceedings. The Civil Asset Forfeiture Reform Act of 2000 (“CAFRA”) enacted a broad range of important forfeiture law reforms, including a statutory right to counsel for indigents through court appointments directed to the Legal Services Corporation. CAFRA’s right to counsel in these civil proceedings holds important lessons for the current civil right to counsel movement and Part IV concludes by identifying and describing five specific lessons that can be drawn from that successful legislative effort.

I. EARLY EFFORTS TO EXTEND A RIGHT TO COUNSEL TO CIVIL PROCEEDINGS

A. The Story of Clarence Gideon

Clarence Gideon, too poor to hire an attorney, was denied a lawyer to defend him against burglary charges brought against him by the state of Florida.\(^8\) Although the Sixth Amendment guarantees an accused the right to counsel in federal criminal proceedings, the Supreme Court had long held that the U.S. Constitution offered no similar guarantee in state criminal proceedings.\(^9\) On the basis of state law, a Florida trial judge denied Gideon’s respectful request for the appointment of counsel, stating to Mr. Gideon in open court:

I am sorry, but I cannot appoint counsel to represent you in this case. Under the laws of the State of Florida, the only time the court can appoint counsel to represent a Defendant is when that person is charged with a capital offense. I am sorry, but I will have to deny your request to appoint

\(^8\) Bruce R. Jacob, *Memories of and Reflections About Gideon v. Wainwright*, 33 STETSON L. REV. 181, 200 (2003). Gideon was charged with breaking and entering a pool room with the intent to commit a misdemeanor on June 3, 1961. This charge was a felony under Florida state law. *Id.* at 201. Gideon was reportedly seen leaving the pool hall with a bottle of wine and his pockets filled with coins. *Id.* at 208.

counsel to represent you in this case.\textsuperscript{10}

Without a lawyer to assist him, Gideon was convicted at trial based largely on shaky eyewitness testimony. Sentenced to five years in prison, Gideon filed a habeas petition to the Florida Supreme Court seeking to attack his conviction on the grounds that he was wrongfully denied assistance of counsel.\textsuperscript{11} After the Florida Supreme Court denied his petition, Gideon mailed a petition for certiorari to the U.S. Supreme Court asking the high Court to accept his case and review whether the denial of counsel was unconstitutional.\textsuperscript{12} The Supreme Court granted review to consider whether its prior holding in \textit{Betts v. Brady} should be reconsidered.\textsuperscript{13}

Notably, one of the first things the Supreme Court did after granting review of Gideon’s case was to appoint a lawyer to assist Gideon in his appeal.\textsuperscript{14} The Court chose a highly respected lawyer, Abe Fortas, to undertake this important task. Although the Court gave no formal reasons for its selection, it may be safely assumed

\textsuperscript{10} Gideon v. Wainwright, 372 U.S. 335, 337 (1963). Arguably, the Florida judge was not correct when he stated that he could only appoint counsel in a capital case. He could have read the \textit{Betts} decision to permit an appointment of counsel in noncapital felony cases under circumstances where such an appointment of counsel is necessary to provide a fair trial to Gideon. See Jacob, \textit{supra} note 8, at 202.

\textsuperscript{11} Jacob, \textit{supra} note 8, at 212–13.

\textsuperscript{12} \textit{Id.} at 214.

\textsuperscript{13} \textit{Betts v. Brady} involved a request for counsel by an accused who was indicted for robbery in Maryland state court. \textit{Betts}, 316 U.S. at 456–57. Betts was too poor to afford a lawyer and his request for appointment of counsel was denied on the basis that Maryland law did not require appointment of counsel except in murder or rape cases. \textit{Id.} at 457. Betts represented himself, was found guilty, and was sentenced to eight years in prison. \textit{Id.} Upon review, the Supreme Court held that the trial court’s refusal to appoint counsel for Betts did not necessarily violate due process guarantees. Using a “totality of the facts” analysis, the Court treated due process as less rigid and more fluid than other guarantees of the Bill of Rights, and held under the facts of the case that the denial of counsel was not offensive to common and fundamental understandings of fairness. \textit{Id.} at 461–73.

that the Court believed that Fortas had the intellectual firepower and ample resources to effectively represent Gideon and, equally importantly, to fully develop the important issues for a proper disposition by the Court.  

Fortas brought powerful advocacy to bear on behalf of Gideon. In his brief to the Court, Fortas argued that the Fourteenth Amendment required that counsel be appointed for an indigent defendant in every criminal case involving a serious offense because the aid of counsel is indispensable to a fair hearing. Urging the Court to reverse its holding in Betts v. Brady, Fortas’ brief concluded with a passage originally written by Erwin Griswold and Benjamin Cohen in a letter to the editor of the New York Times published shortly after the Supreme Court’s ruling in Betts:

[A]t a critical period in world history, Betts v. Brady dangerously tilts the scales against the safeguarding of one of the most precious rights of man. For in a free world no man should be condemned to penal servitude for years without having the right to counsel to defend him. The right to counsel, for the poor as well as the rich, is an indispensable safeguard of freedom and justice under law.  

15 Gideon’s request to the Supreme Court was for a “competent attorney to represent [him] in this Court.” BERNARD SCHWARTZ, SUPER CHIEF 459 (1983). At the Court’s conference, Chief Justice Warren suggested that Abe Fortas be appointed. Id. See also ANTHONY LEWIS, GIDEON’S TRUMPET 47 (1964), which briefly discusses the appointment of counsel for Gideon, noting that former law clerks to the justices are often appointed, as are law professors and established practitioners. But, “like other matters decided by the Supreme Court, the choice of a lawyer for an indigent petitioner is entirely in the bosom of the justices,” Lewis writes. Id. The appointment of Fortas has double significance. First, the appointment of any counsel reaffirms that in the proceedings before it, the Supreme Court believes that indigent petitioners should be represented. Second, it is not simply a matter of having any counsel; the Court values the participation of a skilled lawyer who can assist the Court in its decision making while advocating for a client.


17 Id. The letter to the editor was dated July 29, 1942 and published on August 2, 1942 in direct response to the Supreme Court’s ruling in Betts v.
On the Road to Civil Gideon

At oral argument on January 15, 1963, Fortas told the justices that “a common man with no training in law cannot go up against a trained lawyer and win; you cannot have a fair trial without counsel.”

Legal assistance provided to Clarence Gideon in his appeal undoubtedly made a difference. In March 1963 a unanimous Supreme Court overruled *Betts* and held that “in our adversary system of criminal justice, any person haled into court, who is too poor to hire a lawyer, cannot be assured a fair trial unless counsel is provided for him.” This basic proposition was deemed so fundamental that the justices wrote in their opinion that “this seems to us to be an obvious truth.”

Today, it is tempting to view the Court’s landmark decision as simply a quick stroke of the judicial pen resulting in sweeping constitutional change. Such an interpretation, however, would not do justice to the difficult battles that preceded the Court’s ruling in *Gideon* and made possible the overruling of longstanding court

---

18 See generally, LEWIS, supra note 15. Fortas argued that no person, however intelligent and smart, could be expected to represent himself effectively and that even Clarence Darrow felt he needed a lawyer when he had criminal problems. See also Jacob, *supra* note 8, at 296 n.477 (2003) (citing Robert J. Aalbert, *From the Classroom: Gideon’s Trumpet*, 12 J. LEG. STUD. EDUC. 321, 326 n.395 (1994)).

19 Justice Douglas called Fortas’ argument the best he had heard. See generally SCHWARTZ, *supra* note 15.

20 *Gideon*, 372 U.S. at 344; see also SCHWARTZ, *supra* note 15 at 458, 460 (describing how *Betts v. Brady*, which held that an indigent defendant did not have a due process right to counsel in noncapital cases unless he could not obtain a “fair trial,” was overruled by the Supreme Court).

21 *Gideon*, 372 U.S. at 344.
precedent. Twenty-one years earlier in *Betts v. Brady*, the Supreme Court held that the Sixth Amendment did not mandate the provision of free counsel to indigent defendants accused of serious crimes in state proceedings. Advocates and academics promptly questioned the wisdom of the *Betts* rule. In the years that followed *Betts*, many states developed their own paths and provided counsel through legislative or judicial means. By the time the Court was called upon in *Gideon* to reconsider its holding in *Betts*, twenty-two states sided with Clarence Gideon in amicus filings supporting his claim to appointed counsel. Over time, the *Betts* rule was largely swallowed by exceptions crafted by the states about evolving standards as to what fundamental fairness principles should require in court proceedings for those too poor to afford a lawyer. Without these developments, it is uncertain when, or even if, the Supreme Court would have reconsidered its prior holding in *Betts v. Brady*.

**B. The New Legal Services Program**

Although *Gideon* gave the poor a right to counsel in serious criminal proceedings, it did not address their need for counsel in civil proceedings. Soon thereafter, advocates for the poor secured the birth of a new federal legal services program that opened the doors to the nation’s courthouses for many of the poor. Congress passed the Economic Opportunity Act of 1964, establishing anti-poverty programs that made federal funds available for legal services to the poor. The Office of Economic Opportunity (“OEO”), led by Sargent Shriver, worked with local communities to solicit initial proposals for legal services funding, and by 1968,

---

22 See *supra* note 13 (describing *Betts v. Brady*).

23 See e.g., Cohen & Griswold, *supra* note 17, at E6.

24 See Laura K. Abel, *A Right to Counsel in Civil Cases: Lessons from Gideon v. Wainwright*, 15 Temp. Pol. & Civ. Rts. L. Rev. 527, 533 (2006) (discussing the development of a right to counsel in the states through an expanding variety of special cases that over time eroded the *Betts* rule such that any serious criminal charge became viewed as a special circumstance warranting the appointment of counsel). See also *Gideon*, 372 U.S. at 350 (Harlan, J., concurring).
On the Road to Civil Gideon

260 OEO programs were operating throughout the United States. After political opposition to the young program took root in California under Governor Ronald Reagan, and later in the Nixon White House when President Nixon appointed outspoken legal services foe Howard Phillips to head OEO for the purpose of dismantling the program, support for an independent legal services program gained support. After a protracted legislative fight about the structure and scope of this new, independent entity, Congress passed the Legal Services Corporation Act of 1974. President Nixon’s signature on the bill would be one of his last official acts before resigning from office that year. Under the new Act, the legal services program would now be administered by a nonprofit corporation governed by an independent board of directors appointed by the President and confirmed by the Senate.

Much of the political opposition to the legal services program was generated by the enormous success that it quickly achieved once the nation’s court houses were finally open to the poor. Attacking long-standing, systemic injustices that preyed upon the poor, federally-funded lawyers scored impressive victories in the Supreme Court and in many of the nation’s circuit and district courts that began to balance the scales of justice that previously tilted strongly in the direction of government and large corporations. In a short period of time, legal services programs won landmark cases that had a profound effect upon the poor. In Shapiro v. Thompson, Goldberg v. Kelly, Fuentes v. Shevin, King v. Smith, and Boddie v. Connecticut, among others, the

---


27 Legal Services Corporation Act, 42 U.S.C.A. § 2996c (West 2010).


Supreme Court delivered vital pronouncements that would not have been possible without the availability of legal services to the poor in the new federal program. Indigent Americans were beginning to enjoy the fruits of counsel for the first time in civil matters; however, they still did not have a right to that counsel.

C. Court Action to Secure a Civil Gideon Falls Short in Lassiter

The principal effort to establish a right to counsel in civil matters came to the Supreme Court eighteen years later in a legal challenge involving the termination of a mother’s parental rights to her child in *Lassiter v. Dep’t of Soc. Serv. of Durham County*.

Abby Gail Lassiter had been accused of not providing proper medical care to William, her infant son. A North Carolina family court adjudicated William a neglected child and transferred his custody from Lassiter to the county department of social services. One year later, Lassiter was convicted of murder charges in an unrelated matter and began serving a lengthy sentence of imprisonment. Three years after removing William from Lassiter’s care, the county department petitioned the court to terminate Lassiter’s parental rights alleging that she had not contacted William for an extended period of time and had left him in foster care for two consecutive years without showing adequate progress at remedying the problems that led to his removal from her custody.

Lassiter was brought from prison to a family court hearing to answer charges that her parental rights to William should be terminated. When Lassiter asked for a postponement of the hearing

---

33 The extraordinary record of success before the Supreme Court only tells a small part of the story. Legal services programs won cases in federal appellate and trial courts that established far reaching legal principles affecting the most essential needs of the poor. See, e.g., Escalera v. N.Y. City Hous. Auth., 425 F.2d 853 (2d Cir. 1970); Javins v. First Nat’l Realty Corp., 428 F.2d 1071 (D.C. Cir. 1970); Edwards v. Habib, 397 F.2d 687 (D.C. Cir. 1968).


35 *Id.* at 20–22.
On the Road to Civil Gideon

in order to obtain counsel, the trial court refused. The judge concluded that, despite Lassiter’s poverty and incarceration, she had been given ample opportunity to obtain counsel for the hearing and that “her failure to do so [was] without just cause.”

Without counsel, Lassiter tried unsuccessfully to represent herself and her parental rights were terminated. On appeal, Lassiter argued that she was entitled to the assistance of counsel under the Fourteenth Amendment’s due process clause since she was indigent and could not afford to hire counsel. A state appeals court found that the assistance of counsel was not constitutionally mandated, and the North Carolina Supreme Court denied review. The United States Supreme Court decided to hear the case. After reviewing prior precedent in Gideon, Argersinger, and In re Gault that required the appointment of counsel where a loss of liberty was at stake, the high Court held that as liberty interests diminish, so does an individual’s right to appointed counsel.

---

36 Id. at 22.

37 The Lassiter hearing was held on August 31, 1978. Lassiter, 452 U.S. at 21. In her own defense, Lassiter tried to cross-examine witnesses against her, but without much success. Id. at 53–56. The judge reminded her several times that she could only ask questions and that her questions were disallowed because they were really arguments, and not questions. Id. at 23. At the Supreme Court, the American Bar Association filed an amicus brief on the side of Lassiter in which it argued that involuntary termination of parental rights cases are prone to error and to ensure a fair hearing, an attorney must be made available at public expense. ABA Brief for Lassiter as Amicus Curiae Supporting Petitioner at 3–4, Lassiter v. Dep’t of Soc. Serv. of Durham Cnty., available at 1980 WL 340036 (No. 79-6432). The ABA’s brief referenced heavily Lassiter’s inability to conduct an effective cross-examination of agency witnesses who testified against her. Id. at 14. The brief argued that “without meaningful cross-examination, the risk of error in these cases substantially increases.” Id. “Judges will rely more heavily upon the state’s unchallenged presentation.” Id.

38 Lassiter, 452 U.S. at 21–24.

39 Id. at 24.

40 See Argersinger v. Hamlin, 407 U.S. 25, 37 (1972) (holding that counsel must be provided even where the crime is petty and the prison term brief).

41 See In re Gault, 387 U.S. 1, 4 (1967) (holding that a juvenile has a right to counsel where his freedom is curtailed by an institutional commitment, even though the proceeding is viewed as civil and not criminal).

42 Lassiter, 452 U.S. at 26.
Court denied Lassiter’s right to counsel after employing a balancing test weighing the due process factors identified in *Mathews v. Eldridge* with the presumption against a right to counsel without a potential deprivation of physical liberty. The Court stated, that “a wise public policy . . . may require . . . higher standards . . .” but the Fourteenth Amendment simply “imposes on the States the standards necessary to ensure that judicial proceedings are fundamentally fair.”

In the Court’s view, Lassiter’s case did not involve expert witnesses, present any specially troublesome points of law, or provide a situation where the presence of counsel could have made a determinative difference in the outcome of the case. While the Court acknowledged that a parent has an important interest in the companionship, care, custody, and management of her child that warrants special deference, a majority of the justices did not believe that the case presented a situation that warranted overcoming the presumption against a right to counsel.

*Lassiter* is often erroneously regarded as standing for the inflexible principle that civil cases not involving a loss of liberty do not require the appointment of counsel. In fact, the Supreme Court held that federal courts should evaluate the need for a court-appointed counsel on a case by case basis, utilizing the *Mathews v. Eldridge* due process factors as a guide. But the reality in busy trial courts is that a case by case approach often prompts judges to adopt an across the board rule from which they rarely deviate. Today, trial courts tend to ignore *Lassiter*’s instruction to inquire whether the *Mathews* factors might require the appointment of counsel in the cases before them and instead treat the *Lassiter*

---

43 *Mathews v. Eldridge*, 424 U.S. 319, 321 (1976) (articulating three elements that must be considered in a due process challenge: the private interests at stake, the government’s interest, and the risk of an erroneous decision).

44 *Lassiter*, 452 U.S. at 31.

45 *Id.* at 33.

46 See Michael Millemann, *The State Due Process Justification for a Right to Counsel in Some Civil Cases*, 15 TEMP. POL. & CV. RTS. L. REV. 733, 734 (2006) (noting that while *Lassiter* created a presumption against the right to counsel in civil cases that do not involve a loss of liberty, it did establish a case-by-case approach to determine whether the presumption had been overcome).
holding as an absolute rule, except in the most extreme of circumstances.47

In dissent, Justice Blackmun sharply criticized the Lassiter majority for adding the value of physical liberty, a “burdensome new layer” in his words, to the standard three-factor due process framework of Mathews v. Eldridge.48 Justice Blackmun disagreed strongly with what he perceived as the Court’s retreat from the proposition that it look at a whole area (“decision making contexts”), and not at individual litigants, when determining if due process requires the appointment of counsel.49 Justice Blackmun noted the Court’s reasoning in Goldberg v. Kelly,50 in which the Court did not look merely at the circumstances of a particular litigant, but rather at welfare recipients as an entire class.51 In Justice Blackmun’s view, procedural norms must be based on the whole context, and not on the specifics of an individual litigant.52 Despite this strong dissent, federal courts have not reconsidered Lassiter’s holding in the decades that have followed and many believe that the Supreme Court is no more likely today to disturb Lassiter’s holding.53

47 See e.g., Jacob, supra note 8, at 202.
48 Lassiter, 452 U.S. at 42 (Blackmun, J., dissenting).
49 Id. at 49–50.
51 See generally id.
52 The ABA resolution recognizes this basic principle and recommends that counsel be afforded in the most important subject matter areas involving basic human needs, rather than attempting to determine on a case-by-case basis the need for legal representation.
In the absence of a right to counsel, the poor have had very limited access to lawyers in civil matters. The litigation successes of the federal legal services program generated substantial political opposition and led to drastic funding cuts in the 1980s under the Reagan administration and again in the 1990s, following the 1994 mid-term elections. In addition, Congress imposed substantial restrictions on the activities of legal services lawyers, cutting off the poor’s access to lawyers in fundamental ways. As inadequate funding levels were reduced even further and the legal services program again became political fodder, millions of Americans were left to handle their legal problems entirely on their own. Although legal aid programs enjoy renewed support today from public and private funding sources, the harsh reality persists that without a right to counsel in civil matters, access to legal help for millions of poor Americans remains beyond reach. Especially in a weak economy, indigent litigants must fend for themselves.


55 For example, Congress prohibited legal services programs receiving federal funds from engaging in class actions, seeking attorneys’ fees, prisoner representation, representing individuals who were being evicted from public housing because they face criminal charges of selling or distributing illegal drugs, most activities involving welfare reform, lobbying, and representing people who are not U.S. citizens with certain limited exceptions. See Omnibus Consolidated Rescissions and Appropriations Act of 1996, Pub. L. No. 104-134, 110 Stat. 1321. In fiscal year 2010, Congress removed the restriction on collecting or retaining attorneys’ fees. See Consolidated Appropriations Act, Pub. L. No. 111-117 (2010).

56 See legal needs studies discussed infra note 54.

57 See, e.g., Terry Carter, Judges Say Litigants are Increasingly Going Pro Se at Their Own Peril, A.B.A.J. (July 12, 2010), http://www.abajournal.com/news/article/judges_say_litigants_increasingly_going_pro_se--at_their_own_/ (reporting on a survey conducted of ABA Judicial Division’s National
On the Road to Civil Gideon

II. The Organized Bar Takes a Bold Stand

With a decisive loss in Lassiter, the movement to achieve a civil Gideon languished for many years. As federal courts became increasingly unreceptive to expansive readings of constitutional protections, advocates generally agreed that a landmark Supreme Court decision mandating a right to counsel in civil matters, similar to what the Supreme Court had done in Gideon, was not likely to be achieved in the near future. Instead, proponents devoted their efforts to expanding access to counsel by creating justice commissions under the auspices of state supreme courts and uniting behind national efforts to increase federal funding for legal services to the poor. These efforts opened up access to counsel for perhaps millions of Americans in a broad range of civil matters, funded by state IOLTA programs, civil filing fees, bar registration fees, and a range of other mechanisms intended to boost legal services to the poor. Despite these important developments, advocates were unable to make substantial process at chipping away at the fundamental justice gap in America which, according to most studies, reveals that only 20 percent of poor Americans in need of legal help are able to be helped with current resources. A study commissioned by the Legal Services Corporation found that for every indigent client who was able to get free legal help from a legal aid office another client entering with an equal need of help was turned away. Without a right to counsel, the justice gap

---

58 Civil Gideon advocates generally agreed that they should avoid federal courts when trying to advance their efforts. See, e.g., Kaufman, supra note 53, at 351 (noting that the conventional wisdom among civil Gideon advocates was to avoid federal courts based upon the assumption that the Supreme Court was too conservative to find a right to counsel in the constitution).

59 See infra notes 81–842 for a study conducted by the American Bar Association and other studies of states citing the statistic; Legal Services Corporation, Documenting the Justice Gap in America: The Current Unmet Civil Legal Needs of Low Income Americans, Legal Services Corp. (2d ed. 2007), available at http://www.lsc.gov/justicegap.pdf (stating that less than one in five of the legal problems experienced by low-income people are addressed with the assistance of a private or legal aid lawyer).

60 See Legal Services Corporation, Documenting the Justice Gap in
remains as large as ever.

A. The ABA’s Right to Counsel Resolution

In 2005, then ABA president Michael Greco commissioned a presidential task force on access to civil justice to study whether counsel should be provided as a matter of right in civil matters to those unable to afford counsel. The task force, chaired by Howard Dana, Jr., concluded that equality before the law has remained a woefully inadequate charity over the past 130 years which has not delivered justice for all. The task force recommended that the ABA support a right to counsel at the public expense in civil matters involving basic human needs. On August 7, 2006, the ABA House of Delegates adopted Resolution 112A calling upon federal and state jurisdictions to provide counsel as a matter of right at the public expense to low-income persons in adversarial proceedings involving “basic human needs,” such as those involving “shelter, sustenance, safety, health or child custody.” The ABA resolution encouraged each jurisdiction to determine appropriate categories where the provision of counsel was most important and to develop local strategies for achieving this goal.

The ABA’s action reignited a national movement in support of...
On the Road to Civil Gideon

a right to counsel that had begun to show promise in some states.64 State and local bar associations agreed to co-sponsor the ABA resolution,65 and academic institutions held conferences to reexamine the legal underpinnings of a civil Gideon.66 The subject again became prominent in academic journals and practitioner publications.67 Many states and localities soon followed this lead,68 amidst new-found optimism that a right to counsel in essential civil legal matters might indeed be obtainable.


65 Among the co-sponsors were the Association of the Bar of the City of New York, King County Bar Association (WA), the Philadelphia Bar Association, the Maine State Bar Association, the Connecticut Bar Association, the Boston Bar Association, the Los Angeles County Bar Association, and a variety of others. For a complete list, see the ABA website, AM. BAR ASS’N, REPORT TO THE HOUSE OF DELEGATES (Aug. 2006), available at http://www.abanet.org/legalservices/sclaid/downloads/06A112A.pdf (last visited Aug. 24, 2010).


68 For example, the Pa. Bar Ass’n passed a right to counsel resolution in Sept. 2007 and commissioned a task force to implement this policy.
B. Experimentation in the States

Many believe that the key to success in the civil Gideon movement lies with growing experimentation in the states. In the twenty-one intervening years between the Supreme Court’s ruling in Betts and Gideon, most states adopted right to counsel statutes for serious criminal offenses, and by the time Gideon was argued twenty-two states agreed to sign on as amici curiae parties supporting Gideon’s position. Only Florida and two other states, Alabama and North Carolina, advocated for the retention of the Betts rule. This clear change among the states signaled an important message that had a significant impact upon the Court. Similar developments over time on the civil side are also likely to have a persuasive impact, especially upon judges who believe the law should evolve slowly as the nation’s views develop.69

One of the more promising state court initiatives in the civil Gideon movement preceded the ABA’s right to counsel resolution. In Frase v. Barnhart,70 a Maryland child custody case, advocates asked their highest state court to decide whether the poor can hope to receive equal treatment as a matter of fundamental constitutional rights if they have no access to legal help. While the Maryland Court of Appeals decided the case on other grounds and did not reach the right of counsel issue, three of the seven justices would have found such a right under the Declaration of Rights of Maryland’s state constitution.71 Not constrained by the Supreme Court’s interpretation of the federal constitution in Lassiter, the

---

69 For a discussion of state statutes guaranteeing a right to counsel in parental termination cases and other civil matters, see Laura K. Abel & Max Rettig, State Statutes Providing for a Right to Counsel in Civil Cases, 40 CLEARINGHOUSE REV. 245 (2006); see generally Rosalie R. Young, The Right to Appointed Counsel in Termination of Parental Rights Proceedings: The States’ Response to Lassiter, 14 TOURO L. REV. 247 (1997).

70 See generally Frase v. Barnhart, 840 A.2d 114 (Md. 2003). See also John Nethercut, Maryland’s Strategy for Securing a Right to Counsel in Civil Cases: Frase v. Barnhart and Beyond, 40 CLEARINGHOUSE REV. 238, 239 (2006) (concluding that challenges of Lassiter based on federal constitutional grounds were unlikely to succeed, but that a greater chance of success existed based upon state constitutional guarantees).

71 Barnhart, 840 A.2d at 126.
three Maryland justices acknowledged the historic path they were proposing by mandating a civil Gideon and the concerns it raised, but they responded with their own question: “What could be more important?”

While state litigation seeking to establish a right to counsel in selected areas of civil needs remains an important thrust of the right to counsel movement, advocates for a civil right to counsel have also turned their attention to a broad range of other advocacy measures. State and local bar associations have adopted similar resolutions to that of the ABA urging increased access to justice through the establishment of a right to counsel, and they have recommended legislative change to amend state constitutions. They have sponsored state and local legislation and have

---

72 Id. at 103, 141.
74 See generally Paul Marvy & Laura Klein Abel, Current Developments in Advocacy to Expand the Civil Right to Counsel, 25 Touro L. Rev. 131 (2009).
75 See, e.g., Alaska Bar Association (Sept. 11, 2008); Hawaii State Bar Association (Dec. 2007); Massachusetts Bar Association (May 23, 2007); Pennsylvania Bar Association (Nov. 2007); Philadelphia Bar Association (original co-sponsor of ABA resolution and additional resolutions, April 30, 2009).
76 See California, Conference of Delegates of California Bar Association (Oct. 2006) (recommending legislation to amend state constitution in order to create a right to counsel where basic human needs are at stake).
77 In October 2009, Governor Schwarzenegger of California signed the Sargent Shriver Civil Counsel Act providing funding for a two year pilot project to provide poor individuals a lawyer in certain high stakes cases (anticipated to include domestic violence claims, child custody cases, and housing matters). See Gary Toohey, A Civil Right to Counsel: Inevitable or Unrealistic, PRECEDENT, Winter 2010, at 23, available at http://members.mobar.org/pdfs/precedent/feb10/civil.pdf (citing California Recognizes Civil Right to Counsel, Creates Pilot Program, BRENNAN CENTER FOR JUSTICE, http://www.brennancenter.org/content/elert/lselert1016091 (last visited Aug. 11, 2010)). See 2009 Cal. Legis. Serv. 457 (West 2009).
established state model statutes. For example, Florida adopted a right to counsel statute in 2005 requiring legal representation for children determined to be eligible for special immigrant juvenile status so that they could apply for that status. In New York, the State Judiciary recently established a program aimed at providing for a right to counsel for homeowners facing foreclosure. The program will be initially implemented in two New York counties and may eventually be put in place throughout the state.

Advocates have also fostered state justice commissions and bar association task forces. These have spurred pilot projects,

---


81 See, e.g., MD. ACCESS TO JUSTICE COMM’N, INTERIM REPORT & RECOMMENDATIONS 1 (2009), available at http://www.courts.state.md.us/ mdjc/pdfs/interimreport111009.pdf (Maryland Access to Justice Commission was created in 2008); N.H. CITIZENS COMM’N ON THE STATE COURTS, REPORTS AND RECOMMENDATIONS 10 (2006) (studying the implementation of civil Gideon); see also Abel, supra note 24, at 534–35 (noting that a number of states have access to justice commissions with high ranking legislators and judges participating).

On the Road to Civil Gideon

academic conferences, and a significant body of published writings. As individual efforts go forward, information and strategies are exchanged as part of a National Coalition for a Civil Right to Counsel which maintains a website and provides assistance to local efforts. Finally, at its 2010 annual meeting, the ABA adopted a Model Access Act to provide a model statute for


Massachusetts launched two pilot projects with the Boston Bar Foundation to explore the impact of full representation in eviction cases. See Pilots, CIVIL RIGHT TO COUNSEL, available at http://www.civilrighttocounsel.org/advances/pilots/ (last visited Jan. 27, 2011).


states and localities to use in their varied efforts to establish and administer a civil right to counsel, accompanied by a report presenting basic principles of a right to counsel in civil legal proceedings.\textsuperscript{87}

Despite a growing number of court and legislative initiatives in states and localities, advocates have experienced difficulty in getting their hands around the most effective next steps that should be taken in pursuit of a civil \textit{Gideon}. Should advocates focus their efforts on state courts in an attempt to get favorable rulings under state constitutional guarantees? Should they focus on lobbying state legislatures for legislation that mandates counsel at the public expense in compelling subject matters that most directly address essential needs? Should they turn to local legislatures where elected officials are most closely tied to basic human needs and may be particularly sensitive to local concerns?

Or, instead, should advocates engage in an aggressive public education campaign that informs citizens that a reading of their Miranda rights which they often hear on television crime shows, about having a lawyer appointed if one cannot be afforded, does not apply to even the most important of civil cases?\textsuperscript{88} Alternatively, should advocates establish bar- or foundation-sponsored pilot projects that will integrate empirical studies to measure the success of their efforts and the true societal costs of not providing counsel when basic human needs are at stake? While there are no certain or easy answers to these questions and each locality must decide for itself which path best matches its own unique needs and concerns, there are lessons to be learned from the efforts of others to expand access to counsel in discrete civil matters.

\textsuperscript{87} The ABA’s Model Access Act for implementation of a civil right to counsel, \textit{MODEL ACCESS ACT} § 104 (2010), and accompanying Basic Principles of a Right to Counsel in Civil Proceedings, \textit{MODEL ACCESS ACT} § 105 (2010), were adopted at the ABA’s annual meeting in August 5–10, 2010 in San Francisco. See \textit{ABA Announcements}, ABA JOURNAL (Mar. 1, 2010), available at http://www.abajournal.com/magazine/article/aba_announcements2/.

\textsuperscript{88} Interestingly, almost four-fifths incorrectly believe that the poor already have a right to legal aid in civil cases. Deborah L. Rhode, \textit{Access to Justice}, 69 \textit{FORDHAM L. REV.} 1785, 1792 (2001).
III. CIVIL FORFEITURE REFORM AND A RIGHT TO COUNSEL FOR INDIGENT HOMEOWNERS

A. Introduction to Civil Forfeiture

The power of the government to seize contraband and obtain its forfeiture is a law enforcement tool as old as the nation.\(^8^9\) In 1983, President Ronald Reagan declared war on drug racketeers in a State of the Union message and promoted strong federal forfeiture measures when he signed the Comprehensive Crime Control Act of 1984.\(^9^0\) While civil and criminal forfeiture had been used to fight drug trafficking since the early 1970s, the 1984 legislation extended the reach of civil forfeiture to real estate that is used to facilitate drug transactions.

Civil forfeiture actions are *in rem* proceedings brought against the offending property on the theory that it is the property that makes possible the illegal conduct. As a result, the action is not brought against the property owner and the guilt or innocence of the property owner is not determinative of the outcome in a civil forfeiture action. The theory behind civil forfeiture is that the property has done wrong by facilitating illegal drug activity and that the government can eliminate illegal drug activity by taking the profit out of drug offenses. At the same time, the government’s use of civil forfeiture as a drug-fighting tool can recoup drug enforcement costs and amass significant sums that are available for future law enforcement activities.

As law enforcement authorities gained financial success through civil forfeiture, they became more aggressive in their use of forfeiture proceedings. Many critics charged that harsh penalties, minimal safeguards, and relaxed procedures provided in civil forfeiture statutes trampled on the rights of ordinary citizens.

---

\(^8^9\) From the earliest days of the Republic, cargo ships were subject to federal forfeiture laws. See historical discussion of forfeiture in Caldero-Toledo v. Pearson Yacht Leasing Co., 416 U.S. 663, 683 (1974).

who were often innocent themselves of any wrongdoing. Federal courts expressed growing concern over government forfeiture practices in the cases before them. The cry for forfeiture reform among interest groups and in the media mounted, leading to the introduction of reform legislation in Congress. Representative Henry Hyde wrote a popular book, entitled *Forfeiting Our Property Rights, Is Your Property Safe from Seizure?*, in which he expressed concern that the war on drugs had become a “series of frontal attacks on basic American constitutional guarantees . . .”

---


93 On June 15, 1993, Representative Henry Hyde introduced H.R. 2417, known as the Civil Forfeiture Act of 1993, the first initiative in a seven year battle to obtain reform that would ultimately culminate in CAFRA. *See Civil Asset Forfeiture Reform Act, H.R. 2417, 103rd Cong. (1993).* The Hyde Bill attempted to heighten the government’s burden of proof, provide for the appointment of counsel and eliminate cost bonds, among other things. *See id.* Congressman John Conyers Jr., a Democrat from Michigan, introduced H.R. 3347, known as the Asset Forfeiture Justice Act, Asset Forfeiture Justice Act, H.R. 3347, 103rd Cong. (1993), after holding hearings on forfeiture reform. *See Department of Justice Asset Forfeiture Program: Hearing Before the Subcomm. on Legislation and National Security of the House Comm. on Government Operations, 102d Cong., 2d Sess. 12–13 (1992).* Conyers’ bill would have eliminated in rem forfeiture proceedings, requiring instead that the owner of property be convicted of a crime upon which the forfeiture is based before allowing the government to forfeit property. *Civil Asset Forfeiture Reform Act, H.R. 2417, 103rd Cong. (1993).*

94 *Henry Hyde, Forfeiting Our Property Rights, Is Your Property Safe from Seizure?* 1 (Cato Institute 1995). Representative Hyde plainly voiced his opinion in his book: “My personal belief, which prompted my writing this book, is that there is an immediate need for restoration of the constitutional
In his book, he called for many reforms, including the appointment of counsel for indigents.95

The demand for forfeiture reform recognized from the start that property owners needed greater access to counsel if property rights were to be adequately protected. Some began to argue that property owners were constitutionally entitled to counsel in civil forfeiture proceedings. Since many aspects of civil forfeiture closely resemble criminal proceedings, and civil forfeiture is acknowledged to be “quasi-criminal” and punitive in nature, advocates argued that property owners enjoyed a Sixth Amendment right to counsel in civil forfeiture cases and were entitled to have counsel appointed for them under the Criminal Justice Act.96 However, the Supreme Court has not deemed forfeiture actions to be criminal proceedings for the purposes of entitlement to counsel,97 leaving to lower federal courts the responsibility of deciding the constitutional dimensions of this issue in individual cases.

B. Court Efforts to Obtain a Right to Counsel

Thus far, efforts through the courts to establish a right to counsel in civil forfeiture proceedings have achieved momentary victories, but not lasting success. In United States v. Bowman,98 a

principles that are debased by the current application of asset forfeiture laws.”

Id. at 4.
95 Id. at 81 (proposing that anyone financially unable to obtain representation in a federal civil forfeiture matter be appointed counsel, paid with funds from the Department of Justice Assets Forfeiture Fund).
97 See Austin v. United States, 509 U.S. 602, 608 n.4 (1993) (holding that some constitutional protections apply to civil forfeiture, but unless a civil forfeiture proceeding is so punitive that it must reasonably be considered criminal, counsel is not required to be appointed). United States v. 7108 West Grand Ave., Chicago, Ill., 15 F.3d 632, 635 (7th Cir. 1994).
district court in Alaska held that the Sixth Amendment guaranteed a lawyer to a claimant in a civil forfeiture proceeding. Unwilling to “exalt form over substance,” the Alaska court rejected the government’s argument that the Sixth Amendment’s right to counsel does not attach because forfeiture actions, being in rem rather than in personam, are not adversarial proceedings. Ultimately, however, the order and opinion in the case were vacated based upon plea arguments entered in the case.99

The legal claim that the Sixth Amendment provides a right to counsel in civil forfeiture proceedings has been rejected by federal appellate courts in the Second,100 Sixth,101 Seventh,102 Ninth,103 Tenth,104 and Eleventh105 circuits on the basis that civil forfeiture statutes authorize the forfeiture of property and not the imprisonment of the property owner.106

Advocates seeking to establish a right to counsel in civil forfeiture proceedings in state courts have not fared any better. In a case of first impression, a Pennsylvania intermediate appellate court held that an indigent property owner was entitled to counsel in civil forfeiture actions under the due process clause of the Fourteenth Amendment.107 The court acknowledged that forfeiture

100 See United States v. 87 Blackheath Road, 201 F.3d 98 (2d Cir. 2000).
103 Acosta v. United States, 130 Fed. Appx. 881 (9th Cir. 2005); United States v. $292,288.04 in U.S. Currency, 54 F.3d 564, 569 (9th Cir. 1995) (finding there is no Sixth Amendment right to counsel because civil forfeiture statutes authorize the forfeiture of property and not the imprisonment of the property owner).
104 United States v. Deninno, 103 F.3d 82, 86 (10th Cir. 1996).
105 United States v. 817 N.E. 29th Drive, 175 F.3d 1304, 1311–12 (11th Cir. 1999), cert. denied, 528 U.S. 1083 (2000).
106 See also United States v. All Funds on Deposit in any and all Accounts, 2009 WL 2424337 (E.D.N.Y. 2009) (holding that a property owner is not entitled to an appointment of counsel because the Sixth Amendment’s protections are confined to criminal prosecutions).
107 Commonwealth of Pennsylvania v. $9,847.00 U.S. Currency, 637 A.2d
law is a “complicated and abstruse” field that poses substantial burdens for a pro se litigant. Influenced by the U.S. Supreme Court’s holding that the Eighth Amendment’s excessive fines clause applied to civil forfeiture proceedings, Pennsylvania’s Commonwealth Court concluded that appointment of counsel was constitutionally required under the Fourteenth Amendment’s due process clause in order to protect an indigent property owner from the likelihood of an erroneous deprivation of his property.

This decision was short-lived. Three years later, the Pennsylvania Supreme Court reversed the Commonwealth Court and held that the due process clause of the Fourteenth Amendment did not require the appointment of counsel for an indigent property owner. Applying the Supreme Court’s balancing test announced in *Mathews v. Eldridge*, the Pennsylvania Supreme Court found the burden to government of providing counsel to a class of property claimants to be substantial and the risk of erroneous deprivation to property owners minimal. Without a liberty interest at risk, the Court held that a proper application of *Mathews* weighed against a property claimant’s entitlement to counsel in civil forfeiture actions. Other state courts have similarly declined to find a right to counsel in civil forfeiture proceedings.

C. Legislative Reform

As judicial efforts to obtain a constitutional right to counsel in

---


108 *$9,847.00 U.S. Currency*, 637 A.2d at 743.


110 *Commonwealth v. $9,847.00 U.S. Currency*, 704 A.2d 612 (Pa. 1997).

111 The Pennsylvania Supreme Court did not consider whether the Pennsylvania Constitution might require counsel under circumstances not required by the U.S. Constitution. *Id.* at 617.

112 See, e.g., *State v. $1,010.00 in American Currency*, 722 N.W. 2d 92 (SD 2009) (holding no Sixth Amendment right attached because the action was not a criminal proceeding, and rejecting the argument that state law required the appointment of counsel on the basis of equity considerations). See also *State v. Halvorson*, 724 N.W. 2d 703 (Table) (Wis. App. 2006); *State v. Hermann*, 719 N.W. 2d 800 (Table) (Wis. App. 2006); *In re Property Seized from Behmer*, 720 N.W. 2d 191 (Table), 2006 WL 1279318 (Iowa Ct. App. 2006).
civil forfeiture actions proved largely unsuccessful, advocacy shifted to the legislature. With a growing number of publicized incidents of forfeiture abuse and increasing concern that innocent owners forfeited property without sufficient legal protection, Congress embarked on a seven year effort to reform the nation’s civil forfeiture laws. These legislative efforts began in the House of Representatives and ultimately obtained consensus in the Senate, culminating in the passage of the Civil Asset Forfeiture Reform Act of 2000 (“CAFRA”).

The new legislation enacted important safeguards to achieve procedural fairness and enable average citizens to contest forfeiture claims brought against their property. The reform statute created a uniform innocent owner defense for all federal civil forfeiture statutes and heightened the government’s burden required to prove that private property was subject to forfeiture. Representative Henry Hyde, a primary architect of CAFRA’s reforms, spoke proudly of the Act’s accomplishments: “It returns civil asset forfeiture to the ranks of respected law-enforcement tools that can be used without risk to the civil liberties and property rights of American citizens. We are all better off that this is so.”

While CAFRA included many important civil legal protections for ordinary citizens, the Act’s expanded right to counsel is considered by some to be among its most important reforms. In proposing legislative reform, the House Judiciary Committee’s report recommended “eight core reforms” to civil forfeiture law, of which the appointment of counsel was listed second only to reform of the burden of proof. Acknowledging that there is no

---

113 See Hyde, supra note 94, at 5–6.
114 See Civil Asset Forfeiture Reform Act, 18 U.S.C.A. § 983(c) (West 2010). Under the new Act, the government is now required to meet a preponderance of the evidence standard in order to demonstrate the forfeitability of property. Id.
116 See 1 David B. Smith, Prosecution and Defense of Forfeiture Cases P 1.02[2]-[3] (Matthew Binder 2010).
118 The eight core reforms proposed in H.R. 1658 are outlined in Report
On the Road to Civil Gideon

Sixth Amendment right to appointed counsel for indigents in civil forfeiture cases, the report nonetheless expresses the Committee’s conclusion that civil forfeiture proceedings are “so punitive in nature that appointed counsel should be made available for those who are indigent, or made indigent by seizure in appropriate circumstances.”

The Hyde-Conyers Civil Asset Forfeiture Reform Act (House Bill 1658) passed the House of Representatives in 1999 by an overwhelming vote of 375 to 48. As negotiations continued with the Department of Justice, Senators Hatch and Leahy introduced Senate Bill 1931. The Senate bill rejected the broader grant of authority for appointed counsel found in the House bill, and instead provided for the appointment of counsel only where an indigent’s primary residence was the subject of the proceeding. Negotiations with Senators Sessions and Schumer led to agreement on thirty substantive changes and the addition of new sections giving additional authority to law enforcement to utilize criminal forfeiture powers. These new additions included the provision of counsel for indigent homeowners seeking to defend against civil forfeiture of their primary residences, utilizing the Legal Services Corporation as the conduit for the implementation of a right to counsel. Promising “to do the right thing on this important issue of fairness,” the Senate bill included a “right to counsel” provision for a limited class of civil asset forfeiture proceedings. This provision was adopted in the final text of the bill that was ultimately signed into law by President Clinton in April of 2000.

106-192 as the following: Burden of proof, appointment of counsel, innocent owner defense, return of property upon showing of hardship, compensation for damage to property while in the government’s possession, elimination of cost bond, adequate time to contest forfeiture, and interest. Id. at 11-19.

119 Id. at 14. While the Judiciary Committee reported H.R. 1658 favorably to the House of Representatives by a vote of 27-3, the dissenting view expressed in the Committee’s report took issue with the expanded provision for appointment of counsel, believing that the bill lacked substantial safeguards to protect against abuse of this provision. The dissenting view also noted that successful challenges to forfeiture were already eligible to recover attorneys’ fees under the Equal Access to Justice Act. Id. at 20, 34.

120 146 CONG. REC. 3656 (2000).
121 See Civil Asset Forfeiture Reform Act, 18 U.S.C.A. § 983(b)(2)(A)
The text of this provision set forth the parameters of this right:

If a person with standing to contest the forfeiture of property in a judicial civil forfeiture proceeding under a civil forfeiture statute is financially unable to obtain representation by counsel, and the property subject to forfeiture is real property that is being used by the person as a primary residence, the court, at the request of the person, shall insure that the person is represented by an attorney for the Legal Services Corporation with respect to the claim. 122

The right to counsel provided for attorney compensation at levels equivalent to that provided for other court-appointed representation, 123 and required that the Legal Services Corporation submit statements of reasonable fees and costs during the course of the representation for court review. 124 At its core, the newly created right was intended to protect the family home against erroneous deprivation under harsh civil forfeiture laws. 125 Significantly, protection of the family home received enthusiastic support from Congress, reflecting deeply-held views that private homeownership represents a cornerstone of the American dream. 126

(122 Id.
123 See id. § 983(b)(3).
124 Id. § 983(b)(2)(B)(i).
125 The adverse impact and severe consequences upon an entire family caused by the forfeiting of a primary residence was clearly on the minds of key senators as compromises were reached leading to this provision. For example, Senator Leahy noted that Vermont state law does not permit the forfeiture of real property that is used as the primary residence of a person involved in the violation and a member or members of that person’s family. See 146 Cong. Rec. 3655 (2000) (statement of Sen. Leahy) (citing 18 V.S.A. § 4241(a)(5)).
126 See Stephanie M. Stern, Residential Protectionism and the Legal Mythology of Home, 107 Mich. L. Rev. 1093, 1095 (2009) (describing theoretical notions of the home’s reign over property law and its connection to one’s personhood). See also Jim Cullen, The American Dream 136 (Oxford Univ. Press 2003) (noting that no American dream has broader appeal than that of owning a home). The premium placed on protecting the family home from a wrongful taking may also been seen in current efforts to provide counsel for
CAFRA expands access to counsel in three important ways. First, it authorizes federal courts to appoint counsel for indigent property owners who are accused of a crime related to the subject of the civil forfeiture proceeding and are represented by court-appointed counsel in the underlying criminal case. In these instances, a court is encouraged, though not required, to appoint counsel for a property owner where the individual has standing to contest the civil forfeiture and asserts a claim in good faith. Lawyers who are appointed to provide representation are compensated at levels authorized by the Criminal Justice Act for court-appointed counsel in criminal proceedings.

Second, upon request of an indigent property owner, the act requires a court to appoint counsel where a primary residence is the subject of civil forfeiture proceedings. In these instances, the court appoints the Legal Services Corporation and, again, the lawyer’s services are compensated at rates equivalent to that set for court-appointed counsel under the Criminal Justice Act. Finally, CAFRA provides a financial incentive to expand access to legal representation by authorizing an award of attorney’s fees to a property owner who substantially prevails against the homeowners in foreclosure proceedings. See, e.g., supra note 80.

In addition to expanding access to counsel, CAFRA eliminated cost bonds which, until their removal, presented onerous obstacles to challenging government seizures of private property. Civil Asset Forfeiture Reform Act, 19 U.S.C.A. § 1608 (West 2010).

128 18 U.S.C. § 983(b)(1)(A) provides that:

if a person with standing to contest the forfeiture of property in a judicial civil forfeiture proceeding under a civil forfeiture statute is financially unable to obtain representation by counsel, and the person is represented by counsel appointed under section 3006A of this title in connection with a related criminal case, the court may authorize counsel to represent that person with respect to the claim.

Civil Asset Forfeiture Reform Act, 18 U.S.C.A. § 983(b)(1)(A) (West 2010). 18 U.S.C. § 983(b)(1)(B) instructs trial courts to take into account “the person’s standing to contest the forfeiture,” and “whether the claim appears to be made in good faith” when determining whether to authorize an appointment of counsel.


129 Id. § 983(b)(2)(A).

130 Id. § 983(b)(2)(A)–(B).
United States in a civil forfeiture proceeding. While the Equal Access to Justice Act (“EAJA”) had already authorized attorney’s fees against the U.S. government under limited circumstances, CAFRA’s grant of authority to a claimant who substantially prevails in civil forfeiture litigation (and is not convicted of a crime related to the subject property) expands the potential for recovery of attorney’s fees against the federal government.131

IV. FIVE LESSONS LEARNED FROM CAFRA’S STATUTORY RIGHT TO COUNSEL

The successful legislative effort to expand access to counsel for property owners in federal civil forfeiture proceedings holds valuable lessons for the civil Gideon movement. The movement to establish a right to counsel proceeded simultaneously in both judicial and legislative forums, but advocates soon found that judicial efforts largely failed while legislative initiatives offered greater promise of success. The balance of this Article identifies and discusses five important lessons that may be drawn from this successful reform effort, with special focus on CAFRA’s achievement of a right to counsel for indigent homeowners whose primary residences are the subject of federal civil forfeiture proceedings. Hopefully, these lessons offer helpful guidance to advocates fighting to secure a right to counsel at the public expense for indigent litigants in a broad range of civil proceedings.

A. Lesson One: Narrative stories of failure which document how lives are shattered and private property seized from ordinary citizens in civil proceedings, without the safeguard of having a lawyer present, provide a powerful catalyst for legislative change.

A powerful catalyst for legislative change is often rooted in compelling narrative stories recounted by ordinary citizens describing incidents of abuse, injustice, and official overreaching.

On the Road to Civil Gideon

“Narratives make a point and persuade people because of the lifelikeness, which is in turn based on a person’s knowledge about how things really happen in the world . . . .” (internal quotations omitted). Lawyers have long relied upon storytelling as a valuable advocacy tool because they recognize that emotion persuades. For this reason, storytelling is routinely used at trial in appeals to juries and also in negotiations, and it is strategically used in legal advocacy and brief writing addressed to the highest courts.

Storytelling is especially effective when it depicts characters that draw listeners into a narrative that facilitates empathy and understanding. It is widely acknowledged that “[s]tories, parables, chronicles, and narratives are powerful means for destroying mindset—the bundle of presuppositions, received wisoms, and shared understandings against a background of which legal and political discourse takes place.” As a result, legal narrative is increasingly taught in law schools in appreciation

132 Bret Rappaport, Tapping the Human Adaptive Origins of Storytelling by Requiring Legal Writing Students to Read a Novel in Order to Appreciate How Character, Setting, Plot, Theme, and Tone (CSPTT) are as Important as IRAC, 25 T.M. COOLEY L. REV. 267, 273–74 (2008).

133 Id. at 276.

134 See Stacey Caplow, Putting the “I” in Wr*ng: Drafting an A/Effective Personal Statement to Tell a Winning Refugee Story, 14 LEGAL WRITING: J. LEGAL WRITING INST. 249, 261 n.41, 42 (2008).

135 While trial instruction materials routinely tout the power of storytelling in convincing a trial judge or jury of the wisdom of the lawyer’s legal position, narrative persuasion is actually used effectively throughout all lawyering and even in more formal and legalistic means of advocacy, such as appellate briefs and oral argument before the highest courts. See, e.g., Richard K., Sherwin, The Narrative Construction of Legal Reality, 18 VT. L. REV. 681, 709–16 (1994) (discussing storytelling that was used by advocates in their brief in Miranda v. Arizona to convince the high Court that the interrogation process was unfair). See also Philip N. Meyer, Are the Characters in a Death Penalty Brief Like the Characters in a Movie?, 32 VT. L. REV. 877 n.4 (2008) (citing Anthony G. Amsterdam’s list of limits in using narrative persuasion in cost-conviction litigation).

136 Meyer, supra note 135, at 1.

of the notion that storytelling affects what we come to regard as truth and reality, and serves as an effective means of motivating us to adopt a proposed remedy to a legal problem. At its core, narrative stories are “enabling and empowering and, indeed, fundamental to how we fashion our beliefs and how we act upon them.” In short, narratives are an important and effective component of advocacy in all of its forms.

Storytelling has long been at the heart of successful legislative initiatives. Narrative stories resonate deeply with the public (and their elected officials) when the lives of ordinary people are seriously injured by governmental action which violates widely held notions of fairness. Storytelling can establish and perpetuate a particular view of reality, and listening to narrative stories allows one to see the world through another’s eyes. Proponents of civil asset forfeiture reform understood this basic tool and used it effectively in their efforts to achieve systemic change.

The highly rated television news show 60 Minutes is a national forum for effective storytelling that has often prompted legislative action. In 1992, the show featured the story of Willie Jones, an African-American landscaper who was stopped at the Nashville airport after paying cash for his airline ticket. While detained at

---


139 Philip N. Meyer, Teaching Writing and Teaching Doctrine: A Symbiotic Relationship? Vignettes from a Narrative Primer, 12 LEGAL WRITING: J. LEGAL WRITING INST. 229, 230 (2006) (discussing why it is important to teach narrative persuasion and providing a brief summary of references to scholarship on legal storytelling and narrative jurisprudence).

140 See Delgado, supra note 137, at 2422. See also Bendavid, supra note 91, at 1.

141 Delgado, supra note 137, at 2439.

On the Road to Civil Gideon

the ticket counter, Mr. Jones was told that no one had ever paid cash for a ticket. Following detention, law enforcement authorities seized $9,000 in cash from his person because, according to police, he matched the profile of a drug courier. Although Mr. Jones explained that he was carrying this sum of cash to buy landscaping materials for his business from a nursery in Texas that required payment by cash, law enforcement authorizes were unpersuaded. In testimony before the House Judiciary Committee, Mr. Jones described what happened to him after being stopped by police at the airport:

I was questioned about had I ever been involved in any drug-related activity, and I told them, no, I had not. So they told me I might as well tell the truth because they were going to find out anyway. So they ran it through on the computer after I presented my driver’s license to them, which everything was—I had—it was all in my name. And he ran it through the computer, and one officer told the other one, saying, he is clean. But instead, they said that the dogs hit on the money. So they told me at that time they was going to confiscate the money.

The agents contended that police dogs had identified traces of drugs on the money, justifying the seizure of Jones’ cash. Mr. Jones was then released by the agents and never charged with a criminal offense. He was told that he could continue on to Texas if he wanted since his plane had not yet departed. Of course, as noted by Representative Henry Hyde in a legislative hearing, Mr. Jones had no reason to continue on to Texas since his money was gone and he could not buy the shrubs that his cash was intended to purchase.

Federal civil forfeiture law required Mr. Jones to post a 10 percent cost bond in order to legally challenge the seizure of his property. However, he was unable to afford this cost. He later

144 Id. See also 60 Minutes, supra note 142.
146 Id. at 7.
147 Id.
filed a civil action in federal court alleging that he was the victim of an unconstitutional search and seizure.\footnote{Jones v. United States Drug Enforcement Admin., 819 F. Supp. 698, 724 (M.D. Tenn. 1993).} The court agreed, and his cash was ultimately returned to him. In ruling for Mr. Jones, a federal judge sounded a strong note of concern, stating “that the statutory [forfeiture] scheme as well as its administrative implementation provide[d] substantial opportunity for abuse and potentiality for corruption.”\footnote{Id.}

The Willie Jones story highlighted the burdensome obstacles that ordinary citizens faced in challenging the government’s seizure of their property. In addition, it raised serious issues of governmental overreaching seemingly authorized by civil forfeiture law. In Mr. Jones’ case, inadequate legal protections also raised troubling questions as to whether civil forfeiture law was contributing to racial profiling by law enforcement authorities.\footnote{See H. R. Rep. No. 105-358, pt. 1, at 23–26.}

While the court did not find sufficient evidence of racial motivation in the Jones’ case, race has played a part in other drug seizure investigations.\footnote{Jones, 819 F. Supp. at 723.}

Congress heard compelling stories in legislative hearings from many victims of civil forfeiture law. In Vermont, civil forfeiture practices garnered considerable legislative and media attention when the parents of a local Vermont family were accused and convicted of federal drug violations. The Mannings and their four children lived on a farm owned by the parents. The federal government brought an action to forfeit the family home based upon the alleged illegal acts of the parents. If the action proved successful, a forfeiture would have deprived the innocent children of their family home. Senator Jeffords, among others, interceded and convinced the U.S. attorney to accept a beneficial trust for the children that would allow them to continue to live on the farm even if the property was confiscated in the forfeiture action.

The Manning children were fortunate to have influential elected officials to speak on their behalf. Many others, however, are not so fortunate, especially when they lack the financial means...
to hire an attorney to protect their interests. The experience of a constituent family compelled Vermont’s Senator James Jeffords to introduce a civil asset forfeiture reform act designed to prevent the “devastating psychological impact that the confiscation of homes can have on the innocent children who live in them.” Senator Jeffords noted that too often property is seized from individuals who are never charged with or convicted of a crime. He expressed special concern for individuals who face difficult burdens when the government’s actions take the form of a civil proceeding which lacks the protections inherent in a criminal case. Senator Jeffords recognized that property owners need the assistance of counsel under these circumstances and he therefore sponsored a provision in CAFRA which authorized the appointment of counsel for a person who is financially unable to obtain representation. Under this proposal, an appointed lawyer would be compensated in an amount equal to that provided to appointed counsel in criminal proceedings.

The adage attributed to former House Speaker Tip O’Neal that all politics is local proved true in Vermont and played an

---


153 139 CONG. REC. S1,655 (daily ed. Nov. 10, 1993) (statement of Sen. James Jeffords), 139 CONG. REC. S 15601, at *S1,655 (LEXIS).

154 Senator Jeffords noted that “as much as 80 percent of the people whose property is seized are never charged with, let alone convicted of, a crime.” Id.

155 Sec. 608(b)(1) and (2) of the Senate bill provided for compensation to be funded by the Justice Department’s Assets Forfeiture Fund established under section 524 of title 28, United States Code, and compensated at rates in accordance with section 3006A of title 18, United States Code, with maximum compensation at $3500 per attorney for representation at the district court level and $2500 per attorney for appellate court representation, similar to maximums set in federal felony cases. Id.

important role in the introduction of legislation intended to remedy abuses of governmental authority back home.\footnote{157}{See Albert Hunt, \textit{Rep. O’Neill, House Democrats' Choice, Looms as Powerful and Assertive Speaker}, WALL ST. J., Dec. 6, 1976.} Senator Patrick Leahy of Vermont put it this way:

I am well aware from incidents in Vermont about how aggressive use by Federal and State law enforcement officials of civil asset forfeiture laws can appear unfair and excessive, and thereby fuel public distrust of the government in general and law enforcement in particular.\footnote{158}{146 CONG. REC. 3,655 (2000) (statement of Sen. Patrick Leahy).}

Compelling stories that prompt legislative action often come from the experiences of ordinary families in an elected official’s home district. According to House Judiciary Committee testimony, Margaret L. Cutkomp was seventy-five years old and a hard-working, frugal woman who chose to save rather than spend her savings.\footnote{159}{Civil Asset Forfeiture Reform Act: Hearing on H.R. 1916 Before the H. Comm. on the Judiciary, 104th Cong. 21 (1996) (statement of King Cutkomp).} She never took a vacation or missed a day’s work in the business.\footnote{160}{Id.} By age seventy, Margaret Cutkomp had acquired ownership of a couple of residential rental properties and had saved a total of around $70,000 which she kept in a floor safe located in her house.\footnote{161}{Id.} She was a holdover from the Great Depression and grew up distrusting banks.\footnote{162}{Id.}

In December 1989, federal law enforcement authorities seized Margaret Cutkomp’s cash savings and three months later her home and two rental properties which she owned.\footnote{163}{Id.} The government never charged Margaret Cutkomp with a crime.\footnote{164}{Id.} Apparently, her
only wrongful act was living next door to one of her adult sons, who purportedly sold marijuana from his home. In legislative testimony, Margaret Cutkomp’s other son, King Cutkomp, described how his mother’s safe was rusted shut and had to be drilled open, a fact very much at odds with the government’s theory that the cash stored in her home was the product of the alleged illegal activity. Indeed, the bills in the safe were mostly old bills from the 1960s and 1970s that showed their age by being covered in mold and mildew.

Nonetheless, Margaret Cutkomp was told by government authorities that she had one-half hour to pack up and get out of her house. While an attorney called by King Cutkomp was able to stop his mother’s eviction from her home, King Cutkomp ultimately determined that it would cost more to go to trial and fight the government than his mother’s seized property was actually worth. A settlement with the government allowed Margaret Cutkomp to keep some of her life-long savings but, according to King Cutkomp’s testimony, the government took most of it and along with it his mother’s “dignity and love for our government.” King ultimately joined reform efforts and urged Congress to re-write civil forfeiture laws to include “proof, fairness and compassion.” He testified that civil forfeiture law was “ruining people’s lives” and he called it “a national disgrace.”

Proponents of forfeiture reform recognized that legal arguments and appeals to noble principles alone would not bring about needed legislative change. They understood that they needed to show in plain terms how civil forfeiture practices victimized ordinary citizens, offended basic principles of fairness, and seized cherished belongings often without any lawyer to help a property owner. Legislative success depended on being able to bring these tragic stories to light, especially as they affected ordinary citizens.

---

165 See id. at 22.
166 Id.
167 Id.
168 Id. at 22–23.
169 Id. at 23.
170 Id. at 24.
171 Id.
who were absolutely innocent of any wrongdoing. As the call for forfeiture reform drew increasing support, Senator Leahy proclaimed that it was “time for Congress to catch up with the American people and the courts and do the right thing on this important issue of fairness.”

To remedy forfeiture abuse, Congress understood that it needed to address the fact that so many property owners did not have counsel and often surrendered their private property rather than oppose the government. As a result, access to counsel became one of the key reforms in overhauling the nation’s civil forfeiture laws. Congress appeared particularly disturbed by the fact that individuals accused of crimes had greater access to counsel than law-abiding property owners who were not charged with any criminal wrongdoing. To restore fairness to forfeiture law, reform legislation needed to increase access to counsel.

In recent years, advocates seeking greater funding for civil legal aid and promoting higher levels of pro bono participation from the private bar have generally appealed to higher callings of the legal profession and to noble objectives of due process of law. Quite understandably, the access to justice movement has touted stories of success which lawyers for the poor have achieved when they represent indigent clients. These stories are generally uplifting: they describe a home saved, a public benefit obtained, or personal safety restored. The stories document happy endings in courts and governmental agencies, made possible because the poor had a trained advocate at their side. The stories are frequently recounted in bar association magazines or legal newspapers, with the goal of boosting lawyer volunteerism and promoting increased funding from public and private sources. In short, success stories make lawyers feel good and help to secure needed participation.

and support from the bar and judiciary.

However, feel good stories are not likely to be the stories that will ultimately mold public opinion on the need for a lawyer at the public expense in a civil case. Civil forfeiture reform teaches us that the narrative stories which are more likely to have a strong impact on public opinion and elected officials are those that describe tragic consequences that befall ordinary families when they do not have a lawyer to protect their most precious interests. Rather than success stories, these are stories of failure. These are the stories of what happens to families, low-income and middle-income, who lose their homes; elderly citizens who are denied critical medical care when they need it most; and parents who lose custody of their children not because the facts or the law are against them, but rather because they simply do not have a lawyer to give them a shot at a fair hearing in our civil justice system. These are the stories in which justice is meted out not based upon a proper application of the law, but instead based upon harsh realities of default in which status and wealth are deciding factors. These are the stories that must be told in public forums, on television, in social media, in general circulation newspapers, and ultimately in the halls of Congress.

The first lesson of civil forfeiture reform is that for legislative change to occur, the public and their elected officials must be confronted with the tragic stories of ordinary lives shattered in our civil justice system, not based upon what is right or fair, but rather because they did not have a lawyer at their side to protect their most fundamental interests.

B. Lesson Two: Strong cautionary statements about inadequate legal protections, voiced by influential courts, and bolstered by academic and popular criticism and supporting empirical data, provide a firm foundation for civil justice reforms.

The narrative stories of citizens victimized by civil forfeiture abuse painted a powerful picture of the problem, but might not have resulted in legislative change without the official imprimatur of influential courts and the detailed writings of academics
describing how civil forfeiture procedures were failing ordinary citizens. In other words, forfeiture reform hinged not only on the telling of narrative stories; these experiences needed to be validated by credible and respected institutions wielding power in our society.

In civil forfeiture reform, official validation of legal deficiencies and systemic unfairness took many forms. Strong statements appearing in the text of decisions by appellate courts in the cases before them provided a powerful and urgent message from conservative institutions of the need for substantial change. Under some circumstances, judicially-crafted language appearing in court decisions can motivate and influence the expectations of social actors. Unquestionably, court victories affect legal discourse and embolden action, while enhancing negotiation postures and mobilizing social movements. Legal action has often been the catalyst for change needed to remedy social injustice. Lawyers are often good social change agents because of their “[l]eadership qualities, forensic ability, talent for reasoning, and knowledge of the legal system,” all of which aid in generating change and nourishing social movements to succeed.

In civil forfeiture cases before them, federal judges expressed deep concern that the legal framework did not adequately protect the legitimate interests of property owners. For example, in United States v. All Assets of Statewide Auto Parts, Inc., the Second Circuit stated, “We continue to be enormously troubled by the government’s increasing and virtually unchecked use of


On the Road to Civil Gideon

civil forfeiture statutes and the disregard for due process that is buried in those statutes." Federal judges voiced their concerns, and Congress listened.

Critical discussion of the harshness of civil forfeiture law also took hold in academic journals, popular books, special interest campaigns, and investigative articles published by general newspapers. Representative Henry Hyde forcefully advanced the need for change in his book *Forfeiting Our Property Rights*, in which he directly questioned whether private property was safe from seizure. He voiced concern about the violation of constitutional protections and the extent to which cherished liberties were lost in civil forfeiture cases. Summarizing his reasons for writing a book on this topic, Representative Hyde stated, “My personal belief, which prompted my writing this book, is that there is an immediate need for restoration of the constitutional principles that are debased by the current application of asset forfeiture laws.” Representative Hyde was not alone. Leonard A. Levy, a constitutional scholar, wrote *A License to Steal* in which he documented forfeiture abuses and expressed views that were highly critical of civil forfeiture practices.

The American Civil Liberties Union ran a media campaign exposing the dangers of civil forfeiture and took out a full-page advertisement in a Sunday *New York Times Magazine* condemning civil forfeiture practices. General circulation newspapers

---

179 United States v. All Assets of Statewide Auto Parts, Inc., 971 F.2d 896, 905 (2d Cir. 1992).

180 See generally 146 CONG. REC. 3655 (2000) (statement of Sen. Patrick Leahy). Frequent references during the legislative process to critical language appearing in court opinions about civil forfeiture practices showed that Congress heard and valued the concerns of appellate courts on this subject. See id.

181 HYDE, supra note 94, at 4.


183 See Advertisement, American Civil Liberties Union, N.Y. TIMES MAG., April 29, 2001, at 135 (containing a picture of Uncle Sam pointing a gun at the reader with the statement below “I want your Money, Jewelry, Car, Boat and House.” The advertisement continued, “The forfeiture laws were designed as a new government weapon in the ‘war on drugs.’ But they’ve done little more than provide law enforcement with a license to steal.”). See also Latest ACLU
devoted substantial investigative resources to exposing the ills of civil forfeiture practices. The *Pittsburgh Press* published a six-day series reflecting ten months of national research on civil asset forfeiture in which reporters reviewed 25,000 drug seizures; interviewed 1,600 prosecutors, defense lawyers, cops, federal agents, and victims, and looked at court documents in 510 cases. The multi-part series concluded that “seizure and forfeiture, the legal weapons meant to eradicate the enemy, have done enormous collateral damage to the innocent.”184

These varied writings were bolstered by empirical data which had a profound impact upon elected officials. Legislators routinely consider policy issues that are informed by empirical research.185 Empirical data plays an important part in justifying the need for reform, especially when it involves research methods that extend beyond quantitative data and include research based upon observation and experience.186 Empirical research is frequently used as a basis for amending statutes in the legislative process,187 and while it does not necessarily provide answers to policy questions, it does raise the level of policy debate and can improve its conclusions.188 If legislators are to make factual assumptions in legislation they sponsor or support, it is important that those assumptions be grounded in fact and empirical data can help to provide them with a better understanding of how laws play out in

---


188 *Id.* at 197.
On the Road to Civil Gideon

the real world.189

Social science research is critical not just to the legislative process, but also to judicial decision making. It occupies a valuable role in the way that courts consider how information about social reality contributes to shaping the way society should be ordered,190 and in how many leading legal questions should be resolved, such as the size of a jury in a criminal case,191 the application of Rule 11 of the Federal Rules of Civil Procedure,192 or capital punishment decision making.193 Many agree that better information through social science data improves public policymaking.194 Certainly, policy wonks urged Congress to base the nation’s drug policy on

189 See Peter J. Smith, New Legal Fictions, 95 GEO. L. J. 1435, 1448 (2007) (discussing the role of empirical data and social science research upon judicial decision making and the formulation of legal rules).


192 See Carl Tobias, Some Realism About Empiricism, 26 CONN. L. REV. 1093, 1098 (1994) (discussing the important role that empirical data on Rule 11’s application and use played in the work of public policymakers, such as the Advisory Committee on Civil Rules, in formulating proposed amendments to the Rule).


194 See, e.g., Tobias, supra note 192, at 1103. There is, of course, considerable debate about the role that social science should play in the decision making of legal institutions. There is understandable tension between social science methodology, which “seeks quantifiable precision” in its measurements and attempts to statistically control for the effects of multiple variables, and the work of legal institutions that operate in “complex circumstances influenced by an indefinite variety of known and unknown factors.” See also Robert F. Schopp, The Nebraska Death Penalty Study: An Interdisciplinary Symposium, 81 NEB. L. REV. 479, 483 (2002). This debate is beyond the scope of this Article.
science and research, not ideology.\textsuperscript{195} While empirical studies are essential, their effectiveness can be undermined if the data is too voluminous, complex, or indecipherable to be really helpful to decision-makers. In the legislative process, a simple statistic that captures the essence of the problem as it affects real people can serve effectively as a rallying call for change. In the movement to obtain civil forfeiture reform, one such statistic emerged as an effective rallying cry: \textit{At least 80 percent of all civil forfeiture cases go unchallenged, without benefit of counsel.}\textsuperscript{196} Over and over again, this statistic reminded legislators of the inescapable fact that the uphill challenge of litigating civil forfeiture cases and defending private property against the government resulted in at least eight out of every ten cases going uncontested.\textsuperscript{197}

Elected officials repeatedly asked law enforcement authorities to explain why there was such a high uncontested rate in civil forfeiture cases. They wanted to know who were the property owners who did not contest forfeiture.\textsuperscript{198} Were property owners walking away from their property as a way of avoiding an indictment?\textsuperscript{199} Or was it just too expensive or difficult to obtain a...
lawyer in order to litigate against the government? This one, simple statistic—that 80 percent of all civil forfeiture cases went uncontested—spoke powerfully and frequently to the fact that something was definitely wrong. Ordinary citizens were stripped of their property simply because they lacked access to the courts to stand toe-to-toe with the government, which, ironically, is always represented by a lawyer in court. The cases were often too expensive to litigate in relation to the value of property at stake.\textsuperscript{200} For low-income citizens who cannot afford a lawyer under any circumstances, they simply had no choice but to walk away even when they had meritorious arguments to present.

At its core, the ability to have a lawyer at one’s side during frightening times when confronting superior governmental resources is what enables citizens to have access to the courts, a promise that is basic to a well-functioning democracy. The truth is that access to our courts is linked closely to having a lawyer by one’s side,\textsuperscript{201} and the absence of counsel unquestionably accounted for such a high uncontested rate in civil forfeiture cases.\textsuperscript{202}

The 80 percent uncontested statistic was deeply disturbing to lawmakers. It posed a haunting question to which there never surfaced an entirely satisfactory answer, other than that there must be a serious flaw in the system. The success of our civil justice system depends upon ordinary citizens being able to tell their side of the story before any official action takes place that deprives them of their hard-earned gains. How is it that eight out of every ten property owners would simply walk away from their property without a fight, an explanation, or a defense?

If individuals were allowing their property to forfeit to the

\textsuperscript{200} For example, one civil forfeiture victim, Richard Apfelbaum, was a salesman carrying slightly less than ten thousand dollars on his way to Las Vegas to gamble when drug enforcement agents stopped him and conducted a consensual search. \textit{See} \textit{Levy}, supra note 182, at 131. Upon finding the money, the agents confiscated the money and left him only thirty dollars to return home. \textit{Id.} Initially, Apfelbaum fought back by posting a bond and hiring an attorney, but over time with attorneys’ fees mounting, he gave up and was quoted as saying “I’m not in a position to spend $10,000 trying to get $9,000 back.” \textit{Id.}

\textsuperscript{201} \textit{Kaufman}, supra note 53, at 372.

\textsuperscript{202} \textit{See} \textit{H.R. REP. NO. 105-358, pt. 1 at 28–29 (1997).}
government without a fight when they had a legitimate defense either because the burdens of the law were set too high or their defenses could not be presented without the help of a lawyer, confidence in the rule of law was clearly at risk. Eric Holder, then a high ranking member of the Justice Department, expressed this vital concern when he testified that “no tool of law enforcement, however effective at fighting crime, can survive for long if the public thinks that it violates the basic principles of fairness and due process that lie at the core of the American system of justice.”

Congressman Hyde stated that he simply wanted to “give the average citizen who is not a sheriff, who does not have a relative in city council,” an opportunity at obtaining due process of law.

The second lesson of civil forfeiture reform is that narrative stories, while powerful, are usually not enough by themselves to bring about needed change. Narrative stories gain real strength when joined with official validation from high sources and respected institutions, such as appellate courts, and when a broad range of voices are raised through academic articles, books, and investigative reports published in leading newspapers. And, significantly, these voices can crystallize behind a powerful, simple statistic which suggests that the game as it is playing out in the real world is fundamentally unfair. In the civil forfeiture context, that statistic was one which presented a haunting question that had no obvious answer other than that the system was simply unfair. A challenge for the civil Gideon movement will be to find an equally powerful statistic that highlights a fundamental flaw in the civil justice system when lawyers for the poor are absent.

This second lesson further illustrates how simultaneous efforts to obtain a right to counsel in both judicial and legislative forums paid off. While court efforts alone did not succeed, the observations and concerns expressed by appellate courts provided a powerful impetus for legislative change and focused national attention on the need for reform to protect basic rights. The civil


Gideon movement, as well, would benefit from continuing validation from appellate courts of the failings of the civil justice system when lawyers are not present, and much more can be said in books, newspapers, and social media about the unfair advantages conferred upon government and well-resourced parties in civil cases simply because an individual on the other side is too poor to afford a lawyer.

C. Lesson Three: Civil forfeiture reform attracted bipartisan support from a broad coalition of diverse interests that made legislative compromise possible and paved the way for the adoption of a right to counsel for indigent homeowners whose primary residences were at risk.

With compelling narrative stories from ordinary citizens and mounting concern expressed by courts, academics, and the media, reform advocates assembled a broad-based coalition of diverse interests that made bipartisan support and legislative consensus possible. The broad-based support for reform came from across the political spectrum and set the stage for reasonable compromise when difficult issues threatened to halt legislative progress. The depth of support from diverse organizations made the threat of a more robust right to counsel appear credible, forcing detractors to reach compromise on a more limited right to counsel for indigent homeowners facing seizure of their primary residence.

The civil forfeiture reform movement brought together the full range of political and philosophical ideology and presented a coalition of organizations that only come together once every ten years. This broad coalition of groups helped to advance legislation that struck the “right balance” in making common sense and adopting fair and equitable procedures. The movement attracted diverse support from such groups as the National Rifle Association, the American Civil Liberties Union, and the U.S.

---

Chamber of Commerce. Like many other successful social movements, this reform initiative achieved results through the “participation of diversely located subjects whose immediate and direct interests might not coincide with those of the group’s agenda.” As Senator Leahy stated after agreement had been reached on the reform package, “[i]t is not often that we see the U.S. Chamber of Commerce, ACLU, NRA, National Association of Criminal Defense Lawyers, American Bankers Association, the Institute of Justice, Americans for Tax Reform, and the American Bar Association joining together on the same side of a legislative effort.”

The reform movement’s themes of promoting fairness and protecting private property owned by innocent citizens resonated with both ends of the political spectrum. The fact that innocent property owners (those not even charged with an offense) enjoyed fewer rights than individuals accused of a crime when it came to gaining access to counsel in forfeiture cases was a troubling thought. By providing an expanding right to counsel, Congress was able to address this problem which in turn lessened the concern that civil forfeiture practices might be adversely impacting racial minorities and encouraging racial profiling practices.

Senator Lindsey Graham supported an expanded right to counsel in defense of one’s property against the federal government. He favored appointing counsel because it guaranteed one’s day in court. In his opinion, standing alone was “no place to be” when one’s property was seized by the government. When fighting the government for one’s own property, Senator Graham stated that it was only right that Congress provide for counsel.

The use of contrasts can be a powerful tool of persuasion. By
On the Road to Civil Gideon

focusing on stark contrasts between a right to counsel at public expense for property owners accused of a crime, from those who had no similar right simply because they faced no allegations of criminal wrongdoing, Congress struck a responsive chord. While a court was authorized to appoint counsel for an accused in a civil forfeiture case if it found certain statutory factors to be present,\(^1\) innocent owners not accused of criminal wrongdoing appeared to enjoy no such protections. As a result, Senator Leahy supported a limited right to counsel permitting courts to authorize counsel to represent an indigent claimant when the claimant is already represented by a court-appointed counsel in connection with a related federal criminal case. He found this to be a fair compromise, one that eliminated “any appearance that the government chose to pursue the forfeiture in a civil proceeding rather than as part of the criminal case in order to deprive the claimant of his right to counsel.”\(^2\)

Representative Barney Frank shared this view, pointing out that the loss of property in a civil proceeding is no less damaging to an individual than if it is lost as a result of a criminal conviction.\(^3\)

There was disagreement in Congress about how far to extend a right to counsel in civil forfeiture cases. Some senators expressed the view that the House of Representatives had gone too far in its bill when it created a general right to counsel such that it was creating legal aid clinics for property owners in civil forfeiture cases.\(^4\) However, where an indigent homeowner’s primary residence was at stake, leading senators recognized that possible eviction or even homelessness might result from the forfeiture of property.\(^5\) As expressed by Senator Leahy, “[w]hen a forfeiture

\(^1\) H.R. Rep. No. 105-358, pt. 1, at 29 (1997) (a court could appoint counsel if it found three factors identified in 983(d)(A) –(C) of Title 18 to be present).


\(^4\) See Civil Asset Forfeiture Reform Act, S. 1931, 106th Cong. § 981A(b) (1999) which takes a different approach from the House on the issue of appointment of counsel. The Senate narrowed the right to counsel to indigent homeowner’s whose primary resident was at stake. See infra note 220.

action can result in a claimant’s eviction and homelessness, there is more at stake than just a property interest, and it is fair and just that the claimant be provided with an attorney if he cannot otherwise afford one. 219

This basic concern, shared across the broad coalition of supporters of reform, led to compromise that provided for a right to counsel at public expense for indigent homeowners whose primary residences are the subject of civil forfeiture proceedings. 220 The new legislation designated the Legal Services Corporation to administer this limited right to counsel for indigent homeowners. The stage for this compromise was set when the House Judiciary Committee included strong overall right to counsel language in the House bill. The Senate reached agreement when the Hatch-Leahy Civil Asset Forfeiture Reform Act 221 was joined with proposed language from the Sessions-Schumer bill, culminating in a Hatch-Leahy-Sessions-Schumer substitute amendment that was passed by the Senate on March 27, 2000. 222 While the appointment of the Legal Services Corporation (“LSC”) for the implementation of this important new limited right to counsel was not something specifically requested by LSC, bipartisan compromise in the Senate emerged as a means of heading off a broader right to counsel that might otherwise have applied to all civil forfeiture actions. 223

The third lesson of civil forfeiture reform is that broad-based

219  Id.
220  See 146 CONG. REC. S1753, S1759-60 (2000). The broad coalition of support led to an agreement on March 26, 2000, with Hyde, Leahy, Sessions, and Schumer coming together. They were influenced by the knowledge that Vermont Law, Vt. Stat. Ann. Tit. 18 § 4241(a)(5), did not permit forfeiture of real property occupied as primary residence of a person involved or a member of that person’s family. Id.; see also Vt. Stat. Ann. Tit. 18 § 4241(a)(5) (2009).
223  Senator Sessions opposed efforts to have CJA counsel provide representation as a matter of right in such cases, concerned that such an extension would be “camel’s nose” to a broader right to counsel. Instead, he agreed to a compromise offered by Senator Leahy of substituting LSC attorneys for CJA attorneys. See DAVID D. SMITH, PROSECUTION AND DEFENSE OF FORFEITURE CASES 11–12 n.14.6 (1999).
support from organizations representing the full political and ideological spectrum enables a controversial bill to move through the legislature, even in the face of powerful opposing interests. Moreover, such broad and deep support can set the stage for reasonable compromise that allows the legislation to move forward. Here, the threat of an unconditional right to counsel in all civil forfeiture proceedings was deemed credible because of the depth of support from diverse organizations and interests that rarely agree. As a result, opposing views were able to reach compromise on a more limited right to counsel when it was specifically targeted to the most serious need (protecting primary residences) of those individuals most in need (indigent homeowners).

The civil Gideon movement has attracted strong support from the organized bar and from those organizations directly involved in the delivery of civil legal assistance to the poor. But the question remains whether the current coalition is broad enough to wield sufficient legislative influence that will be needed when legislation proposing such a right is challenged by opponents on philosophical or cost grounds. At those times, will the movement to establish a civil Gideon be said to have the support of diverse interests that come together only once every ten years? If the answer is likely not, efforts should be undertaken to expand the base of support, looking for ways to strike common ground with business, governmental, religious, and other interests not always associated with efforts to support the delivery of legal services to the poor. Civil forfeiture reform demonstrates the value of having this support and offers hope that it is indeed possible.

D. Lesson Four: A proposed right to counsel in civil forfeiture proceedings was not viewed as an end in of itself, but rather as a means of insuring that other needed civil forfeiture reforms, if adopted, would succeed.

Lawyers readily understand the vital role that they play in achieving the goals of their clients and, not surprisingly, therefore regard a right to counsel as a noble end in of itself. The current right to counsel coalition certainly ties its raison d’etre to achieving
this procedural goal, even though more substantive social justice interests are likely the primary concerns of this movement. This approach makes good sense to members of the legal profession whose training and experience strongly inform their perspective. But this view threatens to be too limiting and may even be counterproductive as the general public does not necessarily place such a high premium on procedural protections. The public is more likely to demand legislative change when substantive concerns violate their notions of justice and fairness and undermine their respect for the law. Lawyers are important, for sure, but they play a secondary role to substantive interests that are front and center.224

This suggests that the public may not be inclined to rally behind a general right to counsel as an abstract principle of justice, however noble it is, but might be willing to demand of their elected officials such a right if they see it as necessarily connected to achieving an overriding substantive interest. In civil forfeiture reform, access to legal assistance was viewed as a necessary adjunct to protecting against the wrongful loss of private property to the government, especially when the property at risk was the family home. Procedure was tied tightly to substance. In other civil matters, the public may not fully understand the need for a lawyer. For example, the public might initially be unsympathetic to a general right to counsel at the public expense in all housing or real estate matters, but might be willing to demand such a right if there is not widespread confidence that legislation designed to eliminate lending practices that threaten homeownership can succeed without a lawyer to assist an indigent homebuyer. In other words, the creation of limited rights to counsel in other areas of civil law might be most viable where the success of important substantive reforms hinges upon the inclusion of such a limited right to

224 Lawyers naturally look to the courts first for support of their argument that there should be a right to counsel because they expect that judges, as lawyers, understand the close relationship between due process of law, fundamental fairness, and an adversary system in which both sides are represented by counsel. When lawyers turn to the legislature for support, it is not as clear that they will receive the same understanding or support about this relationship, especially as the percentage of lawyers serving as legislators declines across the country.
counsel to make sure that important legislative objectives succeed.

This was the case in civil forfeiture reform. While access to counsel was viewed as an important part of the reform package, it was not viewed as an end in of itself. Rather, the enactment of CAFRA represented the culmination of a decade-long effort to enact a broad range of reforms to insure that property was not wrongly taken from individuals by the government. As one ACLU representative testified, the appointment of counsel for indigents was absolutely needed in order to ensure that individuals would be able to avail themselves of the other reforms in the bill.225

This is not to suggest, however, that the public does not place a high value on the abstract principle of requiring access to legal help in civil matters, especially on behalf of poor or unsophisticated individuals who lack the means to battle superior governmental resources. But, in the case of civil forfeiture reform, legislators understood that representation by counsel was directly connected to the overall fairness of the entire adjudicatory framework. In other words, no matter how hard legislators tried to design fair civil forfeiture procedures, the process might never really be fair if a property claimant was forced to represent herself against the government.226 Chairman Henry Hyde embraced this view when he noted that in a democracy the means can be as important as the ends.227 Even a fairer law and simplified procedures might not succeed if a property owner had to face it alone.228


226 See Civil Asset Forfeiture Reform Act: Hearing on H.R. 1835 Before the Committee on the Judiciary, 105th Cong. 135 (1997) (statement of David Smith, National Association of Criminal Defense Lawyers) (“No matter how fair the formal civil forfeiture procedures are, the process can never really be fair if a claimant is forced to represent herself.”).


However, not everyone shared this basic belief. Eric Holder, then a deputy attorney general in the Department of Justice, testified that the proposed appointment of counsel in civil forfeiture proceedings was one of the two most objectionable provisions of the House version of the reform bill.\(^{229}\) He expressed concern that it was an incentive for abuse and, if adopted, would overburden courts with appointment motions. Some executive agency representatives also expressed concern that an expansion of access to counsel would increase the number of cases on crowded federal court dockets and encourage litigation of plainly forfeitable property interests.\(^ {230}\) Holder argued that the attorney’s fees provision in the Equal Access to Justice Act ("EAJA") provided sufficient protection for property owners. Under that federal law, attorney’s fees can be awarded to a prevailing party against the government if the government’s position is not substantially justified. While the Justice Department supported some aspects of forfeiture reform, it did not support the appointment of counsel as a matter of right.\(^ {231}\)

Law enforcement officers also testified against the proposed appointment of counsel in civil forfeiture actions, arguing that the appointment of counsel for indigents would divert significant assets to the criminal defense bar,\(^ {232}\) and only encourage attorneys
On the Road to Civil Gideon

739
to look for court appointments to file frivolous claims. 233 One South Carolina sheriff called the bill an “entitlement program for lawyers,” arguing that the law-abiding citizens should not be forced to pay for legal services for wealthy drug dealers and criminal syndicates to defend their criminal activities. 234 In short, some law enforcement representatives and government officials opposed a right to counsel on the basis that the potential for abuse was too great, there were insufficient safeguards, and federal law (EAJA) already provided for fee shifting under limited circumstances.

On the other hand, the American Bar Association strongly endorsed civil forfeiture reform, including the provision of counsel for indigents, on the basis that civil forfeiture law disregarded basic principles of due process. 235 The ABA relied upon the Second Circuit’s strong language in the *U.S. v. All Assets of Statewide Auto Parts* case, expressing concern about the lack of fairness in civil forfeiture proceedings. 236 Criminal defense experts argued that the appointment of counsel would provide a legally trained champion to get seized property back into the hands of the lawful owner, representing the first step toward achieving fundamental due process. 237


237 Civil Asset Forfeiture Reform Act: Hearing on H.R. 1835 Before the
Although he objected to an expanded right to counsel, Stefan Casella, a justice department official responsible for civil forfeitures, conceded that it was more important than ever that forfeiture laws operated fairly, that citizens enjoy access to the courts, and that property interests of innocent owners be fully protected.\textsuperscript{238} Describing the historical context for the nation’s use of civil forfeiture powers, Casella testified that federal forfeiture laws were used primarily in the past to forfeit items that had no legitimate basis, such as pirate ships, contraband goods, and whiskey stills. The war on drugs, however, had expanded the use of federal forfeiture laws to people’s primary possessions, including their homes, cars, and cash. The application of forfeiture laws to legitimate items of fundamental importance to citizens raised the stakes, according to Casella, and with it the need to ensure fundamental fairness.

With powerful interests on both sides of this question, a limited right to counsel emerged as the means to make civil forfeiture procedures fairer for ordinary citizens and to attack the high incidence of uncontested actions. With only a limited right to counsel to fund, Congress easily identified a ready source of funding in the proceeds of forfeited property. Property forfeited to the government had proven profitable and the civil asset forfeiture fund netted millions of dollars for law enforcement budgets and special drug fighting initiatives. These funds could accommodate relatively modest expenditures for counsel fees to assure that the operation of the forfeiture program was fundamentally fair and that real property belonging to indigent homeowners was protected from erroneous deprivation.

Of course, cost remains a huge concern when discussing the enactment of a right to counsel. Thus far, attempts to place a dollar figure on a right to counsel in civil matters are extremely

speculative.\textsuperscript{239} Especially in difficult economic times, this is not an easy issue to tackle. This is why it is so important to quantify through statistically reliable studies the savings that can be achieved by providing counsel. Some studies suggest that spending public dollars on civil legal services can save the public triple or quadruple the amount that would need to be spent later if counsel is not provided.\textsuperscript{240} One recent Texas study found that every dollar spent in the state for indigent civil legal services generated an additional $7.14 in total spending, $3.56 in output, and $2.20 in personal income.\textsuperscript{241} This resulted in an additional $457.6 million in business spending and the creation of 3,171 jobs.\textsuperscript{242} However, there are still relatively few economic studies of this kind and there are difficulties in measuring precisely such financial gains. Alternatively where, as here, a particular subject area of the law provides a readily-tapped fund, such as the federal civil asset forfeiture fund, the financial cost question is a much easier one to answer.

As the foregoing discussion makes clear, powerful interests are likely to disagree about the wisdom, desirability, or affordability of a civil right to counsel. This is why it may be helpful to reframe the basic question. The fourth lesson of civil forfeiture reform is that a civil right to counsel may be more widely acceptable if it is viewed not as the primary objective, but rather as the means by which to ensure the success of other important legislative goals. The prospects of establishing a right to counsel might also be improved by limiting that right to cases involving the most serious

\textsuperscript{239} Kaufman, supra note 53, at 384–85 (suggesting that even an estimate of five to nine billion dollars per year may be inadequate, based upon a calculation of fifty to ninety million eligible clients each year, spending an average of $100 per client).

\textsuperscript{240} See Bernice K. Leber, \textit{The Time for Civil Gideon is Now}, 25 TOURO L. REV. 23, 26 (2009) (citing a New York City Department of Social Services report that finds that for every dollar spent on indigent representation in eviction proceedings, four dollars in costs related to homelessness are saved).


\textsuperscript{242} \textit{Id.} at 3.
needs, at least initially. In a legislative forum, the focus should be on the substantive objectives of housing the poor, nourishing young children, or protecting the elderly. A right to counsel should not be the banner headline, but rather the means to making these laudable goals possible.

E. Lesson Five: The adoption of a limited right to counsel at public expense in civil forfeiture cases serves as further proof that Congress values the essential role that lawyers play in achieving important legislative goals and that providing lawyers makes a significant difference in case outcomes.

In Gideon, the Supreme Court acknowledged that the right to be heard would in many cases be of little avail if it did not include the right to be heard by counsel. Certainly, this turned out to be true in the retrial of Clarence Gideon’s criminal case. After the Supreme Court’s ruling, Gideon was retried by Florida authorities, but this time the trial court appointed an experienced lawyer to represent him.243 By some accounts, Gideon’s appointed lawyer skillfully picked apart suspect eyewitness testimony on cross-examination, suggesting in the process that the eyewitness who originally testified against Gideon was actually a lookout for a group of young men who were the real perpetrators of the crime.244 Gideon’s lawyer also introduced new key evidence from a taxi cab driver who drove Gideon from the pool hall on the day of the burglary and was able to refute the prosecution’s contention that Gideon was carrying stolen items when he left the pool hall. Counsel did a masterful job of defending Gideon245 and at the

243 Interestingly, the ACLU offered to represent Gideon on retrial, but Gideon wanted an experienced local lawyer. With Gideon’s agreement, the Florida court appointed W. Fred Turner, a lawyer with an excellent reputation for criminal defense representation in the court’s jurisdiction. See Jacob, supra note 8, at 257. The Gideon trial occurred five months after the Supreme Court’s ruling in Gideon.

244 See ANTHONY LEWIS, GIDEON’S TRUMPET (Random House 1964); Jacob, supra note 8, at 265.

245 Jacob, supra note 8, at 269 (Turner did a masterful job of defending
On the Road to Civil Gideon

743

conclusion of the trial the jury acquitted Gideon after only one hour of deliberation.\textsuperscript{246}

There is little doubt that a trained lawyer makes an important difference in legal proceedings. Empirical studies demonstrate that litigants in court and agency proceedings have a significantly greater chance of success when they are represented by counsel.\textsuperscript{247} Recent articles report that “studies of courts and administrative agencies consistently show that indigent litigants without counsel routinely forfeit basic rights, not due to the facts of their cases or the governing law, but due to the absence of counsel.”\textsuperscript{248}

In civil forfeiture actions, courts have readily acknowledged the importance of counsel because the legal framework is complicated, the cases are fact-intensive, and claimants must assert affirmative defenses or otherwise they will be waived. Needless to say, the government is always represented by a lawyer in civil forfeiture cases. The truth is that in civil forfeiture cases, having

\textsuperscript{246} LEWIS, supra note 244, at 237 (reporting that the jury went out at 4:20 pm and at 5:25 pm there was a knock on the door between the courtroom and the jury room, following which the jurors filed in and the verdict was announced).


\textsuperscript{248} Gary Toohey, Missouri Bar Ass’n, A Civil Right to Counsel: Inevitable or Unrealistic, PRECEDENT, 2010, at 25–26, available at http://members.mobar.org/pdfs/precedent/feb10/civil.pdf (citing BOSTON BAR ASS’N TASK FORCE ON EXPANDING THE CIVIL RIGHT TO COUNSEL, GIDEON’S NEW TRUMPET: EXPANDING THE CIVIL RIGHT TO COUNSEL IN MASSACHUSETTS (2008)).
counsel matters because there is so much that can be said.\footnote{Oversight of Federal Asset Forfeiture: Its Role in Fighting Crime: Hearing Before the Subcomm. on Criminal Justice Oversight of the S. Comm. on the Judiciary, 106th Cong. 85 (1999) (statement of Roger Pilon, Vice President for Legal Affairs, Cato Institute) H.R. 1835, June 11, 1997 Testimony of Roger Pilon.}

The first decade of experience applying CAFRA’s right to counsel for indigent homeowners in civil forfeiture cases appears to lend support to the basic proposition that with counsel, the chances of improved outcomes and of preserving basic human rights are increased,\footnote{See generally Boston Bar Ass’n, Task Force on Civil Right to Counsel, supra note 248; Douglas L. Colbert, Ray Paternoster & Shawn Bushway, Do Attorneys Really Matter? The Empirical and Legal Case for the Right of Counsel at Bail, 23 CARDOZO L. REV. 1719 (2002) (concluding that when lawyers are appointed at pretrial bail hearings, large numbers of defendants avoided incarceration). See also Engler, Connecting, supra note 85, at 92 (concluding that available studies consistently reveal the importance of representation as an important variable in improving success rates and that disparities in case outcomes between represented and unrepresented parties shows that the presence of lawyers impacts case outcomes).} whereas without counsel, litigants face worse outcomes.\footnote{See Carter, supra note 57, available at http://www.abajournal.com/news/article/judges_say_litigants_increasingly_going_pro_se_at_their_own_/ (reporting the view of judges that self-representing is increasing and producing worse outcomes for litigants).} The first years under CAFRA have produced surprisingly few court appointments; arguably, the sample of cases remains too small to draw any firm conclusions. Nevertheless, the results are instructive.\footnote{When CAFRA was passed, the Congressional Budget Office projected that an automatic right to counsel involving primary residents of indigents would increase annual spending by approximately one million dollars over a five-year period. This suggested that, at fixed compensation levels established by the statute, LSC would be expected to represent as many as 285 eligible homeowners each year. See Louis S. Rulli, The Long Term Impact of CAFRA: Expanding Access to Counsel and Encouraging Greater Use of Criminal Forfeiture, 14 FED. SENT’G REP. 87, 89 & n.32 (2001). For reasons that are unclear and beyond the scope of this article, that projection has not come to pass. Instead, there were only twenty-eight LSC appointments during the entire period from 2000–2007. See Response to Freedom of Information Request from Legal Services Corp. to author (May 13, 2009) (on file with author).} What limited results are available point...
clearly toward the conclusion that the presence of counsel affects outcomes.

During the period of 2000 through mid-2007, federal courts appointed the Legal Services Corporation only twenty-eight times to represent indigent homeowners whose primary residences were the subject of civil forfeiture proceedings. Without a right to counsel, one might expect that roughly 80 percent of the cases would be uncontested or, in other words, roughly twenty-four of the twenty-eight cases would result in a forfeiture of property to the government by default. With appointed counsel, however, all of the cases were contested. In at least two of the cases, default judgments that had been entered originally by the court were set aside after counsel were appointed to represent indigent homeowners.

In many of the cases, the person alleged by the government to have engaged in drug trafficking and the homeowner were not the same person. Of the cases in which it was possible to discern from court filings the relationship between the alleged wrongdoer and homeowner, it appears that the alleged wrongdoer was the homeowner only 38.1 percent of the time, while in the majority of the cases (or 61.9 percent of the time), the alleged wrongdoer was someone other than the homeowner. Of the cases in which the alleged wrongdoer was not the homeowner, the alleged wrongdoer

---

253 *See* Response to Freedom of Information Request from Legal Services Corp. to author (May 13, 2009) (on file with author).

254 Admittedly, this is a broad generalization drawn from the 80 percent overall uncontested rate in civil forfeiture proceedings discussed previously in this Article. To know whether this high uncontested rate applies equally to civil forfeiture cases involving primary residences of property claimants would require extensive empirical analysis that is beyond the scope of this article.


256 It was possible to discern the relationship between the alleged drug trafficker and the homeowner in twenty-one of the twenty-eight cases. Of those twenty-one cases, the alleged drug trafficker and the homeowner were the same person in eight of the cases. *See* Response to Freedom of Information Request from Legal Services Corp. to author (May 13, 2009) (on file with author).
was a family member in 33.3 percent of the cases, involved in a relationship with the homeowner (but was not a family member) in 23.8 percent of the cases, and totally unrelated to the homeowner in 4.7 percent of the cases. Assuming that the government is able to show a nexus between the illegal drug activity and the home, the relationship between the alleged wrongdoer and the homeowner is very important because of statutory defenses, such as the innocent owner defense, that may be asserted to prevent forfeiture. A homeowner should be able to prevail in court so long as he or she can prove they did not know of or did not consent to drug activity at the property. The pool of cases decided under CAFRA so far suggests that a meritorious innocent owner defense was available to homeowners in 61.9 percent of the cases, thereby validating the critical importance of having a lawyer present to assert and prove such a defense.257

Case outcomes in this small universe of civil forfeiture cases also support the important role that a lawyer plays once provided to an indigent litigant. In some cases, the assistance of counsel enabled homeowners to hold the government to its legal burdens and ultimately to have the government’s forfeiture action dismissed.258 Even where criminal activity may have taken place, homeowners were still able to save their homes because they had not personally engaged in wrongdoing and their lawyers asserted innocent owner defenses that would have been otherwise waived.259 Here, legal help provided an effective means of checking government power and fulfilling congressional intent that innocent homeowners not forfeit their homes to the government.

In still other cases, where homeowners were allegedly responsible for the drug activity at the property, representation by a lawyer permitted the parties to explore creative resolutions that

257 In thirteen of the twenty-one cases, the property claimant (homeowner) was not the alleged wrongdoer. See Response to Freedom of Information Request from Legal Services Corp. to author (May 13, 2009) (on file with author).
halted wrongful activity and met legitimate governmental interests, but at the same time protected the interests of innocent children residing in the property. In cases such as these, the real victims of a forfeiture of the family home to the government are the innocent children who then face shelter disruption, educational displacement, and even homelessness. With help from a lawyer, an indigent homeowner is not inclined to just give up, but is able to negotiate for a possible settlement of the forfeiture action that offers the promise of protecting his or her children from falling further into poverty. This could take the form of holding the real property in trust for the children, selling the property and placing proceeds from the sale in a trust for the children’s educational needs, or even refinancing the home to pay the government a sum of money as a substitute for forfeiture while still retaining the family home for the benefit of the children. The presence of counsel on both sides of the forfeiture action allows for meaningful negotiations and amicable resolutions that often better serve competing societal interests.

This type of beneficial outcome was achieved in a North Carolina case in which a homeowner and his girlfriend resided in the family home with their minor children. The police arrested the homeowner for growing marijuana at the property after they were called to the home in response to a domestic violence complaint. The government commenced a civil forfeiture action against the real property and, without counsel for the homeowner, a default judgment was entered in favor of the government. After a lawyer was appointed under CAFRA to represent the indigent homeowner, the default judgment was set aside. Soon thereafter, the homeowner died leaving the home to his children in his will. Counsel raised defenses to the forfeiture action on behalf of the children as property claimants and negotiated a resolution that


261 Id.

authorized the sale of the property with the net proceeds to be placed in trust for the minor children.\textsuperscript{263} Without counsel, it is likely that the default judgment would not have been set aside and the children would have been dislodged from their home without receiving any financial support from the equity of the property.

It is difficult to draw firm conclusions from the relatively small number of case outcomes during the first seven or so years under CAFRA in which indigent homeowners received lawyers as a matter of right. Still, it is telling that of twenty-one cases in which a final outcome has now been reached (out of a total of twenty-eight appointments), three cases resulted in outright dismissals of the forfeiture actions and fifteen more resulted in amicable settlements that brought modest or substantial benefit to the homeowner, including the retention of family homes in many of the cases. In approximately 43 percent of the cases reaching final resolution, homeowners held on to their property without any payment or with just a relatively small payment of settlement monies to the government.\textsuperscript{264} These overall outcomes would not

\textsuperscript{263} Consent Order and Judgment of Forfeiture May 6, 2008, Real Prop. Located at 130 High Rock Acres Drive, Black Mountain, N.C., 2007 WL 1959245 (court order available from the PACER court document database).

\textsuperscript{264} In reviewing the twenty-eight cases provided by the Legal Services Corporation in response to the author’s freedom of information request, outcomes were coded in seven classifications: actions still ongoing, rulings in favor of the U.S., rulings in favor of the property claimant, settlements in favor of the U.S., settlements slightly in favor of the property claimant, settlements substantially in favor of the property claimant, and unavailable (documents sealed). Codings were based entirely upon document review obtained from the PACER court document database. Admittedly, discretion was used in classifying settlement outcomes as favorable to the U.S. or slightly or substantially favorable to the property claimant. The category of settlement outcome favorable to the U.S. was used when all of the property in question was forfeited to the U.S., with the claimant receiving nothing, or when the settlement amount was the equivalent of the value of the property. The category of settlement outcome slightly favorable to the property claimant was used when the claimant had to pay a substantial amount in lieu of forfeiting the property, or when the property was sold with the claimant receiving a right to some portion of the proceeds. The category of settlement outcome substantially favorable to the property claimant was used when the claimant only had to pay a modest amount in lieu of forfeiture of the property, or if the property was dismissed
have been possible without a right to counsel.

The fifth lesson from civil forfeiture reform is that case outcomes are directly affected by having a trained lawyer at one’s side. Congress understood this basic proposition when it insisted that indigent homeowners have a right to counsel as a vital part of its reform legislation. It recognized that providing counsel to the poor offered indigent homeowners the best chance of ensuring that reform measures succeeded, that an unacceptably-high uncontested rate in civil forfeiture actions decreased, and most significantly that family homes received adequate legal protection from erroneous government forfeiture. Moreover, the early data demonstrates that concerns expressed to Congress that a right to counsel would be abused or would unduly burden the courts were unfounded. While final tallies of case outcomes are still unfolding, the early data suggests that Congress’ judgment was absolutely correct.

CONCLUSION

Federal civil forfeiture reform offers important lessons for the civil Gideon movement. In the twenty-one years between the Supreme Court’s rulings in *Betts v. Brady* and *Gideon v. Wainwright*, there were substantial developments in the states that expanded the provision of counsel at public expense in criminal matters and created a climate that made the Supreme Court’s landmark decision in *Gideon* possible. With each local expansion, the national movement for a right to counsel in serious criminal proceedings grew stronger.

Almost thirty years after the Supreme Court’s decision in *Lassiter v. Department of Social Services*, the movement to obtain a civil right to counsel where basic human needs are at stake has found new momentum. The enactment of a statutory right to counsel as part of federal civil forfeiture reform adds significantly to this movement and provides valuable lessons on the dynamics of successful legislative change. While CAFRA provides a statutory right to counsel only for indigent property owners whose primary

from the proceeding as part of the settlement agreement. *See* Response to Freedom of Information Request from Legal Services Corp. to author (May 13, 2009) (on file with author).
residences are at risk, and does not protect all property owners in federal forfeiture proceedings (or for that matter in the much larger number of state forfeiture cases where the property interests of indigents are at greatest risk), the enactment of this limited federal right to counsel is still a powerful illustration of the vital role that lawyers play to ensure that the poor have meaningful access to the civil justice system and are able to protect their most important interests.

*Lassiter v. Department of Social Services* is most assuredly the civil parallel to *Betts v. Brady*. Just as *Betts* before it, *Lassiter* tilts the scales away from justice for millions of poor Americans when their most basic human needs are at stake. If judges are to “administer justice without respect to persons, and do equal right to the poor and to the rich,” as their oath requires, they need to preside over fair contests in which the poor have the benefit of a lawyer.\(^{265}\) The road to a civil *Gideon* may be difficult and long, but with each newly-created civil right to counsel in diverse subject areas of state and federal law, our nation draws closer to delivering on its promise of equal justice under law.

---

\(^{265}\) Supreme Court justices each take this oath as provided for in Title 28, Chapter I, Part 453 of the U.S. Code Oaths of Justices and Judges, 28 U.S.C.A § 453 (West 2010).
COLLECTING UNPAID ASSESSMENTS: THE HOMEOWNER ASSOCIATION’S DILEMMA WHEN FORECLOSURE IS NO LONGER A VIABLE OPTION

Rachel Furman*

I. INTRODUCTION

On the three-hour flight from Florida to New Jersey Mr. Albert Furman has plenty of time to contemplate his current challenges.¹ He is no longer concerned with snowy driveways in the winter or driving for miles to visit friends.² In fact, he abandoned these difficulties years ago when he moved to Florida and embraced the luxury of swimming pools and golf courses open year round.³ He also enjoys several other amenities in his housing development—a clubhouse, a gym, road maintenance, trash collection, and landscaping services.⁴ But now, on his flight to New Jersey, a less fortunate reality becomes clear to him.⁵ When he flies back home to Florida in less than one week, the assessments owed to his Homeowner Association (“HOA”) might increase.⁶ He may

* J.D. Candidate, Brooklyn Law School, 2012; B.A., Middlebury College, 2008. I would like to thank Jacki Belin, Mark Furman, Carly Furman, and Tomás Brennan for their continuous support and encouragement. I would also like to thank my Papa Al for inspiring me to pursue my Note topic. I also appreciate the assistance of the members of the Brooklyn Law School Journal of Law & Policy.

¹ Interview with Albert Furman, author’s grandfather, in Short Hills, N.J. (June 3, 2010).
² Id.
³ Id.
⁴ Id.
⁵ Id.
⁶ Id.
experience a decrease in services provided by his HOA; services previously guaranteed by the timely payment of his assessments. 7 He will go to the next HOA meeting to hear that more and more homeowners are unable to pay their dues and that the HOA has no way to recover that lost money. 8 As a member of the HOA for his community, this very real threat to his quality of life and investment in his home will preoccupy Mr. Furman during his week in New Jersey. 9 And he is not alone. 10

Currently, about one-fifth of the United States population lives in HOA communities. 11 As of 2010, there are 309,600 association-governed communities, 24.8 million housing units, and 62 million residents living in these communities. 12 HOA residents range in age and economic diversity, depending on the HOA community. 13 Some communities offer less costly living alternatives for first-time buyers because the similarity in housing structures allows for

---

7 Id.
8 Id.
9 Id.
12 Industry Data National Statistics, supra note 10. “Association-governed communities include homeowners associations, condominiums, cooperatives and other planned communities. Homeowners associations and other planned communities account for 52–55 percent of the totals above, condominiums for 38–42 percent and cooperatives for 5–7 percent.” Id.
Collecting Unpaid Assessments

inexpensive production and cheaper sales.\textsuperscript{14} Other communities attract wealthier individuals by offering services and regulations above those offered by local governments, such as enhanced security systems and landscaping services.\textsuperscript{15} Further, some HOA communities cater to a specific age group, such as retired individuals.\textsuperscript{16}

Unfortunately, the erosion of HOA communities has become a reality for many non-delinquent HOA residents, who watch helplessly as the association struggles to collect assessments from delinquent homeowners.\textsuperscript{17} Failure to collect assessments results in reduced amenities, increased fees for homeowners who diligently pay their assessments, and the possible demise of the community.\textsuperscript{18} Mr. Furman’s dilemma is thus representative of a problem that plagues millions of Americans, and requires a timely and workable solution.

This Note argues that HOAs should require that all homeowners pay an initial security deposit in order to incentivize homeowners to pay assessments on time and to ensure that the burden of unpaid assessments falls to the defaulting homeowner. Part II provides background information on HOAs, and considers their organizational structure and financing techniques. Part III

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{14} See Lacy, supra note 13; Roger Showley, \textit{Architect Who Brought Modernism to the Masses Shares Lessons of His Success}, \textsc{SIGNON SAN DIEGO} (Sept. 19, 2010, 7:24 PM), http://www.signonsandiego.com/news/2010/sep/19/william-krisel-architect-stars-and-first-time-buye/ (explaining why tract houses are less expensive to build and are therefore less expensive to purchase for first-time buyers).
\item \textsuperscript{15} Blakely, supra note 13, at 198.
\item \textsuperscript{16} Id. at 198, 202.
\item \textsuperscript{18} See id.; Gemma Giantomasi, Note, \textit{A Balancing Act: The Foreclosure Power of Homeowners’ Associations}, \textsc{72Fordham L. Rev.} 2503, 2512 (2004) (noting the possibility that “nonpayment could cause the HOA’s cash flow to become unstable and unpredictable, which would defeat the very reasons for the creation of a community having mutual covenants, restrictions, and obligations”) (internal quotation marks and citation omitted).
\end{itemize}
\end{footnotesize}
explains the current problem HOAs face in relying on foreclosure as a means to compel payment from its members. Part IV considers the advantages and disadvantages of alternative means of enforcing payment, such as better drafting in the HOA’s governing documents and rearranging lien priority, ultimately finding that the risks of these methods outweigh the benefits. Part V concludes that using security deposits as a means of enforcing assessment payments provides the greatest protection for HOAs and their members by incentivizing all members to pay their assessments on a timely basis.

II. BACKGROUND ON HOMEOWNER ASSOCIATIONS

HOAs are organizations created by developers where the members possess an interest in common area assets. The HOA is charged with ensuring common areas are properly developed and maintained for the use and enjoyment of its members. HOAs aim to protect property values by maintaining the property itself and by preserving the character of any structures located on the property. They fund maintenance of the property by burdening the individual properties with a servitude demanding that the property owner pay assessments to the HOA. HOA communities require that homeowners pay assessments proportionally to cover the costs the association expends in maintaining the community.

Assessments are not optional charges or membership dues, but rather represent a proportion of the HOA’s total expenses divided among its members to cover the costs of both running the association and providing services to the residents.

---

20 See Giantomasi, supra note 18, at 2505.
21 See id.
22 RESTATEMENT (THIRD) OF PROP.: SERVITUDES § 6.5 cmt. a (2000).
23 See id. § 6.5 (1)(a).
24 Giantomasi, supra note 18, at 2509–10. Further, each HOA must collect
Collecting Unpaid Assessments

provide the sole source of income for HOAs, and HOAs create their operating budgets based on the amount of assessments they collect from each homeowner. Accordingly, if a homeowner fails to pay his assessments—or worse, if several homeowners fail to pay their assessments—the HOA loses money necessary to perform its central functions and the other homeowners are forced to bear the negative consequences.

Given their dependence on assessments for survival, HOAs treat collection of assessments as a critical matter. The HOA secures its assessments by attaching a lien to each individually owned property. The homeowners are aware of the lien because the HOA includes this information in its governing documents given to the homeowners prior to purchase. Since courts recognize the importance of assessments for HOAs’ survival, they usually read governing documents to contain an implied lien for unpaid assessments, even if the documents do not explicitly state that such a lien exists.

Through collection of assessments, HOAs provide their communities with all-inclusive amenities and aim to preserve and enhance property values. For example, HOAs offer benefits, such as capital improvements, landscaping, and pool care. Additionally, HOAs supply certain services usually provided by local government, such as road maintenance, utility services, street

assessments in accordance with state law and its own covenants. See id. at 2510.

25 See id. at 2512.
26 See id.
28 Id. §§ 6.5 (1)(b), 6.5 cmt. d.
29 See id. § 6.5 cmt. b (2000); Giantomasi, supra note 18, at 2509.
31 See Rich, supra note 19; Joyner, supra note 11.
32 Joyner, supra note 11.
and common area lighting, and garbage collection. HOAs can even enact stricter regulations governing the community than can cities or counties, often ensuring greater protection against atypical housing designs or failure to maintain individual premises. Many communities also provide recreational areas, including pools, tennis courts, gyms, clubhouses, and golf ranges, for homeowner use and enjoyment; this proves especially valuable since homeowners likely cannot afford to fund these luxuries independently. Therefore, without the collection of assessments, the HOA cannot successfully maintain and repair the communities’ amenities and homes, resulting in decreased housing prices for the community and an overall decline in appearance and communal facilities.

III. THE CURRENT PROBLEM FACING HOMEOWNER ASSOCIATIONS

The inability to maintain community aesthetics and amenities presents a common problem for HOAs that are unable to collect assessments from delinquent homeowners. In the past, HOAs could often force delinquent homeowners to pay their unpaid assessments through the threat or use of foreclosure. HOAs used

33 See id.; Giantomasi, supra note 18, at 2521. Cities and counties often allow HOA to take over such roles since this provides an attractive alternative to raising taxes. See Rich, supra note 19.

34 See Rich, supra note 19. Additional examples of the HOA’s authority include its ability to “dictate house paint colors, lawn-mowing schedules and parking policies for recreational vehicles.” Id.

35 See RESTATEMENT (THIRD) OF PROP.: SERVITUDES § 6.5 cmt. b (2000); Rich, supra note 19. Additionally, providing recreational benefits helps drive up property values in the community, thus enhancing the wealth of the unit owners. See RESTATEMENT (THIRD) OF PROP.: SERVITUDES § 6.5 cmt. b (2000) (noting that “[d]eteriorating common property and facilities . . . are likely to depress property values, decreasing property-tax revenues as well as the wealth of the property owners”).

36 See RESTATEMENT (THIRD) OF PROP.: SERVITUDES § 6.5 cmt. b (2000); Joyner, supra note 11.

37 See Anderson, HOA Groups in Arizona, supra note 17.

Collecting Unpaid Assessments

to rely on foreclosure as a final enforcement mechanism in order to force payment or to replace the delinquent homeowner with a homeowner able to pay his fees.\textsuperscript{39} Members faced with the possibility of losing their house would often find means to pay their unpaid assessments.\textsuperscript{40} HOAs who relied on the threat of foreclosure to collect unpaid assessments currently face a difficult situation. Taking ownership of a house in a poor economic environment often presents a greater financial burden to the purchaser than simply letting the homeowner remain in his home despite his inability to pay the bank or HOA.\textsuperscript{41} This problem arises in two similar situations.

\textbf{A. Two Current Foreclosure Situations: (1) When the Homeowner Fails to Pay Both His Mortgage and His Assessments and (2) When the Homeowner Pays His Mortgage but Fails to Pay His Assessments}

In the first situation, the homeowner defaults on his mortgage and cannot pay the bank.\textsuperscript{42} Often in this situation, the homeowner who has stopped making his mortgage payments has also stopped paying his HOA assessments.\textsuperscript{43} Given that in poor economic conditions mortgage values often exceed housing values, coupled with the reality that there are rarely buyers, the bank is disinclined to foreclose even though the homeowner is not making his mortgage payments.\textsuperscript{44} Buyers are scant due to problems such as

\textsuperscript{39} See id.
\textsuperscript{40} See Rich, \textit{supra} note 19.
\textsuperscript{41} See Liz Pulliam Weston, \textit{Your Lender Doesn’t Want Your House}, MSN MONEY (Oct. 22, 2008), http://articles.moneycentral.msn.com/Banking/HomeFinancing/YourLenderDoesntWantYourHouse.aspx.\textsuperscript{\textsuperscript{infra }Part III.A.}
\textsuperscript{42} Keith Jurow, \textit{Are HOA Foreclosures a Necessary Tool or an Extortion Racket?}, BUSINESS INSIDER (July 28, 2010, 8:27 AM), http://www.businessinsider.com/are-hoa-foreclosures-a-necessary-tool-or-an-extortion-racket-2010-7.
\textsuperscript{43} Id.
“rapid price depreciation” and “extreme unemployment issues,” which currently shape the economic environment. Foreclosure is thus an unattractive option because banks will have a difficult time finding a buyer and in the interim will be forced to pay the taxes on the house.

HOAs can try to force the bank to foreclose on the delinquent homeowner’s unit in the situation where the bank is reluctant to foreclose and the homeowner has stopped paying both his mortgage and his assessments. The HOA can foreclose its own lien on the unit and have an auction sale subject to the first mortgage owned by the bank. The unit buyer then purchases the unit with the original mortgage in place. In the past, the HOA’s foreclosure process typically frustrated the bank, since the bank hesitated to allow someone else to live in the unit aside from the original borrower. This would often force the bank to foreclose in

101386052; Weston, supra note 41.


47 Telephone Interview with Ira Goldenberg, Adjunct Professor of Law, Brooklyn Law Sch., Partner, Goldenberg & Selker LLP (Oct. 7, 2010).

48 See Nick Timiraos, Condo Boards Take On Lenders: Chasing Unpaid Dues, Associations are Foreclosing on Units Seized by Banks, WALL ST. J., Jun. 18, 2009, at A3.


50 The bank is not as comfortable with a purchaser who takes title to the property subject to an existing mortgage. If the new purchaser stops making his mortgage payments the bank can take the property, but there is no further personal liability for the purchaser, unless the contract between the purchaser and the grantee explicitly states such liability exists. See Bayou Land Co., 924 P.2d at 152 (“If the buyer acquires the land subject to the encumbrance, the land
Collecting Unpaid Assessments

order to replace the old borrower with a new borrower who has obligations directly to the bank. The bank consequently had an increased incentive to initiate its own foreclosure action or continue the pending foreclosure process with renewed vigor. The HOA could begin the process of foreclosing its own lien in hopes that by simply initiating the process, the bank would be forced to act in order to ensure that it finds a new buyer with a new mortgage.

However, in times of economic instability, banks are generally less responsive to HOAs’ foreclosure proceedings. First, even if the HOA forecloses its own lien, it likely takes time to find a new buyer for the reasons discussed above. The problem of limited buyers is further compounded by the fact that the buyer must take the property subject to the pre-existing mortgage. This discourages many buyers from purchasing since now they take ownership of an encumbered unit rather than taking ownership free and clear. The bank is likely aware of the difficulty the HOA faces in finding a buyer and thus does not feel pressured to act

continues to secure the obligation but the buyer is not personally liable for the debt. However, the grantor or seller remains personally liable.”); Anne E. Melley, *Purchase of Property Subject to Lien or Mortgage*, 63 *Tex. Jur. 3d Real Estate Sales* § 58 (2011) (“Where property is purchased subject to an outstanding lien or mortgage, but without an express or implied assumption thereof, the purchaser does not become personally liable for the debt so secured by the property.”). This provides less incentive for the new purchaser to make timely mortgage payments since no personal liability attaches if he fails to do so. See id. Banks are aware of this and thus prefer that the individual living in the house subject to the mortgage is the original mortgage holder. Telephone Interview with Ira Goldenberg, *supra* note 47.

51 Telephone Interview with Ira Goldenberg, *supra* note 47.
52 Id.
53 Id.
54 See Saito, *supra* note 44 (explaining that in times of economic difficulty, lenders are unmotivated to foreclose).
55 See id.; *supra* Part III.A.
56 See *Bayou Land Co.*, 924 P.2d at 152.
57 See id. (“When a seller transfers real property that is encumbered by a deed of trust, the buyer may either take the property subject to the encumbrance or assume the obligation secured by the encumbrance itself.”).
immediately to foreclose on its lien.\(^{58}\)

Additionally, even if the HOA does find a buyer, the bank might not feel the same pressure to foreclose on the unit as it does in a strong economic environment.\(^{59}\) The bank is likely satisfied that the new buyer in the unit is paying the mortgage because now the bank can recover the mortgage on the unit.\(^{60}\) Given that many banks experience great difficulty finding buyers\(^{61}\) and are often left to pay the taxes on the house in the interim between instituting foreclosure and an eventual sale,\(^{62}\) the bank is in no rush to continue a foreclosure action once the HOA has already foreclosed its own lien.\(^{63}\) Even if the HOA exercises its power to foreclose on its own lien in hopes of causing the bank to continue or initiate its own foreclosure proceeding, the bank likely will not take any immediate action.\(^{64}\) Further, it might take the HOA a long time to find a buyer for the encumbered unit for which it now assumes responsibility.\(^{65}\)

In the second situation the homeowner might still be paying his mortgage, but has failed to pay his assessments.\(^{66}\) In this situation the bank has less involvement since the homeowner continues to

\(^{58}\) This is problematic for the HOA, since the purpose of the HOA instituting its own foreclose action was to force the bank to act quickly in its foreclose action. Telephone Interview with Ira Goldenberg, supra note 47.

\(^{59}\) Id.

\(^{60}\) See Bayou Land Co., 924 P.2d at 152; supra note 50 and accompanying text. While supra note 50 explains that banks prefer that the individual living in the house is the original mortgage holder, in times of economic instability the bank’s preference will likely shift. The banks priority in poor economic climates is ensuring that it receives the mortgage, regardless of whether the individual paying is the original mortgage holder. As supra note 50 explains, the new individual in the unit will still be responsible for paying the mortgage even though he does not assume personal liability for the mortgage.

\(^{61}\) Saito, supra note 44.

\(^{62}\) See Winokur, supra note 46.

\(^{63}\) See Saito, supra note 44; Weston, supra note 41.

\(^{64}\) See Saito, supra note 44.

\(^{65}\) See id.; Weston, supra note 41.

\(^{66}\) See Joyner, supra note 11 (“Some who have fallen behind on homeowners association fees—despite being current on their mortgages—have been sent into foreclosure by associations determined to collect.”).
Collecting Unpaid Assessments

pay his mortgage.67 This presents serious problems for HOAs, since some HOAs face delinquency rates of fifteen percent or greater, which leads to serious revenue losses.68 At this point, the HOA can decide to foreclose on its own lien in order to collect the unpaid assessments.69 However, the same problems that the bank faces in initiating foreclosure—it must pay taxes while trying to find a new buyer—exist when HOAs initiate foreclosure to collect unpaid assessments.70 Thus, it is an undesirable option for most HOAs.

B. The Problems Associated with Both Foreclosure Situations

Assuming either the bank or the HOA begins the foreclosure process, that respective party is responsible for paying taxes on the house until it finds a buyer.71 As Jim Svinth, the chief economist for LendingTree.com72 explains, foreclosure represents a bad option for lenders because they “lose money holding that house . . . . They have to maintain it, insure it, market it . . . until it sells.”73 In a depressed economy this is financially more burdensome than simply letting a homeowner stay in his house, despite his inability to pay his mortgage or his assessments.74

Further, even if the bank forecloses and finds a buyer, the HOA often still struggles to collect unpaid assessments.75 In poor

---

67 See generally Joyner, supra note 11 (explaining that HOAs, rather than banks, foreclosed on individuals who failed to pay assessments but still made timely mortgage payments to the bank).
68 Jurow, supra note 42.
69 See id.
70 See Winokur, supra note 46; Weston, supra note 41.
71 See id.
73 Weston, supra note 41.
74 See id.
75 See Dawn Wotapka, New Home Sales Stuck at Rock-Bottom, WALL ST. J. (Oct. 27, 2010, 2:37 PM), http://blogs.wsj.com/developments/2010/10/27/new-home-sales-stuck-at-rock-bottom/ [hereinafter Wotapka, New Home Sales Stuck at Rock-Bottom] (explaining that home sales are currently extremely low). Given that sale prices are so low, junior interests, such as the HOA’s lien for
economic conditions, the home will often be sold at a low price and the bank will be paid first, leaving little or no funds left over for the HOA. 76 While not completely uniform across the country, as “[s]ome jurisdictions give an association lien some priority over previously recorded liens in recognition of the importance of preserving the association’s income stream,” 77 most jurisdictions subordinate assessment liens to mortgage liens. 78 Once a house is sold through foreclosure, the bank gets its money first, and leftover funds are then distributed to other creditors, assuming leftover proceeds exist. 79 Given the present difficulty with securing a buyer, houses in foreclosure tend to sell at very depressed prices. 80 Accordingly, not enough money remains from the proceeds of the sale to pay the bank in full for its mortgage, let alone provide enough money to pay subordinate lien holders, such as the HOA. 81

Unpaid assessments often do not become the obligation of the bank due to the language of the HOA’s governing documents. 82 In unpaid assessments, are unlikely to recover proceeds from the sale. See, e.g., Bankers Trust Co. v. Bd. of Managers of Park 900 Condo., 81 N.Y.2d 1033, 1036 (1993) (a lien for unpaid assessments is subordinate to the bank’s mortgage); Coral Lakes Cmty. Ass’n v. Busey Bank, N.A., 30 So. 3d 579, 585–86 (Fla. Dist. Ct. App. 2010).

76 See supra note 75 and accompanying text.
78 See, e.g., Bankers Trust Co., 81 N.Y.2d at 1036; Coral Lakes Cmty. Ass’n, 30 So. 3d at 585–86.
79 See Coral Lakes Cmty. Ass’n, 30 So. 3d at 583–84.
80 Wotapka, New Home Sales Stuck at Rock-Bottom, supra note 75.
81 See supra note 75 and accompanying text. Florida provides an extreme example of how a foreclosure sale fails to help the HOA collect its unpaid assessments; in Florida the foreclosure can extinguish subordinated liens and the buyer does not always assume responsibility for the unpaid fees. See Crossland Sav. Bank, FSB v. Saffer, 580 N.Y.S.2d 813, 815 (Sup. Ct. Orange County 1992) (holding that “on the foreclosure of a first mortgage of a residential unit the lien for unpaid common charges is extinguished”).
82 See generally Coral Lakes Cmty. Ass’n, 30 So. 3d 579. For example, the court in Coral Lakes Cmty. Ass’n v. Busey Bank, N.A. stated “that because of the Declaration’s plain and unambiguous language subordinating any claim for unpaid HOA assessments to a first mortgagee’s claim upon foreclosure or deed in lieu of foreclosure, it controls and absolves the Bank, as first mortgagee, from
Collecting Unpaid Assessments

fact, many HOAs include a provision in their governing documents to “induce lenders to aid homeowners in purchasing property in the community by awarding them priority over the HOA’s claims for unpaid assessments.”

Banks are more comfortable giving mortgages to homeowners in HOA communities when the association assures the bank that, in the event of foreclosure, the bank will receive its money first. Additionally, homeowners find it easier to secure mortgages with this language in the HOA’s governing documents, thus helping the community attract more buyers. In the past, when foreclosure afforded HOAs a greater chance of recovering unpaid assessments, HOAs benefited from this relationship between the HOA, the homeowner, and the bank, since it brought more homeowners, and thus more income, into the community. However, now that foreclosure fails to guarantee the HOA any proceeds, HOAs face the negative consequences of such drafting in their governing documents.

Another difficulty with foreclosure is that often a long delay exists between when the bank sends the homeowner a foreclosure notice and when the foreclosure process concludes. In states such

liability for any assessments accruing before it acquires the parcel.” 

Additionally, the court in Shields v. Andros Isle Prop. Owners Ass’n explained that “[r]estrictions found within a Declaration are afforded a strong presumption of validity, and a reasonable unambiguous restriction will be enforced according to the intent of the parties as expressed by the clear and ordinary meaning of its terms[.]” Shields v. Andros Isle Prop. Owners Ass’n, 872 So. 2d 1003, 1005–06 (Fla. Dist. Ct. App. 2004). If a HOA has explicit language subordinating its lien to the bank’s lien, courts will give strong preference to upholding the language in the document. See id.

Coral Lakes Cmty. Ass’n, 30 So. 3d at 581 n.1.

See id. at 584.

See id.

See id. (noting that the HOA originally included language in its governing documents prioritizing the bank’s mortgage over its lien for unpaid assessments in order to “entice lenders to finance purchases in its community” and thus bringing more buyers into the community).

See Wotapka, New Home Sales Stuck at Rock-Bottom, supra note 75 (noting the depressed prices of houses in foreclosure).

See J. Craig Anderson, Arizona HOAs Turn to Hardball Tactics; Critics Call System Inequitable, Some Owners Hounded to Pay Dues; Others Not, ARIZONA REPUBLIC (Aug. 30, 2010, 12:00 AM), http://www.azcentral.
as Florida, where foreclosure cases currently overwhelm the courts’ resources and time, the courts move slowly processing all their foreclosure cases. The Florida Supreme Court Task Force on Residential Mortgage Foreclosure Cases, charged with providing guidance and solutions for mortgage foreclosure cases, described the situation in Florida as follows:

Picture this: the biggest road out of town. Now imagine it is rush hour. In a thunderstorm. Add that it is also a hurricane evacuation. A lane is closed due to construction delayed by budget impacts. Imagine the traffic jam. The clearest description of the impact of the foreclosure crisis and the following recession on Florida’s courts can be summarized by that picture. Imagine every car is a case . . . . The enormous increase in foreclosure filings has overwhelmed [the courts’] resources in many circuits and represents a caseload traffic jam that the infrastructure cannot meet in a timely and efficient manner . . . .

Since courts severely lack resources, they simply cannot move efficiently through foreclosure cases. The community suffers as a

---

89 Florida is an example of a judicial foreclosure state. See generally FLORIDA SUPREME COURT TASK FORCE ON RESIDENTIAL MORTGAGE FORECLOSURE CASES, FINAL REPORT AND RECOMMENDATIONS ON RESIDENTIAL MORTGAGE FORECLOSURE CASES (Aug. 17, 2009), available at http://www.floridasupremecourt.org/pub_info/documents/Filed_08-17-2009_Foreclosure_Final_Report.pdf [hereinafter TASK FORCE ON RESIDENTIAL MORTGAGE FORECLOSURE CASES]. In judicial foreclosure states, foreclosures are processed through the courts, while in non-judicial foreclosure states, statutes establish foreclosure requirements rather than the courts. Catherine Rampell, Answers to Your Questions on the Foreclosure Crisis, Part II, N.Y. TIMES, Oct. 21, 2010, http://economix.blogs.nytimes.com/2010/10/21/answers-to-your-questions-on-the-foreclosure-crisis-part-ii/. Therefore, the problems associated with the courts’ inefficient management of foreclosure cases relate only to those states requiring judicial foreclosure.

90 See generally TASK FORCE ON RESIDENTIAL MORTGAGE FORECLOSURE CASES, supra note 89 (explaining that delays exist in the foreclosure process due to lack of resources in the courts).

91 Id. at 4.

92 See id.
Collecting Unpaid Assessments

result of the slow pace of foreclosure cases because, during the proceedings, delinquent owners fail to pay assessments, and Florida does not require banks to pay the unpaid assessments until title has passed.93

Similar to the courts, the banks also proceed very slowly in foreclosure cases, impeding the HOA’s ability to collect unpaid assessments.94 If the foreclosure process takes a long time to complete, which is likely given that the “[r]epeated postponement of foreclosure-sale dates has become commonplace,”95 the homeowner is responsible for all the fees in the interim, which he likely will not pay.96 Amanda Shaw, president of Phoenix-based community-management firm Associated Asset Management, described how Phoenix is currently grappling with this problem.97 She explains that “[o]ver the past 18 months [as of August 2010], 120,000 Phoenix-area HOA members have received foreclosure notices, which often prompt the homeowners to move out in anticipation of a trustee sale and the loss of their home.”98 However, in most cases the lender still needs to complete the foreclosure process, and in the interim the original homeowner who moved out is responsible for paying the assessments.99

Even when the bank takes a long time to complete the foreclosure process, the bank is often not responsible for the assessments on the unit.100 For example, in U.S. Bank Nat’l Ass’n v. Tadmore, the Condominium Association101 filed a motion to

---

93 See id. at 26.
94 See Timiraos, supra note 48.
95 Anderson, Arizona HOAs Turn to Hardball Tactics, supra note 88.
96 See id.
97 See id.
98 Id.
99 See id.
100 See, e.g., Deutsche Bank Nat’l Trust Co. v. Coral Key Condo. Ass’n (at Carolina), 32 So. 3d 195, 196 (Fla. Dist. Ct. App. 2010); U.S. Bank Nat’l Ass’n v. Tadmore, 23 So. 3d 822, 823 (Fla. Dist. Ct. App. 2009). The bank might not be responsible for unpaid fees at all, or alternatively, the bank is not responsible for the fees during the pendency of the foreclosure action. See U.S. Bank Nat’l Ass’n, 23 So. 3d at 823.
101 While this case discusses a Condominium Association, the statute used to reach the decision applies to all mortgagees who “acquire title to a unit by
“compel the Bank to either proceed with its foreclosure action or to commence paying the $939.56 monthly maintenance fee for the decedent’s condominium unit to the Association.”102 The Association made this motion after one year of waiting for the Bank to foreclose on the property, during which time the Association was unable to collect any assessments owed by the unit owner.103 The Association claimed the court should use its equity power to order the bank to pay the assessments.104 The court rejected the idea that equity demanded that the bank pay assessments while the foreclosure proceeding was pending.105 Accordingly, the bank did not assume responsibility for paying assessments even in the face of a lengthy foreclosure proceeding.

However, the court did explain that the bank “may be required to pay condominium maintenance fees after it acquires title and then only in limited amount.”106 Yet, this does not offer the HOA much relief.107 In Florida, for example, the bank does not pay the association during the foreclosure proceedings, which can take approximately eighteen months or longer from the time the homeowner fails to make mortgage payments.108 Then the amount owed to the HOA is statutorily reduced to a small percentage of the expense the HOA incurred to maintain the property.109 This demonstrates that even those states providing some redress for HOAs who suffer during drawn-out foreclosure proceedings only allow HOAs to collect a minimal percentage of unpaid assessments

foreclosure” and thus applies to a situation where a bank forecloses on a unit owned by a HOA as well. See FLA. STAT. ANN. § 718.116 (West, Westlaw through ch. 274 (end) of the 2010 2d Regular Sess. of the 21st Legis. and ch. 283 of the 2010 Special “A” Sess. of the 22d Legis.).

102 U.S. Bank Nat’l Ass’n, 23 So. 3d at 823.
103 See id.
104 See id.
105 See id. at 823–24.
106 Id. at 823.
107 See Timiraos, supra note 48; TASK FORCE ON RESIDENTIAL MORTGAGE FORECLOSURE CASES, supra note 89, at 26.
108 Timiraos, supra note 48.
109 TASK FORCE ON RESIDENTIAL MORTGAGE FORECLOSURE CASES, supra note 89, at 26.
Collecting Unpaid Assessments

after the bank completes the foreclosure sale.110

The long delay between initiating foreclosure and completing the foreclosure process only seems to be worsening.111 In October 2010, banks began “suspending tens of thousands of foreclosures in the 23 U.S. states that handle foreclosure through the court system . . . pending a review of their paperwork practices.”112 The banks’ foreclosure practices were called into question after it became apparent that banks employed “robo-signers,” who signed thousands of documents each month without appropriate review or notarization.113 This creates verification problems for the bank.114 First, the robo-signers likely did not check every affidavit to ensure that the bank employing them actually owned the mortgage.115 Second, although the notary step is supposed to occur when the robo-singer signs the affidavit, it often occurred at a much later date.116 While the banks suspended the foreclosures to improve the verification process, this does not indicate that the banks will forgive the unpaid mortgages.117 Instead, “the document issue will delay the foreclosure, but won’t change the outcome.”118 Accordingly, HOAs are harmed because homeowners, while still responsible for their assessments, often stop paying once they receive foreclosure notice.119 The longer the foreclosure process takes, the longer the HOA goes without collecting its

110 See id.
112 Id.
113 Id.
115 Id.
116 Id.
117 Whelan, supra note 111; see also Shell, supra note 114 (“This is likely to mean delays to most people who are in foreclosure. This is not a silver bullet. The reality . . . is that most people signed a note, owe the money, and cannot pay.”) (internal quotation marks omitted).
118 Whelan, supra note 111.
119 See Anderson, Arizona HOAs Turn to Hardball Tactics, supra note 88.
assessments.120

IV. ALTERNATIVE MEANS OF ENFORCING PAYMENT: BETTER DRAFTING IN THE HOAS’ GOVERNING DOCUMENTS AND REARRANGING LIEN PRIORITY BASED ON COOPERATIVE LAW

Several hurdles exist that bar HOAs from collecting unpaid assessments through the foreclosure process in declining value markets.121 Yet, a range of solutions also exist for HOAs trying to collect unpaid assessments.122 The HOA can use better drafting techniques in creating its governing documents to explicitly state that it will be paid before the bank in a foreclosure proceeding or that the bank is responsible for paying a portion of the unpaid assessments after the sale.123 Another option for the HOA is to

120 This proposition assumes the homeowner stops paying his assessments once he receives notice of foreclosure.
121 See, e.g., Timiraos, supra note 48; Task Force on Residential Mortgage Foreclosure Cases, supra note 89, at 26.
122 There are several techniques HOAs can use to try to collect unpaid assessments, such as the imposition of late fees, sanctions, fines, and penalties. See Hyatt, Condominium and Homeowner Association Practice, supra note 30, at 120, 157–58. However, these methods offer little protection in a situation where the homeowner has stopped paying her mortgage, vacated the premises, and yet is still responsible for paying her assessments. Another option for the HOA is to withhold services. This too, however, leaves the HOA in a less-than-ideal situation. Now the homeowner who cannot pay her maintenance fees suffers the consequences, but so too does everyone else living in the community who must watch the subject property deteriorate. This has an adverse effect upon the neighborhood as a whole, both visually and economically. See id. at 159. HOAs can try other techniques as well, such as publishing a list of those homeowners who failed to pay their assessments in hopes that shame might convince the homeowners to pay; however, statutes now prohibit this tactic. See Lisa Magill, Posting Debtor Lists to Collect Delinquent Condo & HOA Assessments, Florida Condo and HOA L. Blog, (June 22, 2009), http://www.floridacondohaalawblog.com/2009/06/articles/assessments/common-expenses/posting-debtor-lists-to-collect-delinquent-condo-hoa-assessments/. Florida statutes also prohibit HOA from mailing anything to a debtor with words written on the envelope intended to embarrass the debtor or calling the debtor very early or very late at night. See id.
123 See Coral Lakes Cmty. Ass’n v. Busey Bank, N.A., 30 So. 3d 579, 585 (Fla. Dist. Ct. App. 2010); Wayne S. Hyatt, Condominium, and Home Owner
Collecting Unpaid Assessments 769

rearrange the lien priority structure between the HOA and the bank by borrowing aspects of cooperative law. However, while both of these options provide some benefits for the HOA, overall they discourage banks from providing mortgages to purchasers in HOA communities, stifling the growth and economic advancement of the HOA and its community. Lastly, the HOA can require that homeowners pay a security deposit upon purchase of a home in a HOA community. This provides the most viable solution for the HOA and its residents because it shifts the burden of unpaid assessments to the delinquent homeowners and incentivizes all homeowners to pay assessments in a timely fashion.

A. Better Drafting in the HOAs’ Governing Documents

The HOA can use explicit language when drafting its governing documents, allowing the HOA to recover its unpaid assessments before the bank recovers its mortgage payments. Alternatively, the HOA can include a provision in its documents stating that, even if the bank’s lien takes priority over the HOA’s lien, the purchaser, likely the bank, is still required to pay at least a portion of the unpaid assessments. HOAs can look to new state statutes for guidance in establishing provisions that demand payments from the purchaser but are not too onerous as to create a disincentive for banks to loan any money to buyers. However, state statutes might not provide enough protection for HOAs. Therefore, HOAs should borrow from the concept that the bank must pay a portion of the unpaid assessments, but should include in


Id. at 15.

See infra Part V.A–B.

See Coral Lakes Cmty. Ass’n, 30 So. 3d at 585.

See TASK FORCE ON RESIDENTIAL MORTGAGE FORECLOSURE CASES, supra note 89, at 26.

See id.
its governing documents a requirement that the bank will pay a high enough percentage of the unpaid assessments to realistically help the HOA in its recovery.

For example, after a foreclosure sale in Florida “the delinquent amount [of assessments owed to a HOA] is statutory reduced to a mere fraction of an association’s expense to maintain the property.”130 According to Florida’s Real and Personal Property laws governing HOAs, the purchaser is responsible to pay:

The parcel’s unpaid common expenses and regular periodic or special assessments that accrued or came due during the 12 months immediately preceding the acquisition of title and for which payment in full has not been received by the association; or . . . [o]ne percent of the original mortgage debt.131

Florida’s current statutory reduction to a mere fraction of what the HOA is owed presents an inadequate solution, since HOAs are unable to recover a fair portion of money owed from unpaid assessments.132 HOAs should therefore use state statutes as a guide, but not as an exact replica, of language to include in its governing documents ensuring an adequate payment from the bank of prior unpaid assessments. By including language ensuring that HOAs receive some payment for unpaid assessments, HOAs can put themselves in a more financially powerful situation with respect to banks, who otherwise are reluctant to give HOAs any money during or after the foreclosure process.133

While this might help the HOA recover more during a foreclosure sale, this benefit is outweighed by the negative consequences resulting from its implementation.134 First, if the

130 Id.
131 Fla. Stat. Ann. § 720.3085 (West, Westlaw through ch. 274 (end) of the 2010 2d Regular Sess. of the 21st Legis. And ch. 283 of the 2010 Special “A” Sess. of the 22d Legis.). However, in Coral Lakes Cmty. Ass’n v. Busey Bank, N.A., the Court notes that this statute did not apply retroactively and cannot be used to contradict the plain language of the governing documents. See Coral Lakes Cmty. Ass’n, 30 So. 3d at 583–84.
132 See Coral Lakes Cmty. Ass’n, 30 So. 3d at 583–84.
133 See generally Coral Lakes Cmty. Ass’n, 30 So. 3d 579.
134 See id. at 584.
Collecting Unpaid Assessments

language in the governing document states that the HOA could collect unpaid assessments before the bank received its money, banks would be reluctant to give mortgages to homeowners looking to live in HOA communities. This presents an unworkable solution since the bank was “the entity that previously made the . . . purchase of the home possible.” Punishing the bank for delinquent homeowners’ failure to pay assessments causes banks to give fewer mortgages to individuals in HOA communities and therefore indirectly hurts the HOA. Second, if the HOA includes language that a purchaser is responsible for paying a portion of the unpaid assessments, this might discourage potential buyers, including the bank, from purchasing or foreclosing the unit. If banks are discouraged from foreclosing on a unit, then this presents the same problem addressed earlier where banks are reluctant to foreclose and the HOA must find another enforcement solution.

B. Rearranging Lien Priority Based on Cooperative Law

Another option for the HOA that shifts the burden of unpaid assessments to the bank is for the HOA to borrow aspects of the law governing cooperatives (“co-ops”). Similar to HOA communities, co-ops are “legal formats for ‘unit ownership’—that is, the ownership of a physically defined portion of a larger parcel of . . . real property.” The owner of a co-op unit is a member of an organization made up of other owners, who all share

---

135 See id. at 585. Banks would be reluctant to give mortgages because if the HOA subordinated the bank’s mortgage to its own lien for unpaid assessments this would “place the economic burden” on the bank, which would prefer not to have the added burden when giving a mortgage to a homeowner. See id.

136 Id.

137 See id. (explaining that language in the HOA’s governing documents giving the bank’s lien priority over the HOA’s lien benefits the community).

138 See id.

139 See supra Part III.A–B.

140 See generally Whitman, supra note 124.

141 Id. at 15.
responsibility for maintaining and managing common areas.\textsuperscript{142}

One of the main differences between co-ops and HOA communities, however, is that in a co-op, a single corporation owns both the units and the common areas.\textsuperscript{143} The corporation owning the co-op provides unit owners with two documents—a proprietary lease\textsuperscript{144} and a stock certificate.\textsuperscript{145} These documents establish the unit owner’s ownership in a share of the corporation and his right of possession of a particular unit. Consequently, instead of owning their individual units, co-op owners own shares in a corporation that leases apartments to its shareholders.\textsuperscript{146}

Since co-op owners own shares in a corporation rather than owning actual real property, laws governing co-ops are vastly different from laws governing HOA communities.\textsuperscript{147} For example, in New York, the state’s Uniform Commercial Code, which relates to commercial transactions, governs co-ops.\textsuperscript{148} The Code “unambiguously gives co-op corporations an automatic first lien on a shareholder’s apartment . . . provided that the bylaws or proprietary lease give the co-op such a lien.”\textsuperscript{149} This differs from laws governing real property, such as ownership of a unit in a

\begin{itemize}
  \item \textsuperscript{142} Id. at 16.
  \item \textsuperscript{143} See id. at 17.
  \item \textsuperscript{144} A proprietary lease is a document authorizing a shareholder’s right to use one of the apartments in a co-op building under the conditions specified in the lease. Jay Romano, \textit{YOUR HOME; Proprietary Co-op Lease: New Model}, N.Y. TIMES, Dec. 17, 2000, http http://www.nytimes.com/2000/12/17/realestate/your-home-proprietary-co-op-lease-new-model.html.
  \item \textsuperscript{147} See id.
  \item \textsuperscript{148} Id.
  \item \textsuperscript{149} Id.
\end{itemize}
Collecting Unpaid Assessments

HOA community. In the real property context, the bank secures the loan it gives to a homeowner to purchase a unit by taking a mortgage on the property, which the bank can foreclose if the homeowner fails to repay his loan. By recording the mortgage, the lender establishes the legal priority of his lien over other liens, such as that held by the HOA for unpaid assessments. Therefore, one critical difference between the corporation’s lien on a unit in a co-op as opposed to the HOA’s lien on a unit in a HOA community is that the corporation’s lien is given priority; if the corporation must recover unpaid fees, it can do so before the bank recovers its mortgage.

Consequently, co-op ownership places a higher burden on the bank, since a co-op corporation’s lien takes priority over the bank’s lien. Lenders, such as banks, are always concerned with the priority of their mortgage loans against the HOAs’ liens for unpaid assessments. Thus, banks loan money to buyers more readily if the bank’s lien is prioritized because the bank is then first in line for repayment if the borrower defaults. Arranging the priority of lien payments so that the bank’s lien is subordinated to the co-op’s lien discourages banks from loaning to individuals in co-ops, and accordingly might not provide an ideal structure for HOAs to emulate.

Without the assurance that banks are paid first, banks hesitate to loan money, and thus potential buyers would struggle to secure mortgages in order to purchase homes in a HOA communities. This ultimately hurts the HOA and the homeowners because,

---

150 See id.
151 See id.
152 Id.
153 Id.
154 Id.
155 Id.
156 Whitman, supra note 124, at 46.
157 See id. at 47.
158 See id. (“[s]ubordination of the association’s lien to the first mortgage . . . is generally thought advantageous to lenders,” and therefore without this advantage, banks are more hesitant to lend).
159 See id.
without the bank willing to lend money, the community faces additional difficulties in selling houses in the community.\footnote{Difficulties in selling houses already exist in poor economic climates due to problems such as unemployment and a general lack of interest in buying homes. Saito, supra note 44.} Property prices for the unsold houses would drop, decreasing the value of the other homes in the area.\footnote{See RESTATEMENT (THIRD) OF PROP.: SERVITUDES § 6.5 cmt. b (2000); TASK FORCE ON RESIDENTIAL MORTGAGE FORECLOSURE CASES, supra note 89, at 16, 26 (explaining that the longer a house remains unoccupied and begins to decline in value, the worse the impact on the community).} Further, with fewer homeowners in the association, the association would likely be unable to collect the assessments necessary to fully operate all of its services. While the bank’s prioritization of its lien over the HOA’s lien is “undesirable from the viewpoint of the association and its other members, since they will usually have to withstand an unusually heavy assessment to make up the delinquency of the foreclosed unit owners,”\footnote{Whitman, supra note 124, at 47.} the HOA might actually fair better with a subordinated lien than if its lien took priority over the bank’s lien. The co-op lien structure therefore might not provide a workable solution for HOAs, since allocating the burden of unpaid assessments to the bank would discourage banks from lending money to would-be-buyers.

Another difference between HOA communities and co-ops is the degree of reliance that exists between members of the community in terms of avoiding foreclosure.\footnote{See id. at 20.} In a co-op, the co-op corporation is typically the mortgager of the blanket mortgage\footnote{A blanket mortgage is a mortgage covering the entire building.} covering the building.\footnote{Whitman, supra note 124, at 18.} Thus, the members of a co-op are more dependent on the other members to successfully finance the building’s mortgage, since all the co-op members together pay off the blanket mortgage.\footnote{Herbert J. Friedman & James K. Herbert, Community Apartments: Condominium or Stock Cooperative?, 50 CALIF. L. REV. 299, 323–24 (1962); see also Whitman, supra note 124, at 20.} If one co-op member defaults, the corporation suffers, as it will be short funds to pay off
Collecting Unpaid Assessments

the mortgage. Additionally, the other members directly suffer since they have to make up the difference. If the other members cannot cover the money owed by the defaulting member, this could trigger foreclosure of the entire co-op.

Contrastingly, in a HOA community, each individual unit owner is his own mortgager on his unit. Therefore HOA communities provide better protection for homeowners against the harms resulting from defaulting homeowners since in HOA communities each individual homeowner is responsible for his own mortgage. Thus, if a homeowner in a HOA community defaults, his default “cannot directly trigger foreclosure . . . of the unit belonging to a non-delinquent owner.”

While HOA communities do provide a layer of protection to non-delinquent homeowners by protecting them from foreclosure against their own units when another member defaults, this benefit dissipates slightly in poor economic climates. First, HOA communities can no longer offer the same protections against foreclosure of non-defaulting individuals’ units since individuals are defaulting at such increasing rates. As more homeowners in HOA communities fail to pay their assessments, the association becomes paralyzed by its lack of funding. Since, for the reasons discussed above, foreclosure is not a viable or helpful option in poor economic climates, HOAs struggle to find ways to replace delinquent homeowners with homeowners able to pay their

167 Friedman & Herbert, supra note 166.
168 See id.; Whitman, supra note 124, at 20.
169 See Friedman & Herbert, supra note 166; Whitman, supra note 124, at 20.
170 See id.
171 See id.
172 Whitman, supra note 124, at 20; see also Friedman & Herbert, supra note 166.
173 Jurow, supra note 42 (explaining that some HOAs face delinquency rates of fifteen percent or more, causing serious revenue shortfalls).
174 See Anderson, HOA Groups in Arizona, supra note 17; Giantomasi, supra note 18, at 2512 (noting that HOA’s cash flow can become severely unstable if homeowners fail to pay assessments).
175 See supra Part III.A–B.
assessments.\(^{176}\) To compensate for the unpaid assessments owed by delinquent homeowners, one of the most feasible options for the HOA is to increase the assessments for homeowners who are able to pay.\(^{177}\) This increase makes it harder for the remaining homeowners to pay the full amount they now owe on their assessments, causing even more homeowners to default.\(^{178}\) In this scenario, defaulting homeowners in a HOA community can cause other homeowners to default on their own mortgages, possibly triggering foreclosure; albeit defaulting members cannot directly trigger foreclosure of non-defaulting members’ houses. Such high default rates on behalf of individual homeowners seriously threaten an association’s solvency.\(^{179}\) Consequently, the advantage that HOA communities possessed over co-ops in terms of preventing foreclosure against non-defaulting individuals becomes more tenuous as the economy suffers and the HOA finds itself subject to the possibility of insolvency.\(^{180}\)

Second, HOAs cannot provide the same level of protection for non-defaulting homeowners against foreclosure in economically challenging times because the HOA has a difficult time evicting delinquent owners and replacing them with owners who can pay their full assessments.\(^{181}\) Without replacing delinquent homeowners, the HOA cannot collect unpaid assessments and the association and the other homeowners suffer and might eventually face insolvency.\(^{182}\) Despite a co-op’s difficulty in protecting non-delinquent members from facing foreclosure if other co-op

\(^{176}\) See Whitman, supra note 124, at 20.

\(^{177}\) See Anderson, HOA Groups in Arizona, supra note 17 (stating that homeowners who diligently pay their fees are forced to face the consequence of increased assessments as a result of delinquent homeowners); Timiraos, supra note 48 (explaining that “with empty pockets, more . . . associations are forced to raise homeowner dues”).

\(^{178}\) See Timiraos, supra note 48 (questioning “[h]ow many of these owners are actually going to be able to pay an additional assessment”).

\(^{179}\) See Whitman, supra note 124, at 20; Anderson, HOA Groups in Arizona, supra note 17.

\(^{180}\) See Friedman & Herbert, supra note 166, at 324; Whitman, supra note 124, at 20.

\(^{181}\) See Whitman, supra note 124, at 20.

\(^{182}\) See id.; Anderson, HOA Groups in Arizona, supra note 17.
Collecting Unpaid Assessments

members fail to pay, a co-op has an easier time quickly evicting delinquent members and thus eliminating this problem. 183 The cooperative can
treat the delinquent member as a tenant 184 and use unlawful detainer or other summary process to evict him . . . . In a [HOA], by contrast, the usual remedy is foreclosure of the association’s lien on the unit, an action which normally enjoys no calendar preference and may take many months to complete.185

Therefore, in poor economies HOAs struggle to provide adequate protection against foreclosure for non-delinquent members since they cannot quickly replace defaulting owners with individuals who can readily pay their assessments.

Despite the difficulties HOAs face in shielding their members from foreclosure based on other member’s inability to pay assessments, HOAs still provide non-defaulting members more protection against foreclosure than co-ops provide their members.186 Unlike the co-op structure, which allows for one member’s default to directly trigger foreclosure of the entire building,187 the HOA structure prevents the default of one homeowner to cause any direct foreclosure proceedings against non-delinquent owners. 188 Adopting the co-op structure therefore might actually endanger non-delinquent HOA members more than it would offer protection. 189

---

183 See Whitman, supra note 124, at 20.
184 For an explanation of reasons why the cooperative can treat the delinquent member as a tenant, see Sun Terrace Manor v. Mun. Court, 108 Cal. Rptr. 307, 307 (Cal. Ct. App. 1973) (considering the approach of Maryland and New York and the opinion of a writer for the California Law Review, which all conclude that “[t]he relationship between the [shareholder-tenant] and the [corporate cooperative] is that of landlord and tenant”).
185 Whitman, supra note 124, at 20–21 n.14.
186 See id. at 20.
187 See id.
188 Id.
189 See id.
Another option for HOAs attempting to collect unpaid assessments is to use security deposits to re-allocate the burden of not paying assessments to the delinquent homeowners. This method ensures that homeowners are incentivized to pay assessments in a timely manner in order to recover their security deposit. It also protects homeowners against paying increased assessments as a result of other members’ defaults, since fewer members default after paying a deposit. The security deposit method allows a HOA to use a defaulting member’s deposit to cure his default, as well as secure payments during the foreclosure process. Finally, homeowners will not lose money by paying a security deposit, as they do by paying increased assessments, but rather will receive interest accruing on their deposit. Therefore, requiring a security deposit from buyers provides a HOA with the greatest security against defaults and insolvency.

A. Background on Security Deposits

A security deposit in the real estate context is most common in a landlord-tenant relationship. It is a deposit of cash or other assets from the tenant given to the landlord for the landlord to use in the event the tenant defaults. Defaults result from failure to
make timely rent payments, failure to perform necessary repairs as instructed under the lease, or failure to vacate the premises at the expiration of the lease. The landlord may adjust the security deposit throughout the term of the lease, in the event the rent increases, in order to continually ensure the deposit covers any unpaid costs. This incentivizes the tenant to pay rent on time and maintain the premises because he will lose his deposit if he fails to do so.

A landlord usually prefers to collect the security deposit from the tenant when the lease is executed and delivered. This ensures that the deposit is made and it provides the landlord with a fund in the event the tenant defaults on his rent at an early date. Further, after the landlord collects the deposit he continuously checks that the deposit level does not drop. For example, if the landlord must use the deposit, whether in part or in total, to cure a default, the landlord requires the tenant to refill the deposit to the original amount by making further payments. "While the requirement that the tenant replenish the security deposit may, in effect, result in trading one default for another, in the absence of such a requirement, the tenant may be deemed to have cured the default by virtue of the application of the security deposit." Accordingly, maintaining a full deposit from each tenant at all times is important for the landlord in order to guarantee funds always exist to cure defaults, and to ensure that once the landlord

197 Id.
198 See id. at 227. Another method of ensuring that the deposit is always capable of covering unpaid assessments is to "set the initial security deposit based on the final escalated rent." Id.
199 See Ralph J. Rohner, Leasing Consumer Goods: The Spotlight Shifts to the Uniform Consumer Leases Act, 35 CONN. L. REV. 647, 719 (2003) ("[T]he economic premise that any front-loaded payments, from a simple security deposit to full payment of the lease obligation . . . reduces the lessor’s risk and thus the consumer’s overall cost for the lease.” The lessor’s risk is reduced because the deposit ensures that the tenant pays.).
200 Saltz, supra note 190, at 228.
201 Id.
202 See id. at 230.
203 See id.
204 Id.
applies the deposit to the deficiency he still possesses a claim against the defaulting tenant.\textsuperscript{205}

Once the landlord receives the security deposit, he becomes responsible for the money.\textsuperscript{206} Several options exist for how the landlord holds the money,\textsuperscript{207} but the tenant prefers that the deposit be held in some form of trust on his behalf and that his deposit not be comingled with the landlord’s other funds.\textsuperscript{208} The landlord should provide information to the tenant regarding how he holds the money, including the rate of interest, if any, that the tenant will receive and the time the interest payments will be made.\textsuperscript{209} When the tenant decides to end the lease, the landlord is required to return the security deposit to the tenant less any money the tenant owes the landlord, unless the tenant vacates or abandons the premises before the lease expires.\textsuperscript{210} In that case, the landlord is not required to return the deposit.\textsuperscript{211} This provides an incentive for the tenant to comply with the terms of the lease in order to recover his full deposit.\textsuperscript{212}

\textbf{B. Benefit to the HOA and its Members of Requiring Security Deposits from Homeowners}

The landlord-tenant relationship and the use of security deposits to establish trust and accountability between the two parties provide a good analogy for homeowners and the HOA. As

\begin{quote}
\textsuperscript{205} See id. at 229–30.
\textsuperscript{206} See generally id. (explaining the landlord’s responsibilities when dealing with the tenant’s deposit).
\textsuperscript{207} The landlord can comingle the security deposit with other funds, isolate the deposit, invest the funds on the tenant’s behalf, or hold the deposit in trust for the benefit of the tenant. See id. at 237–38; R. Wilson Freyermuth, Are Security Deposits “Security Interests”? The Proper Scope of Article 9 and Statutory Interpretation in Consumer Class Actions, 68 MO. L. REV. 71, 81–82 (2003). Further, the tenant might require the landlord to “periodically prove that the funds are intact.” Saltz, supra note 190, at 238.
\textsuperscript{208} See Saltz, supra note 190, at 237–38; Harrison, supra note 194, at 3.
\textsuperscript{209} See Harrison, supra note 194, at 3.
\textsuperscript{210} Id. at 3–4.
\textsuperscript{211} See id.
\textsuperscript{212} See Benjamin, supra note 192, at 186.
\end{quote}
Collecting Unpaid Assessments

demonstrated in the landlord-tenant context, the use of a security deposit provides a strong incentive for tenants to comply with the rules set forth in their lease. 213 Similarly, HOAs looking to find new methods to force compliance in relation to paying assessments could use the security deposit method. 214 First, once the homeowner hands over a security deposit to the HOA he becomes more incentivized to make timely assessment payments. 215 If he fails to make timely payments, the HOA uses the defaulting homeowner’s deposit to cure the deficiency and requires that the homeowner replenish the lost funds. 216 If the homeowner cannot replenish the funds, he recovers less of his security deposit if he decides to move. 217 If, however, the homeowner continues to timely pay his assessments, he receives his total deposit back from the HOA plus any interest that his deposit earned while held in trust on his behalf. 218

Second, by requiring a security deposit, the HOA receives a cushion in case the homeowner fails to pay. 219 The HOA can apply the security deposit to cure the homeowner’s default, protecting the association and other homeowners. 220 Third, if the HOA applies the deposit to cure the delinquent assessment payments and the homeowner subsequently fails to replenish his deposit, this provides notice to the HOA that it might have to consider foreclosure. However, unlike the foreclosure process explained above, 221 if the HOA has a security deposit it can continue to use the funds in the deposit to pay the delinquent assessments throughout the foreclosure proceedings. 222 Even if the funds

213 See Saltz, supra note 190, at 225.
214 See generally id. (noting the considerations of landlords and tenants in determining an appropriate security deposit amount).
215 See supra note 199 and accompanying text.
216 See Saltz, supra note 190, at 229–30.
217 See Harrison, supra note 194, at 3.
218 See Saltz, supra note 190, at 236–37; Harrison, supra note 194, at 3.
219 See Saltz, supra note 190, at 230.
220 See id.
221 See supra Part III.A–B.
222 See Saltz, supra note 190, at 230 (explaining how the landlord can use the funds in the deposit to make up for unpaid rent payments).
eventually run out because of the slow pace of foreclosure, at least for a period during which the foreclosure case is pending the HOA can collect for unpaid assessments. Therefore, the HOA can once again rely on the foreclosure mechanism with the added cushion from the security deposit to protect it in the event of a delayed foreclosure sale.

While utilizing a security deposit provides several benefits for HOAs looking to incentivize homeowners to pay assessments on time, the increased cost the homeowner must incur upon the initial purchase presents an important complication to this landlord-tenant model. In the HOA context, the HOA would have to require a fairly large security deposit to provide enough protection against defaults. This is especially true if the HOA has to initiate foreclosure on a unit since the foreclosure process can take a protracted amount of time. The homeowner would therefore need to obtain a large amount of funding to put towards the security deposit. This requires taking out a larger mortgage, which banks might hesitate to provide, or taking out the same mortgage but scaling down on the purchase price of the house and applying some of the money towards the deposit. Prospective

223 See id. at 227 (explaining that landlords base security deposit payments on the price of the rent). HOAs looking to impose security deposits would similarly base the deposit on the price of the assessments.

224 See Anderson, HOA Groups in Arizona, supra note 17.

225 See generally Gregory C. Chow & Sharon Bernstein Megdal, An Econometric Definition of the Inflation-Unemployment Tradeoff, 68 AM. ECON. REV. 446 (1978) (explaining that tradeoffs require giving up some of one item in order to gain more of another item). While the article discusses tradeoffs in the inflation–unemployment context, the same general principle of tradeoff as an economic theory applies here. The tradeoff situation occurs when a homeowner must pay a large security deposit, assuming a fixed available income, because now the homeowner has less money to spend on purchasing the house. Alternatively, the homeowner can avoid the tradeoff problem by taking out a larger mortgage, since then the homeowner does not have to dedicate part of the original mortgage to the security deposit payment but rather gets a new mortgage that reflects the price of the security deposit. This is assuming the homeowner can obtain a larger mortgage from the bank, which the bank may be reluctant to grant. Louis Uchitelle, Bankers Expected to Stay Hesitant to Lend for Years, N.Y. TIMES, July 5, 1991, http://query.nytimes.com/gst/fullpage.html?res=9D0CE2D91738F936A35754C0A967958260&scp=1&sq=banks%20
Collecting Unpaid Assessments

homeowners are likely to be dissatisfied with these options, since they require borrowing more money or living in a smaller home. However, these disadvantages are offset against the benefits to the homeowner of requiring a security deposit.

While funding the security deposit might present an initial hardship in terms of obtaining finances, homeowners are likely more willing to pay an increased cost up front and be assured that their assessments will not rise at a later date. This is especially true in light of the current predicament homeowners face, where they are required to pay increasing assessments to cover the costs of delinquent homeowners. Homeowners might actually prefer to pay a pre-determined rate up front, rather than an undetermined rate at an unknown later date. Thus, paying an initial deposit actually provides the homeowner with greater security and puts him in a better position to allocate his budget according to costs that he can predict. Further, even though the homeowner pays an upfront deposit, he will not lose this money, unless he fails to pay

227 See E. THOMAS GARMAN & RAYMOND E. FORGUE, PERSONAL FINANCE 272 (9th ed. 2008) (“Insurance is a mechanism for transferring and reducing pure risk through which a large number of individuals share in the financial loss suffered by members of the group as a whole. Insurance protects each individual in the group by replacing an uncertain—and possibly large—financial loss with a certain but comparatively small fee.”); see generally Mila Kofman & Lee Thompson, Consumer Protection and Long-Term Care Insurance: Predictability of Premiums, GEO. U. LONG-TERM CARE FINANCING PROJECT (Mar. 2004), http://ltc.georgetown.edu/pdfs/consumer.pdf (explaining that individuals purchase insurance and pay premiums in order to protect themselves against paying greater, unexpected costs at a later date). Similarly, the security deposit method would ensure predictability by requiring an upfront payment, and then protecting homeowners against unknown, potentially larger costs at a later date resulting from other homeowners’ failure to pay assessments.
228 See, e.g., Anderson, HOA Groups in Arizona, supra note 17; TASK FORCE ON RESIDENTIAL MORTGAGE FORECLOSURE CASES, supra note 89, at 26.
229 See supra note 227 and accompanying text.
230 See id.
his assessments. Assuming the homeowner pays his assessments, he is entitled to receive his entire deposit back when he moves plus any interest his deposit earned. Contrastingly, if the homeowner does not pay the deposit upfront, but instead is subjected to increased assessments later as a result of other homeowners failing to pay their assessments, the homeowner simply loses the money spent on the increased assessments.

Additionally, since homeowners have less of an incentive to pay their assessments on time without a security deposit, a greater probability exists that homeowners will default if the HOA does not use the security deposit method. Even assuming the worst scenario—that the HOA has commenced foreclosure on a unit due to unpaid assessments but has not completed the process before it extinguishes the funds from the security deposit—the HOA and other homeowners should still suffer less of a financial burden than they would without the deposit, since presumably fewer homeowners will find themselves defaulting on assessments after paying a deposit. The HOA therefore should not find itself in the predicament outlined above, where several members of the community fail to pay assessments and threaten the solvency of the association, forcing it to raise assessment costs for other members. Consequently, the homeowner is better off paying the security deposit upfront, despite the initial difficulty obtaining the funding, because it will help decrease the chances of assuming the financial responsibility of the delinquent homeowners at a later date.

VI. CONCLUSION

Considering the options available to a HOA to strengthen its ability to collect assessments, the security deposit method provides the greatest incentive for homeowners to pay assessments on time.

---

$^{231}$ See Harrison, supra note 194, at 4.
$^{232}$ See Saltz, supra note 190, at 236–37.
$^{233}$ See Benjamin, supra note 192, at 186.
$^{234}$ See supra Part III.A–B.
$^{235}$ See supra note 227 and accompanying text.
Collecting Unpaid Assessments

and to ensure that the burden of unpaid assessments falls to the defaulting homeowner. The delinquent homeowners represent the ideal group to bear the burden of unpaid assessments, since they created the burden initially. Paying a security deposit also encourages homeowners to quickly cure any defaults. Thus, HOAs should consider requiring a security deposit from all homeowners upon purchase of a unit in HOA communities. With money available to cure defaults and protect the HOA and its members, Mr. Furman and other similarly situated homeowners can once again enjoy the benefits of living in a HOA community.
NEW YORK AND DIVORCE: FINDING FAULT IN A NO FAULT SYSTEM

Lauren Guidice*

INTRODUCTION

As the millions of children of “fault” divorces in New York can attest, the damage involved in declaring fault in order to dissolve a marriage is palpable and enduring. Many are not old enough to understand why the intricate process may take years to complete, or why their mothers assert “cruel and inhuman” treatment when their fathers never seemed like the monster that phrase made him out to be. For over two decades, New York remained stuck in an archaic fault-based system1—requiring couples to point fingers and air dirty laundry in order to divorce—while the rest of the country progressed and developed unilateral divorce systems.2

The fault requirement of the dissolution of marriage in New York has long been considered to be unnecessary and damaging.3

* J.D. Candidate, Brooklyn Law School 2012; B.A., Fordham University 2008. I would like to thank my family and friends for their constant support, particularly my parents, who taught me that divorce and family are not incompatible. I would also like to thank my editors for their insights throughout the writing and editing of this Note.

1 See The Editors, Is New York Ready for No-Fault Divorce?, N.Y. TIMES ROOM FOR DEBATE BLOG (June 15, 2010), http://roomfordebate.blogs.nytimes.com/2010/06/15/is-new-york-ready-for-no-fault-divorce/ (stating that “New York was the longtime holdout [on no-fault divorce], since South Dakota passed its law in 1985”).

2 South Dakota was the second-to-last state to pass a no-fault divorce law in 1985. SL 1985, ch. 207 § 1; See also, The Editors, supra note 1.

The looming threat of a fault trial can harm the children of divorce by dragging out proceedings and airing allegations of marital fault, many of which are untrue. Additionally, scholars and practitioners criticize the fault requirement for adding unnecessary expense to the already financially taxing experience of divorce. Though fault trials are rare, they can last from two days to several weeks. As supervising judge of the matrimonial division in Nassau County, Judge Robert A. Ross, said, “[fault trials] are never pleasant.”

On August 15, 2010, in response to decades of proposals and complaints, Governor Patterson signed a three-bill package that overhauled New York divorce laws. The package places New York’s divorce laws on equal footing with that of the forty-nine other states that do not require fault in order for a couple to dissolve their marriage. The reform comes twenty-five years after South Dakota, the most recent state to pass a no-fault divorce law.


4 Id. “The Commission recognized the need to change the very culture of the system and to make explicit recommendations to reduce trauma, cost and delay.” Id.

5 Id.


7 Id.

8 Id. (quoting Nassau County Judge Robert A. Ross, who stated, in regards to the “horrible” nature of fault trials: “You may have friends called, girlfriends and boyfriends called, people who are alleged to be girlfriends and boyfriends . . . having to listen to these things can sometimes be an overwhelming experience”).

9 N.Y. DOM. REL. LAW § 170(7) (McKinney 2010); See also Denise M. Champagne, Devil’s in the Details of No-Fault Divorce in New York, DAILY RECORD, July 14, 2010.

10 Champagne, supra note 9.
law, did so in 1985. The bill encountered significant setbacks during its passage through the State Legislature and prior to gaining the endorsement of Governor Patterson. As Governor Patterson noted at the signing of the three-bill package, the reform brings New York divorce laws “into the twenty-first century.”

Prior to the passage of this law, divorce in New York was a long and contentious process that could not be achieved unilaterally. The closest that parties could get to a “no-fault” option was to mutually agree to a divorce, submit to a “trial” separation, and then convert that separation into a divorce after one year of living apart. If a couple could not mutually agree to a divorce, then they would have to seek a divorce on fault-based grounds. Such “fault” options included adultery, cruel and inhuman treatment, abandonment, and constructive

---

11 Cynthia Lee Starnes, Mothers as Suckers: Pity, Partnership, and Divorce Discourse, 90 IOWA L. REV. 1513, 1538 (2005) (“In 1985, South Dakota became the last state to adopt a no-fault divorce ground.”).

12 Id. (noting that the current bill’s passage through the Senate in June 2010 marked the best chance “in more than two decades” of passing a no-fault divorce provision).


15 Section 170(5) of New York’s Domestic Relations Law requires a couple to live apart pursuant to a separation judgment or decree for a period of one or more years and show “satisfactory proof” that they have substantially performed the terms and conditions of the judgment or decree before they will be granted a divorce. N.Y. DOM. REL. § 170(5) (McKinney 2010). Similarly, section 170(6) requires a husband and wife to live apart for a period of at least one year pursuant to a written agreement of separation before they will be issued a divorce. Id. See also Joel Stashenko, No Fault Companion Bill on Maintenance Raises New Concerns, 244 N.Y. L.J. 1 (2010).

16 N.Y. DOM. REL. §§ 170(5)–(6) (McKinney 2010).

17 Id.
abandonment. This route typically involved extended and costly litigation, at the end of which a divorce was not even guaranteed. At other times, couples and their attorneys would resort to outright perjury in order to establish the requisite “fault” necessary for a divorce.

Under the new divorce law, divorce can be granted to an unhappy spouse who declares under oath that his or her marriage has been “irretrievably” broken for at least six months. Other issues related to the divorce, such as child custody and distribution of property, must be resolved by the parties or determined by the court before a divorce will officially be granted.

The State Legislature also passed two divorce finance reforms that will affect post-divorce maintenance awards. The amendments to sections 237(a), 237(b), and 238 provide that the spouse with the greater economic worth will pay for attorney’s fees, expert fees, and any additional costs, disbursements, or

---

18 Id.
20 Champagne, *supra* note 9 (“[I]ndividuals have been known to fabricate allegations in order to establish grounds”); In 1934, the New York Mirror created a sensation with its series, “I Was the Unknown Blonde in 100 New York Divorces!” (discussing the way in which New York lawyers worked around the adultery-only divorce grounds prior to divorce reform in 1960). J. Herbie DiFonzo & Ruth C. Stern, *Addicted to Fault: Why Divorce Reform Has Lagged in New York*, 27 Pace L. Rev. 559, 571–72 (2007).
21 2010 N.Y. Sess. Laws, ch. 384 (McKinney 2010), adding section 170(7) to New York Domestic Relations Law which reads as follows:

The relationship between husband and wife has broken down irretrievably for a period of at least six months, provided that one party has so stated under oath. No judgment of divorce shall be granted under this subdivision unless and until the economic issues of equitable distribution of marital property, the payment or waiver of spousal support, the payment of child support, the payment of counsel and expert’s fees and expenses as well as the custody and visitation with the infant children of the marriage have been resolved by the parties, or determined by the Court and incorporated into the judgment of divorce.

22 Id.
23 N.Y. Dom. Rel. Law §§ 237(a)–(b), 238 (McKinney 2010).
New York and Divorce

expenses. Further, the amendments to Domestic Relations Law section 236 will govern the provision of temporary maintenance in matrimonial actions. The new amendments took effect on October 12, 2010.

The passage of a no-fault cause of action for divorce and a new maintenance formula raise the specter of a judicial process in which fault does not play a role. However, couples divorce for independent and varied reasons—no two divorces are alike. If divorce law is to “do justice,” fault should matter in the division of assets, even where it is properly excluded from the reasons for divorce. By allowing fault to be taken into account when determining maintenance awards, New York courts would maintain the requisite authority to provide equitable post-divorce settlements regardless of which party desired the divorce.

Further, divorce cannot be granted under the new law until issues such as equitable distribution, spousal maintenance, child support, counsel and expert fees, child custody, and visitation are either resolved by the parties or determined by the court. Thus, in order to protect themselves from the “one-size fits all” divorce that can occur in no-fault jurisdictions, New Yorkers should enter into premarital agreements.

24 Id.
25 “Temporary maintenance” is the term used for the monetary support provided during the interim of a matrimonial action. N.Y. Dom. Rel. Law § 236 (McKinney 2010).
26 2010 N.Y. Sess. Laws, ch. 384 (McKinney 2010). In order to study the effects of the new temporary maintenance formula, a Commission was set up and will submit its findings to the state legislature sometime in mid-2011. Stashenko, supra note 15 (“The Law Revision Commission will study the setting of maintenance levels statewide to determine if courts are failing to take into account financial factors that could unfairly disadvantage one spouse or another.”).
28 Id.
29 N.Y. Dom. Rel. § 170(7).
30 Or, post-marital agreements, which are entered into by couples that are already married and wish to negotiate the terms of their divorce. 41 Am. Jur. 2d
become more and more streamlined, premarital agreements offer New Yorkers the most equitable solution to an increasingly litigation-less area of the law.

Part I of this Note will explore the historical development of divorce law by examining the traditional fault based divorce law and the subsequent movement away from the law in other jurisdictions. Part II will discuss the fault-based system under New York’s former Domestic Relations Law and address the inadequacies of the fault based divorce system. Part III will examine the subsequent changes that will occur as a result of New York’s divorce reform bill, including no-fault divorce and divorce finance changes. This section will explore the implications of no-fault divorce laws on maintenance, including how the new formula provides for calculating temporary maintenance awards under New York’s new no-fault divorce reform package, and how that will affect the parties to divorce. Additionally, this section will address the possible financial effect New York’s divorce laws will have on couples seeking to divorce. Part IV will address potential inadequacies in the new divorce law, and recommend both judicial prescriptions and precautionary measures that should be taken by spouses prior to entering into marriage. Ultimately, this Note maintains that faults should play a role in maintenance proceedings and the distribution of marital property. Following the example of many other states, New York should allow marital misconduct to be taken into account in determining maintenance awards and the division of property.

I. HISTORICAL DEVELOPMENT OF DIVORCE LAW

Divorce law developed rapidly throughout the late nineteenth and early twentieth centuries. Viewed in much the same light as

Husband and Wife § 107.

31 N.Y. DOM. REL. LAW § 236 (b)(1) (McKinney 2010).

32 N.Y. DOM. REL. LAW § 236 (McKinney 2010) (Amendments to section 236 of New York Domestic Relations Law were enacted, governing the provision of temporary maintenance in matrimonial actions).

New York and Divorce

tort actions, divorce was seen as a remedy for those spouses who had been wronged by their partner.\textsuperscript{34} Thus, the fault-based divorce system was the dominant scheme in most jurisdictions until the mid-twentieth century.\textsuperscript{35} However, as social values and mores changed, strict fault-based divorce regimes were replaced by unilateral, no-fault systems.\textsuperscript{36} By 1985, almost every jurisdiction in the United States had a unilateral divorce system in place. The singular exception to this overwhelming support for unilateral divorce was New York.

A. Traditional Fault Based Divorce

Divorce was not recognized under English common law until the mid-nineteenth century, when divorce jurisdiction was removed from the ecclesiastical courts to the civil court system.\textsuperscript{37} At that time, divorces were only granted in cases of adultery.\textsuperscript{38} By the late nineteenth century, the laws regulating marriage and divorce in the United States varied drastically among the states.\textsuperscript{39} Though the laws differed, each jurisdiction recognized divorce or judicial separation on limited grounds—all of which involved some degree of fault.\textsuperscript{40} Although the most frequently used grounds for divorce were “[a]dultery, extreme cruelty, or desertion,” many states permitted “insanity, conviction of a crime, habitual drunkenness and drug addiction” to be used as grounds for divorce.\textsuperscript{41}

\textsuperscript{35} JOHN DE WITT GREGORY ET AL., UNDERSTANDING FAMILY LAW 237 (3d ed. 2005).
\textsuperscript{37} See LYNN CAROL HALEM, DIVORCE REFORM: CHANGING LEGAL AND SOCIAL PERSPECTIVES 12–17 (1980).
\textsuperscript{38} Evans, supra note 34, at 472.
\textsuperscript{39} BLAKE, supra note 36, 116–29.
\textsuperscript{40} Id. (noting the differences in approaches to divorce that were adopted).
\textsuperscript{41} GREGORY ET AL., supra note 35, at 237.
Under the fault-based system, the primary concern was “the strong public interest in preserving marriage.”\textsuperscript{42} The system was founded on the belief that requiring proof of fault would limit access to divorce and further the state’s interest in preserving marriages.\textsuperscript{43} Divorce was “conceived as a remedy for the innocent against the guilty.”\textsuperscript{44} This strict notion of “wrong-versus-right” led many courts to refuse to grant divorce in cases where both parties were guilty of marital misconduct.\textsuperscript{45}

The limited availability of divorce began to wane in the second half of the twentieth century as jurisdictions across the country began to authorize divorce without regard to fault.\textsuperscript{46}

Nevertheless, New York’s general divorce law, which solely allowed for divorce on the ground of adultery, remained “immune to revision.”\textsuperscript{47} Attempts to reform New York laws by broadening the grounds for divorce were continuously defeated.\textsuperscript{48} However, by 1930 the New York Legislature had successfully defined five grounds for annulment: infancy,\textsuperscript{49} bigamy, lunacy and idiocy, force or fraud, and physical incapacity (impotence).\textsuperscript{50}

\textbf{B. The Emergence of No-Fault Divorce}

The limited availability of divorce left many couples trapped in

\textsuperscript{43} Evans, \textit{supra} note 34, at 473.
\textsuperscript{44} Brewies v. Brewies, 178 S.W.2d 84, 85 (Tenn. Ct. App. 1944) (holding that a divorce ordinarily will not be granted where both parties are equally at fault).
\textsuperscript{45} The result that occurs is ironic: when both parties have a ground for a divorce, neither has a right to divorce. \textit{Id. See generally} J.W. Bunkley, \textit{The Doctrine of Recrimination in Divorce Law}, 20 \textit{Miss. L.J.} 327 (1948). This was known as the doctrine of recrimination. \textit{Id.}
\textsuperscript{46} \textit{Gregory et al.}, \textit{supra} note 35, at 237 (discussing the collapse of conventional divorce law and the development of no-fault divorce laws).
\textsuperscript{47} DiFonzo & Stern, \textit{supra} note 20, at 564.
\textsuperscript{48} \textit{Blake}, \textit{supra} note 36, at 201–04. Between 1900 and 1933, over fifteen bills were sponsored and rejected. \textit{Id.} at 201.
\textsuperscript{49} Being under the age of consent.
\textsuperscript{50} \textit{Blake}, \textit{supra} note 36, at 66–67.
irretrievably broken marriages. Over time, the fault-based divorce system led to the widespread corruption of the family law system, which legislatures and judges across the nation had come to ignore. Unhappy couples resorted to collusion—fabricating evidence of marital misconduct in order to establish the requisite ground upon which divorce could be granted. In a notorious example, Dorothy Jarvis was hired in over one-hundred divorce cases to play the role of “the other woman.” Jarvis staged adulterous hotel rendezvous with New York husbands. A photographer would catch the “adulterous couple,” thus providing the necessary evidence for the unhappily married New Yorkers to obtain a divorce.

Eventually, the widespread practice of “collusion and perjury” by couples wishing to divorce challenged the legitimacy of the family law system. In response, California enacted the country’s first no-fault divorce law. Signed by Governor Ronald Reagan in 1969, this revolutionary legislation inspired similar
changes across the country as the no-fault approach grew in popularity.61

By the mid-1980s, all states had some form of a no-fault provision integrated into their divorce law.62 Most states enacted no-fault divorce laws that cited “irreconcilable differences” or “irretrievable breakdown” as the basis for marital dissolution.63 Neutral grounds for dissolving marriage reduced the moral stigma associated with divorce.64 In addition, a large majority of no-fault regimes began to permit unilateral divorce.65

Social scientists studying family relationships in the middle of the twentieth century suggested that “marriages broke up in [the] context of conflicts in attitude, personality, or other difficulty on both sides, rather than as a result of fault by one spouse and innocence by the other.”66 Advocates of no-fault divorce relied on this proposition to bolster the argument that traditional grounds for divorce (i.e., adultery, desertion, etc.) were symptoms of a deteriorating marriage rather than the causes.67 It was argued that

---

61 GREGORY ET AL., supra note 35, at 237 (describing California’s divorce reform law as the forefront of the “divorce revolution”).

62 Either as the sole basis for divorce or as an alternative to the traditional fault based system. See Evans, supra note 34, at 474. See also Scheu v. Vargas, 778 N.Y.S.2d 663, 663 (Sup. Ct. 2004) (stating that, in New York, no-fault divorce applies only where there is a previous decree of separation or a written separation agreement, as required by the Domestic Relations Law provision listing grounds for divorce; otherwise, a divorce may be granted only if fault is established pursuant to one or more of the grounds set forth in the provision).

63 Evans, supra note 34, at 474.

64 Id.

65 This meant that one spouse could terminate the marriage without the consent of the other. Michael Grossberg, How to Give the Present a Past? Family Law in the United States 1950-2000, in CROSS CURRENTS, supra note 58, at 3, 7.

66 Id. at 18.

67 Id. at 17–18.
parties should not be required to assert such symptoms in order to finalize a divorce when the causes of the breakdown were often multifarious and equally distributed between the parties.68

In the wake of California’s “legal and cultural transformation,”69 no-fault divorce reform quickly became a prevalent topic in family law discourse throughout the country.70 In 1969, the same year as California’s groundbreaking divorce reform, the Federal Uniform Marriage and Divorce Act (“UMDA”) was approved,71 which called for irretrievable breakdown of the marriage to be the sole ground of divorce.72 Federal encouragement of no-fault divorce practices along with the independent adoption of no-fault divorce regimes among the states highlights the (almost) universal belief that the fault requirement to the dissolution of marriage was outdated and unnecessary. However, despite the fact that every jurisdiction in the country would adopt unilateral divorce systems throughout the 1970s and 1980s, it would take New York nearly a half-century to follow suit.73

C. Why Change Took So Long In New York

In the years leading up to the present overhaul of its divorce laws, New York remained unable to advance past legislative wrangling into legislative action.74 New York’s inaction can be attributed in part to the influence of the Catholic Church and to the

68 Id.
70 See ASSEMB. COMM. ON JUDICIARY, supra note 60, at 132 (1997) (describing California’s passage of a no-fault option as the launch of a “legal revolution”).
71 UNIF. MARRIAGE & DIVORCE ACT § 302 (1970). Although only eight states have adopted the UMDA—Arizona, Colorado, Illinois, Kentucky, Minnesota, Missouri, Montana, and Washington, “the act has greatly influenced the terms of divorce reform for many states.” DiFonzo & Stern, supra note 20, at n.223.
73 THE EDITORS, supra note 1.
74 DiFonzo & Stern, supra note 20, at 567.
many dedicated women’s organizations in the legislature. Such forces underlie the main reason for New York’s delayed progress in divorce reform, which was simply that it was decades behind the rest of the country.

New York deferred divorce reform legislation several times between 1900 and 1960. Scholars have suggested that the “stubborn adhesion” of New York’s lawmakers to a single ground for divorce reflected “not so much a stern sense of duty as an inability to give the problem of marital law more than fitful attention.” Distracted by social and religious forces, decades of World Wars, and economic depression, New York legislators stood stagnant on the issue of divorce reform.

By the 1950s, New York had the “lowest recorded divorce rate in the country.” However, the frequency of annulments, migratory divorce, separations, and desertions combined to increase the state’s total marital disruption far beyond the national average. In the 1960s, “this dichotomy between ‘law-as-statute

---

75 See generally id.
76 See generally id.
77 BLAKE, supra note 36, at 64–79.
78 Id. at 64.
79 DiFonzo & Stern, supra note 20, at 567. Between the years 1900 and 1933, fifteen different legislators sponsored bills to modernize New York’s divorce law by adding grounds such as cruelty and desertion. BLAKE, supra note 36, at 201. However, each bill was “buried in committee.” Id.
80 DiFonzo & Stern, supra note 20, at 576.
81 New York had the nation’s highest annulment rate, comprising approximately one-third of all annulments in the United States. HERBERT JACOB, SILENT REVOLUTION; THE TRANSFORMATION OF DIVORCE LAW IN THE UNITED STATES 35 (1988).
82 The practice of leaving one’s state or country in order to take advantage of more lenient divorce laws and secure a divorce. BLAKE, supra note 36, at 1–4, 152–59.
83 Unhappy New Yorkers were willing, in significant numbers, to leave home to get divorced. DiFonzo & Stern, supra note 20, at 576. By 1922, scholars noted that nearly one-third of all New York divorces had been obtained out of state. Id. Because of Constitutional full-faith and credit requirements, New York Courts accepted these out of state divorces so long as both parties were present. Id. at 573. And, for those New Yorkers who could not afford to leave the state in order to end their marriages, annulments provided another
New York and Divorce

and law-in-action” eventually proved to be too much for the New York Legislature. New York’s first major divorce reform occurred as a result of a 1961 report by the Joint Legislative Committee on Matrimonial and Family Laws. Known as the “Wilson Committee,” the group discovered what Family Law judges and practitioners had long known; the State’s formal divorce law had “spawned a set of practices inconsistent with any notions of sound public policy.” The Committee realized that by framing the need for divorce reform in such a way as to stress the “public policy need for consistency between the law in practice and the law as written,” it could successfully evade the moral debate of divorce itself.

The subsequent divorce reforms—New York’s first—occurred in 1966. The new law provided for divorce on the grounds of adultery, cruel and inhuman treatment, abandonment for two or more years, confinement in prison for three or more years, or living apart for a period of two years or more pursuant to an

route to dissolving their unions. In New York, annulments accounted for 25% of marital dissolutions during World War II and nearly 50% after 1950. Id. at 574. In addition to traveling in order to get a divorce, or resorting to annulments, other couples who wanted to dissolve their marriages simply resorted to deserting their partners. Id. at 575. In 1940, the percentage of white married women (excluding widows) who did not live with their husbands was approximately 30% greater in New York than the rest of the country. Id.  

84 DiFonzo & Stern, supra note 20, at 576.  
85 Id.  
86 JACOB, supra note 81, at 37.  
87 Id. at 35–37.  
88 DiFonzo & Stern, supra note 20, at 578.  
89 Again and again, the Committee was made aware of how ‘the formal law had spawned a set of practices inconsistent with any notions of sound public policy.’ Viewed in this way, divorce became something other than grist for the conservative-liberal debate. Because the issue was framed as ‘disjunction between law-on-the-books and law in action,’ divorce reform could be presented as a much-needed ‘procedural’ change.  
Id. See also, BLAKE, supra note 36, at 212.  
90 Id. at 577.
agreement or a judicial separation decree. Although limited in reach, New York had passed its first no-fault divorce law.

Further, when the New York Court of Appeals upheld the legality and retroactivity of conversion divorce in Gleason v. Gleason in 1970, it became clear that New York was moving in a positive direction insofar as its divorce laws were concerned. Notably, the Court’s reasoning explicitly rejected any state interest in compelling couples to remain together in marriages that were clearly non-operational. The Court concluded that the purpose of the “no-fault provision” allowing for conversion divorce was to remove issues of misconduct from the Court’s consideration. Arguably, the Gleason Court extended support to no-fault divorce based on “moral and social grounds.” No-fault divorce in New York was on its way.

II. NEW YORK’S FAULT BASED DIVORCE SYSTEM

New York’s fault-based divorce system was rife with issues that often led to bitter and divisive battles between couples wishing to end their marriages. Often, a couple’s reasons for divorcing did not neatly fall into one of the statute’s specific grounds for divorce. Additionally, judicially prescribed conditions for each ground of divorce often made the process more complex and

---

91 HALEM, supra note 37, at 258–59. Judicial separation decrees require a showing of marital fault. Id. Alternatively, parties may separate pursuant to a written agreement, filed with the Clerk of the Court, without a finding of fault. Id. In judicial separations, either the innocent party or the party against whom a judgment has been made may apply for conversion. Id.

92 See Gleason v. Gleason, 256 N.E.2d 513, 516 (N.Y. 1970) (holding that the action for a conversion divorce could be maintained even by the person who was found at fault in the original separation action). “Conversion divorce” occurs when two parties separate pursuant to a separation agreement or decree by the court. The terms of the agreement or decree must be abided by and the parties must live separately for a designated period of time—at which time the separation may be “converted” into a divorce by the Supreme Court. See id.

93 See id.

94 Id.

95 HALEM, supra note 37, at 266.

96 See Miller, supra note 3, at 551.
New York and Divorce resulted in uncertain outcomes. Further, while the financial stability of the divorcing parties was often a motivating factor behind courts’ refusals to grant divorce petitions, such considerations frequently resulted in inequitable decisions. Ultimately, New York’s defunct divorce finance laws were a major catalyst for the current overhaul of the state’s domestic relations laws.

A. Marital Fault

The fault-based divorce system subsisted in New York until the amendment of subsection seven to Domestic Relations Law section 170 in October of 2010. Under the prior New York Domestic Relations Law section 170, an action for divorce could only be granted if the couple could establish one of the six grounds for divorce established in 1966.97 While only 23 percent of divorces in 2009 were contested in New York, there was still a possibility of a contested divorce and the ensuing fault trial, with long, drawn out witness testimonies on the alleged wrongdoing of the parties and painful attacks on each party’s credibility.98

The following cases illustrate the difficulty of requiring fault in order to divorce and demonstrate how courts often refused to grant divorces. While it appeared that New York courts strictly enforced fault criteria, the underlying financial concerns at play in divorce proceedings were often the central consideration in courts’

97 Cruel and inhuman treatment, abandonment (also constructive abandonment), imprisonment of one of the parties for a period of three or more years, adultery, and living separate and apart pursuant to a separation judgment or agreement. N.Y. DOM. REL. § 170 (McKinney 2010). According to the New York State Department of Health, the most frequently cited grounds for divorce in 2008 (the most recent year for which statistics are available) were abandonment and cruel and inhuman treatment, respectively. DEP’T. OF HEALTH, INFO. FOR A HEALTHY N.Y.: TABLE 50: DIVORCES BY COUNTY OF DECREE AND LEGAL GROUNDS FOR NEW YORK STATE - 2008, available at http://www.health.state.ny.us/nysdoh/vital_statistics/2008/table50.htm. Out of 52,619 divorces filed for in 2008, abandonment accounted for over 36,000 of the grounds with cruelty a distant second at just over 10,000. Id.

98 Kolker & Hurtado, supra note 6.
reasoning. In *Hessen v. Hessen*—the first case to highlight the various considerations evaluated by courts when determining whether conduct constitutes cruel and inhuman treatment sufficient to grant a divorce—the court held that physical or mental abuse warrants a divorce only if the conduct makes cohabitation improper or unsafe. In *Hessen*, the Court of Appeals noted that the “cruel and inhuman treatment” ground for divorce could not be satisfied by the mere breakdown of the marital relationship, but would require evidentiary proof of cruelty. What is particularly noteworthy in *Hessen* is the Court’s focus on the respective ages of the husband and wife seeking divorce, and on the duration of the marriage. The Court denied the Hessens a divorce, stating that the appearance of misconduct, “which in a matured marriage might fail to justify a finding of substantial misconduct, may justify or even compel an inference of substantial misconduct in a newer marriage.” The case law that follows *Hessen* further demonstrates the absurdity of New York’s divorce laws.

The strict evidentiary standard required to end a marriage of long duration is illustrated by *Palin v. Palin*. Although the wife in *Palin* proved that she had been verbally abused, threatened, and physically attacked by her husband, the court held that this abuse was insufficient for a finding of cruel and inhuman treatment. The court noted that the treatment did not suggest anything more than “unpleasantness.” In fact, New York courts have held that

---

100 *Id.* (emphasis added).
101 *Id.*
102 *Id.* at 895 (stating that for marriages of long duration it is proper to apply the admonition “for better or worse”).
103 *Id.*
104 *Id.* at 893–95. The Court took into account that an older couple may experience the “deleterious effects of old age on the physical and mental disposition,” which may create problems within a marriage, and that changes in family situations like the departures of grown children from the household or family tragedy, may create difficulties. *Id.*
106 *Id.*
107 *Id.* (granting the wife a divorce based on adultery but rejecting the divorce on grounds of cruel and inhuman treatment).
isolated acts of violence are not sufficient for a divorce based on cruel and inhuman treatment.\textsuperscript{108}

In a well-publicized case that occurred in 2010,\textsuperscript{109} a Nassau County court denied a husband’s divorce request on the grounds of cruel and inhuman treatment despite the fact that his wife had threatened him with a samurai sword.\textsuperscript{110} Evidence during the lengthy fault trial showed that the wife had abused the husband for years.\textsuperscript{111} The constant abuse resulted in the husband’s sleep deprivation because he feared “sneak attacks” from his wife.\textsuperscript{112} The wife once brought the sharp tip of a samurai sword “within inches” of her husband’s face; this extreme event precipitated the husband’s request for a divorce.\textsuperscript{113} However, noting that the marriage was one of long duration, the court denied the husband a divorce.\textsuperscript{114} The court’s opinion also noted that the husband failed to testify that the alleged incidents “so endangered his physical or mental well-being as to render it unsafe or improper for him to cohabit with his wife.”\textsuperscript{115}

\textsuperscript{108} See, e.g., Wenderlich v. Wenderlich, 311 N.Y.S.2d 797 (N.Y. App. Div. 1970) (finding “proof . . . that defendant struck the plaintiff the morning of September 1, 1967 . . . insufficient to establish the cause of action”); Concetto v. Concetto, 377 N.Y.S.2d 164 (N.Y. App. Div. 1975) (holding that “[t]he proof, however, failed to establish that the name-calling and the two isolated acts of alleged violence, to which the husband testified, so endangered his physical or mental well-being as to render it unsafe or improper for him to cohabit with his wife”); Rabinowitz v. Rabinowitz, 321 N.Y.S.2d 934 (N.Y. App. Div. 1971) (explaining that “[e]vidence showing that marriage was marked by lack of harmony, frequent quarrels and occasional strife, all adding up to degree of incompatibility, still fell far short of statutory requirements”).


\textsuperscript{111} Id. at *5–6.

\textsuperscript{112} Id. at *3.

\textsuperscript{113} Id. at *5. During the lengthy “fault” trial, the couple’s daughter testified that the tip of the sword was “extremely sharp” and, had she not intervened, her mother could have seriously injured or killed her father. Id.

\textsuperscript{114} Id. at *9.

\textsuperscript{115} Id. (citing Hessen v. Hessen, 308 N.E.2d 891 (N.Y. 1974)).
These cases illustrate the difficulty of proving cruel and inhuman treatment as grounds for a divorce in New York. Although the average divorce case lacks the drama of a samurai sword, many spouses suffered physical or emotional abuse that did not rise to the court-defined level of cruel and inhuman treatment under New York law.\textsuperscript{116} Asserting other grounds for divorce was often just as complicated and ineffective.

New York courts made equally outrageous holdings under the awning of constructive abandonment grounds for divorce. In order for a spouse to assert constructive abandonment,\textsuperscript{117} the refusal to engage in marital relations must have persisted for at least one year prior to the commencement of the action.\textsuperscript{118} However, in \textit{Hammer v. Hammer}, the Court asserted that the failure to file for a divorce on the grounds of constructive abandonment after an extended period of time could bar a spouse from claiming constructive

\textsuperscript{116} See, e.g., Gross v. Gross, 836 N.Y.S.2d 166, 168–69 (N.Y. App. Ct. 2007) (holding that reprehensible and highly offensive behavior does not necessarily establish cruel and inhuman treatment ground for divorce where wife testified that her husband had forced himself on her sexually, trapped her inside the marital bedroom on occasion, and threw her against the walls of their home); E.D. v. M.D., 801 N.Y.S.2d 233, 234 (N.Y. Sup. Ct. 2005) (holding that the evidence of physical contact between the parties was isolated and minimal and not grounds for cruel and inhuman treatment where husband was frequently argumentative and struck the wife on occasion but where the wife had not sought medical treatment for any injuries); S.C. v. A.C., 798 N.Y.S.2d 348, 352 (N.Y. Sup. Ct. 2004) (holding that, with regard to the incidents of physical contact between the parties, none were more than minor and incidental to the behavior complained of and were insufficient for grounds of cruel and inhuman treatment).

\textsuperscript{117} Constructive abandonment is one of the most popular grounds for divorce. See Dep’t of Health, \textit{supra} note 97 (listing the number of divorces in New York in 2008). A refusal or failure to engage in marital relations, to rise to the level of constructive abandonment, must be unjustified, willful, and continued, despite repeated requests from the other spouse for resumption of cohabitation. See Silver v. Silver, 677 N.Y.S.2d 593, 594 (N.Y. App. Ct. 1998).

New York and Divorce

abandonment altogether. The Hammer Court held that where a husband had not threatened to terminate his sexless marriage, but allowed it to continue for ten years, he impliedly consented to the status of the relationship and could not claim constructive abandonment as grounds for divorce.

Although Hammer involved a case where the couple had not engaged in intimate relations for ten years, New York courts have rejected constructive abandonment claims involving much shorter periods. In Breckinridge, the couple mutually declared that their marriage had been sexless for three years, and each asserted that the lack of marital relations was due to the other’s lack of interest. The court held that because both parties implicitly agreed to eliminate marital intimacy, neither could be at fault and no divorce could be granted.

By ignoring any multitude of reasonable causes for a couple to remain together despite their lack of marital relations, New York courts arguably encouraged couples to surrender their marriages early. Couples with legitimate reasons for staying together such as raising children or financial inability to separate could lose the ability to divorce on the grounds of constructive abandonment at a later date. These cases are merely a small sample of New York divorce cases. The inconsistent, unpredictable nature of the decisions clearly illustrates the necessity of a unilateral and “faultless” ground for divorce.

B. Alimony

The effect that New York’s alimony laws had on individuals in

---

120 Id; see Bunce v. Bunce, 426 N.Y.S.2d 105 (N.Y. App. Div. 1980) (constructive abandonment claim based on refusal to engage in marital relations for fifteen years).
122 Id. at 137.
123 Id.
124 Id. at 138.
the wake of divorce—particularly women—was a significant factor behind the Court of Appeals’ holding in *Hessen* and other cases like it.\textsuperscript{125} The old Domestic Relations Law deprived alimony and occupancy of the marital home to a spouse against whom a divorce had been granted.\textsuperscript{126} With this in mind, New York courts decided that special thought should be given to the deprivation of support rights that would follow a fault determination against dependent spouses.\textsuperscript{127} For example, because the husband in *Hessen* was the one who filed for divorce and the one who asserted cruelty against his wife, the Court felt that it would be inappropriate to deprive her—a dependent older woman—of support by granting the husband a divorce absent truly grievous misconduct.\textsuperscript{128} The consideration for the financial welfare of spouses post-divorce sheds a more reasonable light on the Court’s holding. In fact, the concern for the financial effect of divorce on spouses was common for most of the grounds for divorce in New York, as the judiciary did not want individuals to become wards of the state.\textsuperscript{129}

Over time, it became clear that dependent wives seeking divorce on the grounds of cruel and inhuman treatment were held to a lower standard of proof because of the inapplicability of the alimony preclusion.\textsuperscript{130} Although this inequity made sense given the


\textsuperscript{126} *Hessen*, 308 N.E. 2d at 410.

\textsuperscript{127} Such as the wife in *Hessen*. Id. at 412 (noting that because of the effect that fault finding has on the opportunity to obtain support and other post-divorce relief, courts should use broad discretion in considering the grant of a divorce based on cruel and inhuman treatment).

\textsuperscript{128} Id. (holding that the economic consequences of granting of the divorce must be treated as an influential factor in determining whether to grant divorce).

\textsuperscript{129} See *Lord v. Lord*, 409 N.Y.S.2d 46, 49 (N.Y. 1978) (noting that the division of assets and alimony would be considered differently if the wife and children were in danger of becoming wards of the state).

\textsuperscript{130} See, e.g., Filippi v. Filippi, 384 N.Y.S.2d 1010, 1011 (N.Y. App. Div. 1976) (taking note of the fact that most courts applied a lesser standard of proof for women seeking divorce on the ground of cruel and inhuman treatment, but that in the case at bar, the court would strictly apply the *Hessen* standard where
financial issues dependent women faced in the event that their financially secure husbands should decide to divorce them, the extreme domestic situations some husbands were incapable of escaping highlighted the need for reform of New York’s divorce finance laws.\footnote{In \textit{Johnson v. Johnson}, the court rejected a husband’s petition for divorce on grounds of cruel and inhuman treatment despite the fact that his wife had been absent from the family home for extended periods and had assaulted him on at least two occasions. \textit{Johnson v. Johnson}, 478 N.Y.S.2d 54 (N.Y. App. Div. 1984) (holding that the wife’s actions did not affect the husband in any deleterious way). In \textit{Denny v. Denny}, the court of appeals denied the plaintiff husband his application for divorce on the grounds of cruel and inhuman treatment despite his wife’s refusal to move with him to another city when transferred by his employer, that he felt dominated by her to the point that his self-confidence was shaken, and that she deprived him of reasonable intimacy. \textit{Denny v. Denny}, 409 N.Y.S.2d 443 (N.Y. App. Div. 1978). In its consideration of the husband’s request for a divorce, the court of appeals took note of the fact that the wife would be unable to obtain alimony if it granted her husband a divorce. \textit{Id.}}

\textbf{C. Equitable Distribution}

One of the primary reasons for arbitrary and inconsistent divorce holdings was New York’s failure to update divorce finance law when the state updated its divorce law in 1966.\footnote{DiFonzo & Stern, \textit{supra} note 20, at 588.} At the time, New York retained the traditional common law view that legal title to property was determinative of its ownership upon divorce.\footnote{Brett R. Turner, \textit{Equitable Distribution of Property} 7 (3d ed. 2005).} Under such a system, the husband, who was typically the homeowner and breadwinner, left the marriage with a majority of the property.\footnote{See Bea A. Smith, \textit{The Partnership Theory of Marriage: A Borrowed Solution Fails}, 68 Tex. L. Rev. 689, 697 n.47 (1990) (citing Jacob, \textit{supra} note 81, at 5).} This inequality in post-divorce assets was typically remedied by awarding the wife permanent or long-term alimony.\footnote{Jacob, \textit{supra} note 81, at 5. In contrast to “title” states, such as New York, a smaller number of states adhere to the “community property” method...}
Thus, in divorce proceedings where fault was asserted against the wife, courts were wary of granting the husband a divorce for fear of leaving the wife with no property or earning potential and without the relief of alimony.\footnote{See, e.g., Hessen v. Hessen, 33 N.Y.2d 406, 410 (N.Y. 1975) (taking note of the fact that, under section 236 of the Domestic Relations Law, “a divorce granted on the basis of the wife’s ‘misconduct’ will deprive the wife of both her rights to alimony and the exclusive occupation of the marital residence”).}

New York struggled with the effects of its divorce finance laws throughout the 1970s, while nearly all other states adopted equitable distribution laws.\footnote{DiFonzo & Stern, \textit{supra} note 20, at 587.} Equitable distribution deemed marriage an economic partnership\footnote{\textsc{Unif. Marriage & Divorce Act} § 160, 9A U.L.A. XX (1970). Using definitions typical of community property systems, the Uniform Marriage & Divorce Act became a prototype for widely adopted equitable distribution statutes, in which courts are directed to make a just division of marital property based on a series of factors.} and aimed to “credit the unpaid work that the typical non-employed homemaker put into the partnership.”\footnote{See Jens-Uwe Franck, \textit{‘So Hedge Therefore, Who Join Forever:’ Understanding the Interrelation of No-Fault Divorce and Premarital Contracts}, 23 \textit{Int’l J. L., Pol. & Fam.} 235, 243 (2009).} However, women’s groups and state legislators in New York feared the excessive discretion granted to judges under the equitable distribution system, and were concerned that distributing property equally would only encourage divorce.\footnote{DiFonzo & Stern, \textit{supra} note 20, at 587.}


\footnote{\textsc{Nancy Cott}, \textit{Public Vows: A History of Marriage and the Nation} 206 (2000).}
New York and Divorce

amended following the 1979 Supreme Court case Orr v. Orr. The Supreme Court fundamentally altered New York’s divorce finance law by determining that gender based divorce statutes, such as the one that governed alimony awards, were unconstitutional due to their disparate treatment of male and female parties. The Court held that spousal support might be awarded to either deserving spouse regardless of gender. The ruling was ultimately incorporated into the Equitable Distribution Law, and courts now award maintenance to the less monied party in order to maintain that spouse’s standard of living in the aftermath of divorce.

D. Maintenance

Property division was tied to alimony, and New York’s change to equitable distribution altered the law of alimony in a fundamental way. As in the Equitable Distribution Law, the legislature replaced alimony with “maintenance” payments. While the practice of alimony was rooted in the concept that the husband was the primary source of financial support, and was thus his wife’s caretaker, maintenance is based on the idea that marriage is an economic partnership. Thus, under the alimony system a husband was charged with financially supporting his ex-wife because she lacked the means to do so herself. Conversely, maintenance payments represented compensation for loss of

142 Orr v. Orr, 440 U.S. 268 (1979) (holding that an Alabama statute authorizing the imposition of alimony obligations on husbands in favor of wives, but not in favor of husbands, was unconstitutional).
144 Orr, 440 U.S. at 280–82.
145 DiFonzo & Stern, supra note 20, at 588.
146 Id.
147 Which is the reason that alimony payments ceased once the ex-wife remarried, since she would then be “taken care of” by her new husband.
earning from the marital partnership.\textsuperscript{148} Maintenance payments were calculated in order to preserve the parties’ standard of living during the marriage and to provide for the non-moned spouse’s “reasonable needs.”\textsuperscript{149} Courts were encouraged to limit the duration of maintenance payments as a result of the legislation, pushing for recipients to become self-sufficient as quickly as possible.\textsuperscript{150} Effectively, New York ended permanent spousal support.\textsuperscript{151}

Despite the equitable change intended by the legislation, the Equitable Distribution Law and subsequent shift away from permanent alimony did not have the effect of alleviating post-divorce economic inequality,\textsuperscript{152} since most divorcing couples have few assets to divide.\textsuperscript{153} In most divorces, the equitable distribution of assets does not offer each spouse enough of a solid economic foundation in the aftermath of divorce.\textsuperscript{154}

It is much more costly to maintain two households than a single marital home. However, the standard of living for men generally increases in the aftermath of divorce while for women, it significantly decreases.\textsuperscript{155} This is because a high earning spouse—traditionally the husband—can recoup his or her old lifestyle over time whereas a low-earning or non-working spouse cannot do so as


\textsuperscript{149} 1980 N.Y. LAWS 434 (codified as amended at N.Y. DOM. REL. LAW. § 236 (B)(6) (McKinney 2009)). Maintenance ceased, under this system, when the wife remarried. N.Y. DOM. REL. LAW. § 236 (B)(6)(c) (West 2010). However, maintenance could continue even if the wife was in a relationship with another man, but not technically married. \textit{Id.} at § 6(d).

\textsuperscript{150} Garrison, \textit{supra} note 148, at 640.

\textsuperscript{151} In 1986, however, the New York legislature amended the Equitable Distribution Law to permit indefinite maintenance payments in situations that involved older and disabled women with no job skills other than homemaking and child rearing. N.Y. DOM. REL. LAW § 236(B)(6) (McKinney 1986) (amended 2010); \textit{See also}, N.Y. JUR. DOM. REL. § 2206 (McKinney 2010).

\textsuperscript{152} DiFonzo & Stern, \textit{supra} note 20 at 589, 598–99.

\textsuperscript{153} Garrison, \textit{supra} note 148, at 662–64.

\textsuperscript{154} \textit{Id.} at 658.

\textsuperscript{155} DiFonzo & Stern, \textit{supra} note 20, at 595–96.
New York and Divorce

easily or quickly, if at all. Thus, the post-divorce economic needs of most non-working or low-earning spouses simply cannot be achieved by equitable distribution of the marital assets. The income of high earning spouses is inevitably a much more valuable “asset” than marital property. It follows that the non-earning spouse, who, with the passage of the Equitable Distribution Law, traded an award of temporary maintenance for permanent alimony, underwent a substantial economic deterioration in the wake of divorce.

III. THE NEW LAW: CRITICISM, PRAISE, AND POTENTIAL PROBLEMS

The adoption of a no-fault divorce option has engendered both excitement and outrage among the various parties involved in this debate. Women’s groups in particular, who blame no-fault divorce systems for the decreased occurrence of domestic violence, are praising the ability of spouses to unilaterally divorce one another as a major step forward. However, opponents point to the United States’ high divorce rate—connecting no-fault divorce to the ease with which couples can dissolve their marriages. Further, confusion over the new law’s maintenance provisions has scholars and practitioners concerned, and will undoubtedly be the cause of future debate.

A. Unilateral Divorce

The key component to New York’s divorce law reform is the addition of Domestic Relations Law Section 170(7), which allows parties to divorce if the marital relationship has irretrievably broken down for at least six months, and one party has so stated under oath. This judgment can be granted only after the issues of equitable distribution, spousal maintenance, child support, counsel

156 DiFonzo & Stern, supra note 20, at 597.
157 Garrison, supra note 148, at 664.
158 See DiFonzo & Stern, supra note 20, at 597.
159 N.Y. DOM. REL. LAW § 170(7) (McKinney 2010).
and expert fees, child custody and visitation are resolved.\textsuperscript{160}

One of the chief advantages of section 170(7) is the ability of one spouse to dissolve the marriage without the consent of the other.\textsuperscript{161} This provision has received a great deal of criticism. Women’s groups such as the National Organization of Women fear that this “divorce on demand” will lead to increased divorce rates and will negatively impact the financial security of women who will be left destitute once their husbands unilaterally divorce them.\textsuperscript{162} Other opponents of the law cite children as the “voiceless third parties in divorce,” and argue that making divorce “easy” gives couples the opportunity to split up despite the fact that repairing their marriage for the sake of their children may be possible.\textsuperscript{163} Others argue that, because women seek divorce more frequently than men,\textsuperscript{164} no-fault divorce disproportionately harms

\textsuperscript{160} Id. These issues may be resolved by either the parties or determined by the Court. \textit{Id.}

\textsuperscript{161} See id.


On this score, a consensus among social scientists has emerged. It distinguishes between children in families where parents are engaged in unremittingly high levels of conflict and children in families where parents are unhappy but have low conflict. In high conflict families, children are better off if their parents divorce. In low-conflict families, however, children are better off if their parents stay together and repair the marriage. Sadly . . . the majority of parental divorces today occur in low-conflict situations.

\textit{Id.}

New York and Divorce

men since men are unlikely to secure custody of children in the aftermath of divorce and are often ordered to make child support payments.\textsuperscript{165}

Conversely, family law scholars and practitioners point out that one of the most surprising aspects of the no-fault divorce system is that it seems to have little, if any, effect on divorce rates.\textsuperscript{166} At the time most states across the country began adopting no-fault divorce systems, the divorce rate doubled.\textsuperscript{167} However, this increase in divorce rates occurred “in equal measure in those states adopting unilateral divorce” as in states where fault-based divorce systems remained in place.\textsuperscript{168} Additionally, proponents of the new law point out that the state with the lowest divorce rate in the country, Massachusetts, has permitted no-fault divorce since 1975.\textsuperscript{169} Although divorce rates may increase in the aftermath of

huffingtonpost.com/Vicki-larson/why-women-want-out-more-t_b_792133.html (discussing a study done by the National Marriage Project of the University of Virginia, which found that two-thirds of all divorces are initiated by women);


\textsuperscript{167} Id.

\textsuperscript{168} Id.

\textsuperscript{169} Cherlin, supra note 162. “[Massachusetts’] divorce rate is low for at least two . . . reasons: [f]irst, its population is highly-educated, and educated people avoid marrying young, which is a risk factor for divorce. Second, it has a large Catholic population, and Catholics are still somewhat less likely to divorce.” Cherlin notes that New York has an equally educated populous with a great deal of Catholics. Id. Therefore, he expects the rate of divorce in New York (which is currently tied for the fifth lowest rate of divorce in the country) to remain low despite no-fault divorce. Id.
New York’s adoption of no-fault divorce, studies of other states show that this increase will likely be temporary.\footnote{Id. New Yorkers waiting to file for divorce, who did not have grounds to do so, will file immediately, which will likely result in a short swell in divorce rates. \textit{Id.}}

Further, advocates of no-fault divorce herald it as one of the best means of combating domestic abuse.\footnote{\textit{See, e.g.}, Stevenson, \textit{supra} note 166.} Reports of domestic violence tend to decrease in states that adopt no-fault divorce.\footnote{Studies show a marked decrease in domestic violence among states that adopted no-fault divorce systems relative to states (such as New York) that did not. \textit{Id.}} Studies show that the decrease in domestic violence is not simply due to the fact that spouses being abused can unilaterally leave their abusive partners, but because “abusive spouses understand that they will be left.”\footnote{The studies show a thirty-percent decline in domestic violence, which indicates that violence in lasting marriages decreased. \textit{Id.}} Additional praise of no-fault divorce systems attribute decreased rates of female suicide to the ability of abused or unhappy women to unilaterally leave their partners.\footnote{\textit{Id.}} Therefore, it appears that no-fault divorce systems tend to provide an important benefit to particularly defenseless members of society.

\section*{B. Maintenance}

The enactment of Domestic Relations Law section 170(7) is only a small portion of the New York divorce reform package.\footnote{See \textit{N.Y. DOM. REL. LAW} §§ 236–38 (McKinney 2010).} The accompanying bills that amend Domestic Relations Law sections 236, 237, and 238 contain provisions to protect the less monied spouse.\footnote{\textit{Id.}} The new laws require the court to provide for temporary support during a divorce proceeding when the parties have unequal financial resources.\footnote{\textit{Id.}} The revision of \textit{pendente lite}
awards, which order the monied spouse to pay interim counsel fees for the non-monied spouse during the course of the case so as to enable her or him to carry on or defend it, were designed to better reflect the circumstances couples face during divorce proceedings. The new laws create a presumption that the less monied spouse will be entitled to the payment of these fees, which will be set by a new formula. Further, the new law revises the way that temporary maintenance awards are calculated by granting justices a mathematical formula to determine presumptive awards. Lastly, the method for determining post-divorce maintenance awards received an overhaul with the addition of extra factors that the court may consider when deviating from the presumptive award.

The new statutory guidelines established for determining temporary maintenance are intended to provide greater consistency and fairness. However, the mathematical formula for awarding presumptive temporary maintenance, and the factors to be considered by courts when deviating from this presumptive award, may be a cause of some confusion for justices. The temporary maintenance formula for presumptive awards requires Supreme Court justices to take two sets of calculations into account. The new mathematical formula for determining temporary maintenance is meant to provide uniformity and predictability. However, the

---

178 Pendente lite awards are the interim support payments provided to the lower income spouse while divorce litigation is pending. 65 AM. JUR. 2D Receivers § 3 (West 2010).
179 The spouse with larger financial net worth.
180 N.Y. DOM. REL. LAW § 236(B)(5) (McKinney 2010).
181 Id.
182 Id.
183 Id.
184 Stashenko, supra note 15.
185 First, the formula requires subtracting twenty-percent of the lower earning spouse’s income from 30% of the higher spouse’s income. (M=0.30(H)–0.20(L)). N.Y. DOM. REL. LAW § 236(B) (McKinney 2010). The second formula calls for the court to subtract the lower earning spouse’s income from 40% of the combined income of both spouses. (M=0.40(H+L)–L). Id. The Court will use the lower of the two figures to set the guideline amount of temporary maintenance. Id. at 5-a(c)(1).
determination of temporary maintenance awards is ultimately at the discretion of the courts, as justices are free to deviate from the presumptive award if he or she “finds that the presumptive award is unjust or inappropriate.” Thus, a method that was intended to take the uncertainty and confusion out of determinations of temporary maintenance, by providing parties with a mathematical formula, is not much changed from that of the old method of calculating temporary maintenance. Additionally, the statute’s vague guidance for justices on how to determine whether a presumptive award is “unjust or inappropriate” will undoubtedly confound justices. What is worse, however, is that the lack of guidance will ultimately create windfalls for some lucky parties while others will be required to make do with the scant terms of the presumptive award.

Additionally, the presumptive temporary maintenance formula only applies to the first $500,000 of income. Any income above this point may not be calculated into the abovementioned mathematical formula but must be considered separately by the court in determining temporary maintenance. This will inevitably be an added source of confusion and possible tension for divorcing New Yorkers. Because New York is one of the wealthiest states in the country, with six of its sixty-two counties among the wealthiest per capita in the nation, it is likely that many justices will have to determine additional guideline amounts of temporary maintenance through the consideration of a variety of imprecise factors presented by the statute. Thus, the new divorce finance laws largely leave wealthy parties to litigate the details of

---

186 Id. at 5-a(e).
187 Stashenko, supra note 15.
188 N.Y. DOM. REL. LAW § 236(B)5-a(b)(5) (McKinney 2010).
189 Id.
191 N.Y. DOM. REL. LAW § 236 Part B, 5-a(c)(2)(i)-(xix) (McKinney 2010). These factors include the marriage’s length, the substantial differences in the parties’ incomes, and the parties’ standard of living established during the marriage. Id.
their pre- and post-divorce maintenance, and this will likely affect many New York couples.

Another foreseeable concern with the law’s new divorce finance provisions lies in the determination of post-divorce maintenance awards. In calculating maintenance, justices are charged to consider an array of factors such as the length of the couple’s marriage, the standard of living of the parties, the earning capacity of the parties and their prospects of employment, and other factors.192 These factors are often subjective and can be difficult to measure, which may allow justices to vary dramatically in their maintenance awards. An “outspoken supporter”193 of New York’s implementation of no-fault divorce, former Appellate Division Justice Sondra M. Miller, noted that “if [I] were still hearing divorce cases, [I] would be confused by how to set maintenance.”194 While the new law set up a Law Revision Commission to study the effect of the new maintenance provisions on the setting of maintenance levels around the state, in the interim, there may be significant differences in maintenance awards that will unduly disadvantage some spouses throughout the state.195

The new maintenance formulas are the new law’s most striking weakness and will likely be the source of much confusion. The reform of New York’s divorce laws was intended to make divorce more equitable and less time consuming and costly. However, the new temporary and post-divorce maintenance provisions offer confusing and potentially time consuming solutions to an already uncertain and drawn out process.

IV. RECOMMENDATIONS AND PROPOSED SOLUTIONS

The enactment of a no-fault option for the dissolution of marriage in New York is a positive development in the domestic

192 N.Y. DOM. REL. LAW § 236 Part B, 5-a(e)(1)-(q) (McKinney 2010).
193 Stashenko, supra note 15.
194 Id. Miller noted: “If I were a judge, does it mean that I consider child support? The law is not clear. It is a very troublesome piece of legislation . . . .” Id.
195 Id.
relations laws of the state. However, while fault based divorce laws
have their evils, they also encourage (or at times force) spouses to
negotiate with one another in order to secure adequate post-divorce
settlements. The default rules provided by no-fault divorce systems
ignore the fact that each divorce is as unique as the circumstances
that brought the couple to desire an end to their marriage. Often
times, one party is at “fault,” and if post-divorce settlements are to
be properly calculated, all of the facts should be considered,
including the blameworthy conduct of the parties. Further, in the
likely event that New York courts continue to consider fault in the
distribution of property and maintenance awards only in the most
egregious cases, New Yorkers contemplating marriage should
enter into premarital or post marital agreements so that they have
the requisite control over their post-divorce lives.

A. WHY A LITTLE FAULT IS A GOOD THING

While “fault allegations and fault trials add significantly to the
cost, delay and trauma of matrimonial litigation,” in many cases
litigants use fault as a tactical advantage in securing a more
equitable post marital settlement.\textsuperscript{196} This “tactical advantage”
relates to the fact that many courts do not adequately provide post-
divorce financial protection for the non-monied spouse.\textsuperscript{197} For true
post-divorce equitable distribution, New York’s divorce finance
laws should take fault into account in situations where fault is
relevant.

The reasons for divorcing and the financial position of the
parties are unique in each divorce. Often some form of fault does
exist, and in order for courts to adequately address the dissolution
of the marriage before them, they should be able to consider all of
the events that led the parties to seek divorce. In keeping with the
Court of Appeals’ holding in \textit{O’Brien}, New York courts only allow
fault to be a consideration in maintenance awards and property
division in “egregious cases which shock the conscience of the

\textsuperscript{196} Miller, \textit{supra} note 3, at 553; See also Kolker & Hurtado, \textit{supra} note 6.

\textsuperscript{197} Peter Nash Swisher, \textit{The ALI Principles: A Farewell to Fault-But What
Remedy for the Egregious Marital Misconduct of an Abusive Spouse?} \textit{8 DUKE J.
New York and Divorce

court.198 In *O'Brien*, the husband sought consideration of the wife’s fault in respect to the equitable distribution of the couples’ marital property.199 However, the court held that there was no suggestion that the wife was guilty of fault sufficient to shock the conscience of the court, as was required for fault to be taken into account in the equitable distribution of property.200

New York courts should reconsider the *O'Brien* standard and allow for the consideration of fault. Of course, this infusion of fault would have no bearing on the actual granting of divorce but would be considered solely for the purposes of maintenance and property distribution. According to the American Bar Association, marital fault is a “factor” in awarding maintenance and dividing property in twenty-five states and the District of Columbia.201 If no-fault divorce is to “do justice,”202 New Yorkers should have the same opportunity.203

By removing the need for fault, New York’s Domestic Relations Law also takes away the strategic position that fault often plays in divorce settlements.

The threat of a fault divorce trial has often been used as a negotiating tool between warring spouses . . . [W]hen a spouse didn’t want to divorce for religious or other reasons, the threat of a trial airing marital disputes or proving the allegations of fault, might be used as a negotiating tactic paving the way for better settlement terms . . . No-fault

---

198 See, e.g., O’Brien v. O’Brien, 66 N.Y.2d 576 (N.Y. 1985). Although the *O’Brien* Court failed to define “egregious misconduct,” lower courts have concluded that in order for conduct to affect equitable distribution, non-economic misconduct must “consist of behavior that falls well outside of the basis for an ordinary divorce action.” Thomas A. Elliot, *Discovery and ‘Non-Egregious’ Marital Fault*, 12 No. 1 N.Y. FAM. L. MONTHLY 3, 4 (Sept. 2010).


200 Id.


202 Wilson, supra note 27.

203 Id.
divorce [] takes this card out of the mix.204

The positive effects of no-fault divorce in New York should not preclude courts from considering the harmful symptoms of real marital fault when structuring the couple’s post-divorce property settlement or maintenance allocation.205

Many courts in jurisdictions where the legislature enacted no-fault divorce laws have held that fault may be considered.206 As is evidenced by the decisions of courts in other jurisdictions, the fault of one or both parties may properly be considered in respect to alimony, spousal support, or property division pursuant to a divorce based on no-fault grounds. Because New York’s current policy of not considering fault during maintenance determinations is similarly non-statutory, its courts should note the reasoning behind the decisions of other jurisdictions and make a common-law based change by allowing the consideration of fault in appropriate circumstances.

Other no-fault states have considered fault in maintenance determinations. In Huggins v. Huggins, the court rejected the husband’s assertion that Alabama’s no-fault divorce statute precluded the consideration of fault in the court’s grant of alimony.207 The court wisely indicated that marital misconduct was a natural aspect of deciding alimony payments if there was fault to be taken into consideration.208 The court held that a trial court might consider fault when making a property division, even if it

204 Kolker & Hurtado, supra note 6.
205 Evans, supra note 34, at 475.
207 Huggins, 331 So.2d at 707–08.
208 Id.; see Sides, So.2d at 679 (allowing fault to be considered); Miller, 361 So.2d at 579 (the Court considered the fault of a husband in committing adultery); Cooper, 382 So.2d at 571 (the Court considered the fault of a spouse in committing adultery).
New York and Divorce

does not grant a fault-based divorce.209 The Edwards court went a step further adding that, “the facts and circumstances of each divorce case are different,” and thus trial courts should consider each of the particular facts and circumstances surrounding the case at bar in dividing property.210 In Edwards, the wife had abused drugs and alcohol throughout the couple’s marriage and was thus entitled to less than equitable distribution in light of the fault she played in the marital dissolution.211

A Connecticut court likewise considered a husband’s fault in connection with an award of maintenance and division of marital property in Sweet v. Sweet.212 The Sweet court noted that, although fault was not a consideration under the state’s divorce statute, the trial court could properly consider reasons for the dissolution of the marriage in making financial awards.213 This supports the aforementioned proposition that no two divorces are alike and that courts have a duty to review each case before them in its entirety.

Further, in Givens v. Givens, the court noted that “[t]he trial court has discretion in the division of marital property, and a just division does not have to be equal, particularly where one party has engaged in misconduct.”214 The court went on to explain that the conduct of parties during the marriage is a factor to be considered when allocating marital property.215 However, it is important to note the Givens court’s caution: although marital fault should be taken into account in dividing marital property, “it should not serve as a basis for ordering excessive maintenance against, or inadequate marital property to, the offending spouse.”216 The

209 Edwards, 26 So.3d at 1260.
210 Id. at 1259.
211 Id. at 1260.
212 Sweet v. Sweet, 462 A.2d 1031 (Conn. 1983).
213 Id.
214 Givens v. Givens, 599 S.W.2d 204, 205–06 (Mo. Ct. App. 1980) (trial court has discretion in division of marital property and a just division does not have to be equal, particularly where one party has engaged in misconduct).
215 Id.
216 Id.; see Hogan v. Hogan, 651 S.W.2d 585, 587 (Mo. Ct. App. 1983) (the conduct of the parties is among the factors to be considered in the division of marital property).
purpose of considering fault in the distribution of property should not be to punish the party at fault, but to allow for all relevant facts to be considered in establishing an equitable post-divorce settlement.

In an apt holding by the Nevada Court of Appeals in *Heim v. Heim*, the Court noted that the concept of fault is consistent with the statutory requirement that property division and alimony awards be just and equitable, and have regard to the respective merits of the parties.217 Similarly, in *Woodside v. Woodside*, a South Carolina court noted that the marital conduct factor becomes important in equitable distribution when the conduct of one party to the marriage throws a burden on the other party that falls beyond the norms to be expected in a marital relationship.218 In such a situation, the court stated, marital misconduct should affect property distribution.219

As Professor Robin Fretwell Wilson wrote, “Americans care why marriages break apart. Infidelity, violence [and] abandonment matter. This does not mean that we must uncritically embrace the old fault-based divorce laws . . . [but] [i]t does suggest [that] we need a prudent and realistic search for new approaches to enacting our shared moral understanding of marriage.”220 In response to its new no-fault divorce provision, New York courts should take heed of the manner in which other jurisdictions examine fault in the distribution of marital property and maintenance awards. Because fault includes more than “extremely outrageous behavior.”221

---

217 *Heim v. Heim*, 763 P.2d 678, 680 (Nev. 1988) (statutory requirement that awards be “just and equitable” applies to awards of alimony as well as to property disposition).


219 *Id.*


New York and Divorce

York courts must reconsider the standard set forth in *O’Brien* and work toward a more equitable system of marital dissolution.

**B. The Argument for Premarital Agreements in No-Fault Divorce Jurisdictions**

Marriage, at its core, is a contract. Because individuals do not enter into contracts without negotiating the terms, entering into the most important contract of one’s life should not be treated any differently or with less care. Couples can choose to write this important marriage contract themselves via premarital agreements, or they can accept the default contracts written by the State Legislature. Accordingly, New Yorkers considering marriage in the increasingly streamlined divorce system that no-fault divorce encourages, should write the terms of their own marriage and divorce by entering into premarital or post marital contracts.

Couples planning to marry have the option of entering into premarital agreements. Premarital agreements are recognized by all fifty states and are becoming more and more popular among couples contemplating marriage. Under early common law, prenuptial agreements were only applicable to the division of

---


223 Additionally, post-marital agreements are designed for couples that are already married and wish to negotiate the terms of a potential future divorce. ROBERT E. OLIPHANT & NANCY VER STEEGH, *WORK OF THE FAMILY LAWYER* 445–47 (2d ed., Aspen Publishers 2008).

224 Only approximately 5–10% of couples enter into premarital agreements in the United States. Beth Potier, *For Many, Premups Seem to Predict Doom*, HAV. GAZETTE (2003), http://news.harvard.edu/gazette/2003/10.16/01prenu.html. The article interviews Heather Mahar, a fellow at the John M. Olin Center for Law, Economics, and Business at Harvard Law School who surveyed students on Harvard’s campus on premarital agreements and divorce opinions. Mahar found that although respondents to her survey were able to correctly identify the national divorce rate (approximately 50%), the percent that thought they would eventually divorce was a mere 11.7%. *Id.* These inconsistencies underlie the reasoning for why the actual rate of couples that sign premarital agreements is only 5–10%. *Id.*
property upon one spouse’s death. However, due to the increase in divorce rates, courts were forced to expand their acceptance of premarital contracts to govern property distribution in the event of divorce.

The 1970 Florida Supreme Court case Posner v. Posner marked a shift in widespread judicial acceptance of premarital agreements. Although the Posner Court recognized an interest in keeping marriages together, it took note of the rising divorce rate and reasoned that “prospective marriage partners . . . may want to consider and discuss . . . the disposition of their property and the alimony rights of [a spouse] in the event their marriage should fail.”

Premarital agreements are particularly beneficial in no-fault jurisdictions because of the divorce system’s ability to trap couples in disadvantageous financial situations during marriage. For example, a spouse who gives up or postpones his or her professional career to stay home with children or concentrate on caring for the marital home makes personal investments that have little value outside of the marriage. In no-fault divorce systems, where default rules are not totally equitable with regard to post marital maintenance or property division, dependent spouses are at the mercy of their financially secure partner. If that partner unilaterally decides to leave the marriage, these marital investments are devalued because they are economically worthless outside the home. Conversely, the partner who makes such

---

225 Oliphant & Ver Steegh, supra note 223, at 445.
226 Id.
227 Posner v. Posner, 233 So.2d 381, 385 (Fla. 1970) (holding that antenuptial agreements, settling alimony and property rights of the parties upon divorce, if conforming to stringent rules prescribed for ante- and post-nuptial agreements settling property rights of spouses and made in good faith and on proper grounds, cannot be said to facilitate or promote procurement of a divorce and are valid as to conditions existing at the time the agreement was made).
228 Id.
229 Franck, supra note 135, at 254.
230 Id.
investments in the marital home and/or childcare cannot simply walk away from the marriage without suffering economic destitution. Thus, spouses who foresee devoting their time to childcare or domestic management should negotiate the post-divorce rules on permanent maintenance or property division by entering into pre or post marital agreement.

Interestingly, courts have come to regard the enforceability of premarital contracts as a natural consequence of the introduction of no-fault divorce. Prior to the widespread availability of no-fault divorce, courts feared that premarital agreements might provide incentive for a spouse to commit fault and exit his or her marriage without consequence. Because the terms of divorce were already provided for in a pre-nuptial contract, courts believed that the couple had little to fear from divorce and thus little reason to behave. However, in the wake of no-fault divorce regimes and Posner, courts have instead come to view premarital agreements as a straightforward way of resolving the financial issues of divorce. Therefore, premarital agreements should be interpreted as a corrective measure for the disadvantages inflicted by the decline of marriage through the already widespread availability of no-fault divorce.

By entering into premarital agreements prior to marriage, parties may insist on stronger rights to alimony or a larger share of marital property than is provided by the default rules. The fewer the rights to maintenance payments or to a share of the property

---

232 Franck, supra note 135, at 254.
233 Id. at 260–61.
234 Frey v. Frey, 41 A.2d 705, 709–10 (Md. 1984) (explaining that antenuptial agreements settling alimony or property rights of parties upon divorce are not per se against public policy and may be specifically enforced). “The old view’s fear that spouses could induce a divorce through fault, without consequences, because the terms of divorce were settled in a [premarital contract], is no longer persuasive because that spouse [can now] seek a no-fault divorce.” Id.; see also LAURA W. MORGAN & BRETT R. TURNER, ATTACKING AND DEFENDING MARITAL AGREEMENTS 33–34 (2001).
236 Franck, supra note 135, at 238.
237 MORGAN & TURNER, supra note 234, at 35–36.
provided by the no-fault default rules, the more contractual protection seems appropriate.\textsuperscript{238} Therefore, with the advent of no-fault divorce and the failure of New York courts to consider marital fault in the division of property and determination of maintenance awards, couples considering marriage should enter into premarital agreements. Negotiating the terms of a potential divorce may be the best route to a secure and equitable marriage.

CONCLUSION

The enactment of a no-fault divorce system is a positive development for New York’s Domestic Relations Laws. By eliminating the strict fault requirements for divorcing spouses, the law ensures that New Yorkers who wish to dissolve their marriage can do so without added time, expense, or perjury. Additionally, studies of no-fault jurisdictions illustrate other advantages to adopting unilateral divorce such as the decrease in incidents of domestic violence and lower rates of female suicide.\textsuperscript{239}

However, the passage of a no-fault divorce system in New York eliminates any bargaining leverage economically disadvantaged spouses have in the wake of divorce. Despite the fact that in some divorce cases there is assignable fault, New York courts fail to take this into account in awarding maintenance and dividing property.\textsuperscript{240} In light of the state’s adoption of unilateral divorce, New York courts must reevaluate the \textit{O’Brien} standard that allows marital fault to be considered only in egregious cases that “shock the conscience.”\textsuperscript{241} Further, with the loss of fault as a negotiating tool, New Yorkers looking to protect themselves and their assets in the wake of divorce should seek alternative solutions such as premarital agreements.

\textsuperscript{238} Franck, \textit{supra} note 135, at 262.
\textsuperscript{239} See, \textit{e.g.}, The Editors, \textit{supra} note 1 (discussing various pros and cons of no-fault divorce).
\textsuperscript{240} Pappas, \textit{supra} note 162.
PHONEY BUSINESS: SUCCESSFUL CALLER ID SPOOFING REGULATION REQUIRES MORE THAN THE TRUTH IN CALLER ID ACT OF 2009

Alicia Hatfield*

I. INTRODUCTION

In 2008, Doug Bates had a terrifying experience when he was forced to defend his home against what he thought were prowlers. After putting his two toddlers to sleep, he and his wife heard noises coming from their backyard. He grabbed a knife and faced the dark to defend his family. Once outside, he quickly found himself blinded by a spotlight and disoriented by a booming command to drop the knife from his hand. As he was tackled to the ground, he wondered what could possibly have caused a SWAT team to surround his home. The answer to that question was Randal Ellis.

* J.D. Candidate, Brooklyn Law School, 2012; B.A., Philosophy, Brooklyn College, 2008. I would like to thank my family and friends for their love and encouragement, especially my husband. I would also like to thank the faculty of Brooklyn law school and the editors and staff of the Journal of Law and Policy for their efforts on this note, especially Lindsay Lieberman for our many meetings. A special thanks is due to Steven Helfont for his editing and advice. Finally, I would like to thank the practitioners who contributed to this note, Mark Del Bianco and Jerry Grant.

2 Id.
3 Id.
4 Id.
5 See id.
6 Hernandez, supra note 1.
Just moments before, Ellis placed a call to 911 with “spoofed” caller identification (“caller ID”) information, making the call appear to have originated from within Mr. Bates’ home, a practice known as “swatting.” After Ellis told the dispatcher that drugs led him to murder his sister, the SWAT team was deployed to Mr. Bates’ quiet California home. Fortunately, the SWAT team handled the situation with caution and no one was injured. While swatting is one of the many illegitimate uses of caller ID spoofing technology that has garnered significant media attention in recent years, there are many legitimate and socially desirable uses of the technology. Nonetheless, Congress introduced multiple anti-caller ID spoofing bills beginning in 2006.

The Truth In Caller ID Act of 2009 (“TICIDA”), which outlaws the use of caller ID spoofing with intent “to defraud, cause harm, or wrongfully obtain anything of value,” was signed into law.

---

8 See id.
9 Hernandez, supra note 1. Ellis was sentenced to three years in prison. Id.
10 See generally 152 CONG. REC. H3386, H3388 (daily ed. June 6, 2006) (statement of Rep. Engel). Most citizens trust the information displayed on caller ID devices since most remain unaware that caller ID spoofing technology exists. Id.
on December 22, 2010. However, successful caller ID spoofing regulation requires more than a statute outlawing illegitimate uses of the technology, most of which were already illegal under existing federal laws. It is imperative that the Department of Justice, the Federal Communications Commission ("FCC"), and Congress accomplish effective regulation of the industry to curb the nefarious aspects of spoofing and preserve legitimate uses. This Note argues that the TICIDA cannot successfully regulate the caller ID spoofing industry because the criminal penalties under the TICIDA are too minimal to deter most illegitimate users; the TICIDA does not expressly criminalize text message spoofing; and the TICIDA does not create comprehensive regulation of the caller ID spoofing industry. In order to maintain the availability of this technology for legitimate users, the Department of Justice should creatively and aggressively prosecute illegitimate users under alternative federal laws when doing so would result in greater deterrence. In addition, the FCC should request that Congress modify the TICIDA to define the term “call” as both voice and text calls expressly so that text message spoofing does not become a successor technology. Lastly, the FCC should promulgate regulations that facilitate the tracing of spoofed calls and create a Do-Not-Spoof list. Part II explains caller ID spoofing and highlights its most common legitimate and illegitimate uses. Part III analyzes state and federal attempts at anti-caller ID spoofing legislation. Part IV suggests steps the Department of Justice, the FCC, and Congress should take in order to maximize deterrence of illegitimate uses and create successful regulation of the caller ID spoofing industry.

II. BACKGROUND

A. What is Caller ID Spoofing?

To understand how caller ID spoofing is accomplished, it is

helpful to first understand traditional caller ID service. When a phone call is placed over the public switched telephone network or on a cell phone, information about the phone call is sent along with the call itself to the called party. The calling party number ("CPN"), one type of data sent with the call, is a ten-digit number that identifies the phone number from which the call is being placed. If the called party subscribes to a caller ID service, the receiving phone company searches its records for the name that corresponds with the incoming CPN, and the caller ID device displays that information. If the caller blocks her caller ID information by dialing *67 before the called party’s number, the CPN will include a marker to communicate to the receiving phone company that the call is intended to be anonymous. In this case, the receiving phone company will not display the caller’s CPN information.

Another piece of information sent along with a phone call is the automatic number identification ("ANI"). The ANI also contains the ten-digit caller number; however, it is not used for caller ID purposes. This information is sent with the phone call regardless of whether the caller utilized *67 call blocking. The ANI enables premium phone services, such as 800 and 900 numbers, to identify which telephone account to charge for

15 The public switched telephone network ("PSTN"), or the plain old telephone service, is the structure that transmits landline phone calls. HARRY NEWTON, NEWTON’S TELECOM DICTIONARY 667 (20th ed. 2004); see also David Roos, How Telephone Country Codes Work, HOWSTUFFWORKS, http://communication.howstuffworks.com/telephone-country-codes1.htm (last visited Feb. 4, 2011).
16 47 C.F.R. § 64.1601(a) (2006).
17 NEWTON, supra note 15, at 148.
20 47 C.F.R. § 64.1601(b) (2006).
21 Id.
22 NEWTON, supra note 15, at 63.
23 Id. at 147.
24 Id. at 63.
incoming phone calls. When caller ID information is spoofed, both the ANI and the CPN are changed to a fake phone number.

1. Traditional Caller ID Spoofing

Primitive forms of caller ID spoofing were possible if the spoofer had sufficient knowledge of the telephone system to manipulate the signals that communicate caller ID information to the caller ID device. The creation of voice over IP (“VoIP”) technology and the availability of web-based commercial spoofing companies has made caller ID spoofing more accessible. When spoofing is accomplished via a commercial spoofing company, the spoofer first pays for a block of minutes in advance to establish an account. To place a spoofed call, the spoofer either calls the company’s 800 number or visits its website. Next, the spoofer enters the number he is calling, followed by the number he would like displayed as the fake caller.
ID number. Depending on the service used, the spoofer may also have the option to record the call or to alter the sound of his voice. It is impossible to trace spoofed calls except by subpoenaing the spoofing company’s records to determine the identity of the customer.

Spoofing is also possible for individual users of VoIP services. VoIP technology utilizes a text-based method for initiating and ending a phone call, known as Session Initiation Protocol (“SIP”). Spoofing is accomplished when the spoofer uses software, often a program known as Asterisk, to alter his SIP information. Multiple websites provide step-by-step instructions for this process. Most VoIP providers prevent their users from altering their SIP information, but some providers do not secure their systems.

2. Text Message Caller ID Spoofing

Text message spoofing takes place when a party changes the “from” information in a text message so that it appears that the text message was sent from a different telephone. In this way, text
Phoney Business

message spoofing is similar to caller ID spoofing, except that it affects mobile phones exclusively. Text message spoofing is usually accomplished through a text message spoofing website, often owned by the same companies that own caller ID spoofing websites. The only way to verify that the text message was spoofed is to look at the alleged sender’s phone records for the absence of the outgoing message.

B. Uses of Caller ID Spoofing

Caller ID spoofing has many uses. Although some states’ approach to caller ID spoofing classifies all caller ID spoofing as illegitimate, Congress has recognized legitimate uses. The floor debates on the TICIDA evince a Congressional intent to secure the availability of the technology for legitimate users. This section discusses which uses Congress labeled as legitimate and illegitimate.

1. Legitimate Uses of Caller ID Spoofing

There are many legitimate users of caller ID spoofing, including business professionals who use the technology to prevent their personal numbers from becoming public and call centers that project incoming phone numbers on their outgoing lines. Doctors


42 Id.

43 Id.


46 See, e.g., 156 CONG. REC. H2522–24 (statement of Rep. Boucher); 152 CONG. REC. H3386, H3387 (statement of Rep. Markey). By spoofing, telemarketing companies are able to provide the recipients of their calls with a viable return number. The Truth in Caller ID Act: Hearing on S. 704 Before the
and other professionals, who must make occasional phone calls from home, use the service to project their office numbers, informing their clients that they are calling while keeping their private numbers undisclosed. Teltech Systems, Inc. (“Teltech”) reports that the service is also useful for journalists who want to keep their personal numbers private. Additional uses of caller ID spoofing include “seeking criminals who have jumped bail, tracking down child support payment deadbeats, . . . providing whistleblowers anonymity in making disclosures,” and facilitating debt collection. Even Congress members use the technology so that outgoing calls display the office’s main number instead of the numbers of their personal lines.

Arguably, the most socially valuable use of spoofing technology is to protect domestic violence shelters and victims, and this has long been a congressional priority when dealing with caller ID technology. In 1995, the FCC passed regulations requiring that individual users have the capability to block their

S. Comm. on Commerce, Sci., & Transp., 110th Cong. 5 (2007) [hereinafter Cerasale Statement] (statement of Jerry Cerasale, Senior Vice-President, Direct Marketing Association). If a consumer attempted to return a telemarketing call to an outgoing line, the number would be busy. Id.


Teltech is the parent company of SpoofCard, the largest caller ID spoofing company. Thomas, supra note 29.


See id.
personal information when making calls by dialing *67. Many states introduced regulations that required blocking options to be provided free of charge, partially out of concern for domestic violence victims. Thus, government officials have consistently tried to minimize the likelihood that caller ID could deliver the location of a domestic violence victim or shelter to her abuser. By using caller ID spoofing, victims are able maintain contact with loved ones safely.

2. Illegitimate Uses of Caller ID Spoofing

Caller ID spoofing also has many nefarious uses. Proponents of caller ID spoofing legislation focused on particularly egregious examples of illegitimate uses during debates. This section discusses the four nefarious uses of spoofing technology most

54 47 C.F.R. § 64.1601(b) (2006).

Carriers must arrange their CPN-based services, and billing practices, in such a manner that when a caller requests that the CPN not be passed, a carrier may not reveal that caller’s number or name, nor may the carrier use the number or name to allow the called party to contact the calling party.

Id.

55 Telephone Firms Give in on Caller-ID Blocking, NEWSDAY, Oct. 9, 1990, at 41; see also Timothy C. Barmann, Cox Communications Cancels Plans to Charge for Privacy Service, PROVIDENCE J., Jan 9, 2001.

56 See, e.g., Bob Wyss, Wires Crossed in Caller ID Blocks Phone Companies Promise Privacy, but Failures Abound, PROVIDENCE J., March 5, 1995, at 1A; Telephone Firms Give in on Caller-ID Blocking, supra note 55. Opponents of caller ID forcefully argued that the service made customers vulnerable by releasing their private information without regard to whether the customer wanted to enroll in the service, and that this was particularly dangerous in the context of domestic violence shelters and victims. Telephone Firms Give in on Caller-ID Blocking, supra note 55.


often cited by proponents during debates on the TICIDA: fraud,\textsuperscript{59} swatting,\textsuperscript{60} harassment,\textsuperscript{61} and political harassment.\textsuperscript{62}

\textit{i. Fraud}

Caller ID spoofing is used to perpetrate fraud in two ways: to make victims believe a spoofer is a trusted entity or to make a trusted entity believe a spoofer is his victim.\textsuperscript{63} In other words, caller ID spoofing can be used either to trick a victim into revealing confidential information, often called phishing, or to verify identity fraudulently when calling an institutional entity.\textsuperscript{64} Phishers use their fraudulently gained information to transfer money from bank accounts, to sell credit card numbers to third parties, or to apply for credit cards or loans in their victim’s name.\textsuperscript{65} Microsoft research estimates that phishers stole over $61 million in 2007.\textsuperscript{66}

Traditionally, phishers used spoofed websites or emails to trick users into entering confidential usernames and passwords.\textsuperscript{67} Now, phishers can also use spoofed caller ID information.\textsuperscript{68} In Sterling,
Phoney Business

Michigan, for example, residents received phone calls appearing to originate from the local courthouse.69 Victims were told they had “missed jury duty” and they would be arrested if they did not immediately provide their Social Security number.70

Caller ID spoofing is also used to fraudulently verify identity to gain access to confidential accounts.71 For example, Western Union has proceeded with cash transfers after credit card thieves spoofed the credit card holder’s caller ID information, making the call appear to have been placed by the individual whose identity was stolen.72

Voicemail hacking is another example of fraudulently verifying identity.73 Many voicemail companies provide users access to their mailbox via caller ID verification without requiring a password.74 This security flaw was famously exposed in 2006, when SpoofCard75 suspended over fifty accounts, including socialite Paris Hilton’s, because of suspected voicemail hacking activity.76 Notwithstanding the account suspensions, the Los Angeles District Attorney investigated voicemail hacking in 2008 after receiving complaints.77 Ultimately, SpoofCard’s parent company agreed to a permanent injunction, requiring that it may no longer advertise that it is “legal in all 50 states, if that is not the

69 Id.
70 Id.
73 See Tom Gilroy, Software Firm, Two Cell Phone Providers Settle DA’s Allegations of Illegal ’Spoofing,’ 9 COMPUTER TECH. L. REP. (BNA) 616 (Dec. 19, 2008).
74 Id.
75 Spoofcard is the largest spoofing company, with over three million customers. Thomas, supra note 29.
76 Robert McMillan, Paris Hilton Accused of Voice-Mail Hacking, INFOWORLD (Aug. 25, 2006), http://www.infoworld.com/d/security-central/paris-hilton-accused-voice-mail-hacking-457. Many gossip columns reported that Hilton used the technology to hack into actress Lindsay Lohan’s voicemail, although Hilton’s spokesman denied the allegations. Id.
77 Gilroy, supra note 73.
case . . . ."78 In addition, T-Mobile and AT&T were enjoined from advertising that pin-free voicemail access was a secure method of verification.79 Despite these efforts, voicemail hacking still occurs.80 In April 2010, the former publicity director for Dolce & Gabbana, Ali Wise, faced up to four years in prison but instead pleaded guilty to hacking into at least four people’s voicemail accounts, listening to and deleting their messages.81

ii. Swatting

Swatting occurs when a spoofed call is placed to an emergency number.82 The caller claims that an emergency is taking place at the location of the spoofed number.83 This process is known as swatting because the goal is to cause the deployment of a SWAT team to the location from which the call appears to originate.84 In 2009, for example, police caught a group of men who had prank called over sixty cities, claiming that hostage situations were in progress.85 The scheme cost the cities over $250,000 in emergency response expenses, claimed over 250 victims, and injured two.86

Emergency service prank calls have serious implications. The calls can overload the emergency response system itself, and can prevent police officers from responding to legitimate calls.87

78 Id.
79 Id.
81 Id. Ms. Wise hacked into one victim’s voicemail over 337 times. Id.
83 Id.
85 Thomas, supra note 29.
86 Id.
Moreover, some jurisdictions have a policy that police must investigate every serious call placed to 911, and some jurisdictions treat every call received as an emergency. Anytime the police respond to an emergency, first responders race to reach the scene as quickly as possible, placing the lives of those in the community, as well as the responders themselves, in danger.

**iii. Harassment**

Many commercial caller ID spoofing companies claim they are intended for entertainment via prank calls to friends, but the service is often used to prank call strangers or make threatening and demeaning calls to enemies. Caller ID spoofing can aid harassers and stalkers by tricking a victim into answering a call. In 2009, for example, a man called three women in the middle of the night claiming that he was inside their houses watching them. He called the women’s cell phones and spoofed their home numbers so that his victims believed he was calling from within.

---

88 Id.


91 See PHONEGANGSTER.COM, http://www.phonegangster.com/ (last visited Feb. 4, 2011) (claiming their services are intended for “fun”); see also Real Stories/Uses, SPOOFCARD, http://www.spoofcard.com/stories (last visited Feb. 4, 2011) (advertising testimonials from satisfied customers who were able to use the service to trick their friends).

92 *Sabin Statement*, supra note 61.

the house.  

iv. Political Harassment

Legitimate political “robocalling” is “one of the most-used political campaign tools.”95 However, Congress is concerned about the increasing use of political robocalls made with the intent to trick voters or prevent them from voting.96 Caller ID spoofing is abused to make it appear that one candidate is placing a phone call, when in fact his opponent is using the call to annoy or confuse voters.97 This happened to democrat Scott Kleeb, who ran to represent the 3rd District of Nebraska in the U.S. House of Representatives in 2006.98 Automated calls were placed with spoofed caller ID information, making it look as if Kleeb’s campaign placed the phone calls.99 Voters received the calls overnight and sometimes repeatedly.100 Mr. Kleeb is not the lone victim of this strategy. In South Carolina, police arrested a local Republican for purportedly organizing spoofed political robocalls.101 Spoofing made the calls appear to be from the democratic candidate’s office.102

---

94 Id.
96 153 CONG. REC. E1286 (daily ed. June 13, 2007) (statement of Rep. Green). Automated phone call devices can place as many as 100,000 calls in one hour. Miller, supra note 95, at 215. One automated phone call company reports they are able to call “10 to 20 percent of the U.S. population on a single day.” Id. at 216.
98 Id.
99 Id.
100 Id.
102 Id. The woman was charged with unlawful use of a telephone under state law. Id.
Phoney Business

There is anecdotal evidence that spoofed robocalls are effective at motivating voters to change their vote.\(^{103}\) For example, in New Hampshire, the National Republican Congressional Committee placed repetitive robocalls, late at night, designed to appear to be from the democratic candidate.\(^{104}\) The local newspaper printed a letter from an angry voter who stated she would not vote for the democratic candidate due to the annoying calls.\(^{105}\)

III. STATE AND FEDERAL LEGISLATIVE APPROACHES

Increased concern over the damage caused by illegitimate uses of caller ID spoofing technology has triggered legislative responses at both the state and the federal level.\(^{106}\) Four states have passed anti-spoofing legislation.\(^{107}\) Congress considered six different versions of the Truth in Caller ID Act and three different versions of the Preventing Harassment Through Outbound Number Enforcement Act since 2006.\(^{108}\) This section discusses the viability of the states’ approaches to anti-spoofing legislation and introduces the federal legislation.

\(^{103}\) Miller, \textit{supra} note 95, at 221.

\(^{104}\) \textit{Id.} These calls did not use caller ID spoofing; rather, it was the content of the calls that was misleading. \textit{Id.}

\(^{105}\) \textit{Id.}


\(^{107}\) Louisiana, Mississippi, Florida, and Oklahoma have passed anti-spoofing legislation. \textit{See infra} Part III.A.

A. State Legislation

Several states have proposed or enacted anti-spoofing legislation.\textsuperscript{109} Louisiana’s anti-spoofing bill provides a private right of action and enables the district attorney to “recover a civil penalty”\textsuperscript{110} when “a caller . . . knowingly insert(s) false information into a caller identification system with the intent to mislead, defraud or deceive the recipient of a call.”\textsuperscript{111} Oklahoma’s Anti-Caller ID Spoofing Act and Mississippi’s Caller ID Anti-Spoofing Act also outlaw the knowing falsification of caller ID information to mislead, defraud or deceive.\textsuperscript{112} In October 2008, Florida enacted the Caller ID Anti-Spoofing Act (“the Florida Act”), which outlawed spoofing to “deceive, defraud, or mislead.”\textsuperscript{113} By using the term “mislead,” these statutes have the practical effect of outlawing all caller ID spoofing and rejecting the concept of legitimate uses.\textsuperscript{114} Since the essence of caller ID spoofing is the ability to mislead called parties about the originating phone number, all uses become illegal under these statutes.\textsuperscript{115}

In 2008, Teltech Systems, Inc.\textsuperscript{116} filed a complaint alleging, among other things, that the Florida Act violated the Commerce Clause.\textsuperscript{117} The Southern District of Florida agreed and granted

\begin{thebibliography}{9}
\bibitem{109} Stolar & Gall, supra note 12.
\bibitem{110} LA. REV. STAT. ANN. § 51:1741.5 (West 2010).
\bibitem{111} Id. § 51:1741.4.
\bibitem{113} FLA. STAT. § 817.487 (West 2010).
\bibitem{114} Plaintiff’s Statement of Material Facts in Support of Motion for Summary Judgment, supra note 49, at 2.
\bibitem{115} Id.
\bibitem{116} See supra note 48 (discussing Teltech).
\end{thebibliography}
Phoney Business

Teltech’s motion for summary judgment. Teltech argued that because of cellular phones and VoIP services, it could not know the location of the called parties. For example, if a Teltech client in New York spoofed a call to a California number, the called party could be in fact located in Florida, rendering the New York client liable under the Florida Act. Thus, Teltech argued it could not conduct its business in any state without fear of violating the Florida Act. The court concluded that the Florida Act had “the practical effect of regulating commerce that occurs wholly outside the state of Florida” in violation of the Commerce Clause.

Notwithstanding Teltech, states have continued to pass legislation similar to the Florida Act. It is likely these statutes will be challenged in federal court, but the outcome is unclear. In 2005, the U.S. Supreme Court denied certiorari on an appeal from a Washington state court ruling that a Washington anti-email spoofing statute, similar in effect to the Florida act, did not violate the Commerce Clause. The Florida court did not find this persuasive in Teltech, but other Districts might. The constitutional uncertainty of the existing state anti-spoofing

---

118 See Teltech, No. 08-61664-CIV-Martinez-Brown.
119 Plaintiff’s Statement of Material Facts in Support of Motion for Summary Judgment, supra note 49, at 4. Likewise, simply blocking Florida area codes would not be sufficient to guarantee that no calls went into or originated in Florida. Id. at 5.
120 Id. at 16.
121 Id.
122 Id.
123 See LA. REV. STAT. ANN. § 51:1741.4 (West 2010); MISS. CODE ANN. § 77-3-805 (West 2010).
124 E-mail from Mark Del Bianco, Counsel, Teltech Sys., Inc., to author (Nov. 30, 2010, 10:04 AM) (on file with author). In November of 2010, Teltech filed a constitutional challenge to the Mississippi Caller ID Anti-Spoofing Act in the Southern District of Mississippi. Id.
126 See Teltech, No. 08-61664-CIV-Martinez-Brown; see generally Defendant’s Response in Opposition to Plaintiff’s Motion for Preliminary Injunction at 5 Teltech, No. 08-61664-CIV-Martinez-Brown, 2009 WL 1614869.
statutes, in addition to the small number of states that have enacted legislation, indicate that uniform and comprehensive regulation of caller ID spoofing must be accomplished at the federal level.

B. Federal Legislation

1. The Truth in Caller ID Act of 2009

On April 6, 2006, the long journey of the Truth in Caller ID Acts began.127 Despite six attempts to pass this legislation, three in the House and three in the Senate,128 the TICIDA did not pass both chambers until December 15, 2010.129 On December 22, 2010, President Barack Obama signed the bill and it became Public Law 111-331.130 The TICIDA amends section 227 of the Communications Act of 1934 (“the Act”), Title 47 of the U.S. Code, to outlaw certain uses of caller ID spoofing.131

The TICIDA forbids any person “to cause any caller identification service to knowingly transmit misleading or inaccurate caller identification information with the intent to defraud, cause harm, or wrongfully obtain anything of value . . . .”132 Caller ID service is defined as “any service or device designed to provide the user of the service or device with the telephone number of, or other information regarding the origination of, a call made using a telecommunications service or IP-enabled voice service.”133

---

129 Id.
131 S. 30.
133 § 227(e)(8)(B). Accordingly, the TICIDA regulates VoIP calls, as well as traditional phone calls. § 227(e)(1).
Phoney Business

Generally, a willful and knowing initial violation of the Act is punishable by “a fine of not more than $10,000 or imprisonment for a term not exceeding one year, or both . . . .”\textsuperscript{134} A second offense is punishable “by a fine of not more than $10,000 or by imprisonment for a term not exceeding two years, or both.”\textsuperscript{135} The TICIDA provides the FCC with enforcement power,\textsuperscript{136} however, if the FCC determines that criminal prosecution is warranted, then it shall notify the United States Attorney, who shall bring the appropriate charges in the proper court.\textsuperscript{137} Within six months, the FCC is to “prescribe regulations to implement” the TICIDA’s provisions\textsuperscript{138} and report to Congress “whether additional legislation is necessary to prohibit the provision of inaccurate caller identification information in technologies that are successor or replacement technologies to telecommunications or IP-enabled voice service.”\textsuperscript{139}

The TICIDA provides specific civil and criminal penalties.\textsuperscript{140} A specific civil forfeiture penalty, in addition to the general penalty under the Act, is “not to exceed $10,000 for each violation, or three times that amount for each day of continuing violation . . . .”\textsuperscript{141} In contrast, a specific criminal fine under the TICIDA, of “not more than $10,000 for each violation, or three times that amount for each day of continuing violation,” is “in lieu” of the general fines imposed under the Act.\textsuperscript{142} However, the general criminal penalties under the Act, of imprisonment or a penalty of both fine and imprisonment, may be imposed in addition to the specific criminal fine.\textsuperscript{143} Lastly, the TICIDA authorizes independent state enforcement via civil action in federal district

\textsuperscript{134} Communications Act of 1934, 47 U.S.C.A. § 501 (West 2010).
\textsuperscript{135} Id.
\textsuperscript{136} § 227(e)(3)(A).
\textsuperscript{137} Communications Act of 1934, 47 U.S.C.A. § 401 (West 2010).
\textsuperscript{138} § 227(e)(3)(A).
\textsuperscript{139} § 227(e)(4).
\textsuperscript{140} See § 227(e)(5).
\textsuperscript{141} § 227(e)(5)(A).
\textsuperscript{142} § 227(e)(5)(B). Criminal penalties are imposed for “willful and knowing violation” of the TICIDA. Id.
\textsuperscript{143} Id.
court. After the FCC receives notice of the proceeding, it may, however, intervene in the state action.

2. Past Proposals for Federal Legislation

The Preventing Harassment through Outbound Number Enforcement Act (“PHONE Act”) was an alternative approach to caller ID spoofing regulation proposed in the House. There have been three House versions of the PHONE Act, of which the first was introduced in 2006. The most recent version, the PHONE Act of 2009, passed the House on December 16, 2009, with a resounding 418 “yeas” to one “nay,” yet the bill died in the Senate. Unlike the TICIDA, which amends Title 47 of the U.S. Code, the PHONE Act would have amended Title 18 by adding section 1041, “Caller ID Spoofing.” Section 1041(a) proposed to outlaw:

---

144 § 227(e)(6)(A). The TICIDA provides:

[t]he chief legal officer of a State, or any other State officer authorized by law to bring actions on behalf of the residents of a State, may bring a civil action, as parens patriae, on behalf of the residents of the State in the appropriate district court . . . to impose the civil penalties for violation of this subsection.

Id.

145 § 227(e)(6)(C).


147 H.R. 1110; H.R. 740; H.R. 5304.


149 Id.

150 H.R. 1110.
knowingly [using] or [providing] to another (1) false caller ID information with intent wrongfully to obtain anything of value; or (2) caller ID information pertaining to an actual person or other entity without that person’s or entity’s consent and with intent to deceive any person or other entity about the identity of the caller . . . .

A violation of subsection (a)(1) would have been punishable by a “[fine] under this title or imprison[ment] not more than 5 years, or both . . . .” A violation of subsection (a)(2) would have been punishable by the same terms but with a one year maximum term of imprisonment. In addition, any person convicted under section 1041 would have been subject to forfeiture of “(A) any property, real or personal, constituting or traceable to gross proceeds obtained from such offense; and (B) any equipment, software or other technology used or intended to be used to commit or to facilitate the commission of such offense . . . .”

IV. THE TRUTH IN CALLER ID ACT OF 2009 DOES NOT SOLVE THE PROBLEMS ASSOCIATED WITH ILLEGITIMATE CALLER ID SPoothing

Although the TICIDA specifically outlaws illegitimate uses of caller ID spoofing, such legislation does not sufficiently respond to the issues caused by the continued availability of the technology. The TICIDA’s weaknesses will encourage illegitimate users, thus undermining those who seek to use caller ID spoofing for valid purposes. Further response is necessary because caller ID spoofing provides valuable and legitimate services to society that ought to be preserved, as Congress repeatedly recognized during floor debates on the TICIDA.

151 Id.
152 Id.
153 Id.
For many legitimate users, caller ID spoofing is the best way to keep caller ID private since the development of two new services has severely limited the effectiveness of *67 blocking.\(^{156}\) One of these services is anonymous call rejection.\(^{157}\) This service intercepts incoming calls from blocked or private numbers, which prevents the phone from ringing and instead notifies the caller that the number dialed does not receive anonymous calls.\(^{158}\) This is problematic for domestic violence shelters and victims because courts sometimes order domestic violence victims to maintain contact with their ex-spouse to facilitate child custody arraignments.\(^{159}\) If an abuser utilizes anonymous call rejection, caller ID spoofing becomes an indispensable tool for ensuring contact while protecting the location of the victim.\(^{160}\)

The second new service is TrapCall.\(^{161}\) This controversial service reveals blocked phone numbers, which raises privacy concerns for legitimate users.\(^{162}\) When a TrapCall subscriber receives a blocked call, he can reject it.\(^{163}\) TrapCall then re-routes the incoming call to an 800 number, revealing the caller’s ANI.\(^{164}\) The call is then sent back to the TrapCall subscriber’s mobile phone with the private number displayed on the subscriber’s caller

\(^{156}\) Knight Statement, supra note 26, at 23.

\(^{157}\) Id.


\(^{162}\) Olson, supra note 159. It is argued that the service increases privacy and prevents harassment by allowing individuals to know the identity of callers. Id.

\(^{163}\) Id.

\(^{164}\) Id. While the call is re-routed, the caller continues to hear normal ringing. Id. See supra Part II.A for an explanation of ANI.
Phoney Business

ID.\textsuperscript{165} Caller ID spoofing is the only way to prevent TrapCall from revealing a private number since spoofing alters both the CPN and the ANI, protecting the private caller’s information.\textsuperscript{166}

Notwithstanding the important privacy gains caller ID spoofing provides, the grave personal and public costs of illegitimate uses require a comprehensive federal response. The TICIDA fails to provide this.\textsuperscript{167} Unfortunately, if the TICIDA fails to curb illegitimate uses, Congress might decide to amend the TICIDA to adopt the states’ approach of caller ID regulation, banning the technology outright.

In order to cure the deficiencies of the TICIDA and maintain the viability of caller ID spoofing, the Department of Justice, Congress, and the FCC should take a number of additional steps. First, federal prosecutors should creatively prosecute illegitimate users by charging the spoiler with the federal crime that best fits the illegitimate activities and provides the highest level of deterrence. Second, the FCC should notify Congress that text message caller ID spoofing will become a successor technology if Congress does not amend the TICIDA to define the term “call” as voice and text calls. Last, the FCC should enact regulations to facilitate law enforcement tracing of spoofed calls and create a Do-Not-Spoof list.

\textit{A. The Department of Justice Should Use Creative Prosecution Methods to Enhance Deterrence}

During floor debates on the TICIDA, Congress members complained that spoofers could use the service legally to realize illegitimate ends.\textsuperscript{168} Thus, it was argued, the TICIDA’s criminal proscriptions were necessary to close that gap.\textsuperscript{169} While the TICIDA now provides a direct method for federal prosecution of caller ID spoofing with the intent to “defraud, cause harm, or

\textsuperscript{165} Olson, \textit{supra} note 159.
\textsuperscript{166} \textit{Id.}
\textsuperscript{167} See Part II.B.2 (explaining illegitimate uses).
\textsuperscript{169} \textit{Id.} at S173–174.
wrongfully obtain anything of value,“170 illegitimate caller ID spoofing was potentially prosecutable under numerous federal laws that provided steeper penalties than the TICIDA. The widespread use of caller ID spoofing technology for illegitimate ends suggests that effective deterrence requires vigorous and creative prosecution, which seeks to enhance the potential criminal penalties that illegitimate users face. Consequently, while prosecutors should charge those who use caller ID spoofing for swatting or political harassment under the TICIDA, for other illegitimate uses, it will be more effective to charge a spoofer under alternative federal statutes. The remainder of this section suggests which laws prosecutors should use when seeking to deter illegitimate users of caller ID spoofing, and the specific contexts in which these laws should be used.

1. Fraud Prosecution Methods

 Committing fraud is generally illegal under many federal statutes that carry longer maximum penalties and larger maximum fines than the TICIDA, and those statutes should be used instead of the TICIDA to prosecute spoofing when appropriate.171 The facts of the fraudulent scheme will determine which federal law applies.172 If the spoofer perpetrated fraud on a bank or financial institution, then the spoofer should be prosecuted under 18 U.S.C. § 1344, making him subject to “a fine of not more than $100,000, imprison[ment] of not more than 30 years, or both.”173 Otherwise, the spoofer should be prosecuted under 18 U.S.C. § 1343, which covers wire fraud generally.174 Section 1343 provides that a party who “devises a scheme or artifice to defraud, or for obtaining money or property . . . by means of wire . . . shall be fined under

172 See generally §§ 1029, 1343, 1344.
173 § 1344.
174 § 1343.
Phoney Business

this title or imprisoned not more than 20 years, or both."

While § 1343 should be used to prosecute interstate phishing schemes accomplished by spoofing, by its terms, it applies only to “communication in interstate or foreign commerce . . . .” Furthermore, some courts have held that where all parties are residents of the same state, “all telephone calls are presumed to be intrastate and, absent any indication otherwise . . . wire fraud is not [present].” The TICIDA avoids this loophole since it does not have an interstate activity requirement. Yet, the maximum penalty under the TICIDA is drastically lower than the maximum penalty under § 1343, thus providing a windfall to the intrastate spoofer charged under the TICIDA. Consequently, intrastate spoofers should be charged under state wire fraud and identity theft statutes that have higher maximum penalties when available.

If the spoofer used caller ID information as an access device to commit fraud, the spoofer should be prosecuted under 18 U.S.C. § 1029. Section 1029(a)(1) makes it illegal to “knowingly and with intent to defraud produce, use, or traffic in one or more counterfeit access devices.” An access device is defined as any

175 Id.

176 Black’s law dictionary defines a “scheme” as “[a]n artful plot or plan, [usually] to deceive others.” BLACK’S LAW DICTIONARY (9th ed. 2009).

177 Communications Act Amendments, 1952, 18 U.S.C.A. § 1343 (West 2010). In addition, “money or property [must] be the object of the defendant’s scheme to defraud.” United States v. Martin, 411 F. Supp. 2d 370, 373 (S.D.N.Y. 2006). It is therefore possible that § 1343 does not include schemes to obtain personal information. Lastly, to be liable under § 1343, the use of the wires “must at least be ‘incident to an essential part of the scheme.’” Id. at 374 (quoting United States v. Altman, 48 F.3d 96, 102 (2d Cir. 1995) (emphasis omitted)).


180 Christine Mumford, Spam Gives Way to Data Breach, Phishing in Contest for State Legislators’ Attention, 8 COMPUTER TECH. L. REP. (BNA) 56 (Feb. 2, 2007).


182 § 1029(a)(1).
“code, account number, . . . mobile identification number, personal identification number, . . . or other means of account access that can be used . . . to obtain money, goods, services, or any other thing of value . . . .”

Caller ID information is a unique number that the spoofer uses to obtain things of value in phishing schemes and in false identity verifications. In a phishing scheme, the spoofer uses caller ID to access personal information, which the spoofer can sell or use to obtain goods or cash. When it fraudulently verifies identity, caller ID is an account number that leads directly to the spoofer obtaining money, or valuable confidential information. Thus, § 1029 should be charged when phishers spoof false caller ID numbers to obtain things of value.

Alternatively, if the phisher pretends to be a trusted entity, he can be prosecuted according to the mask he wears or the identity of his victims. Those who falsely assume the identity of a government official should be prosecuted under 18 U.S.C. § 912, which carries a criminal penalty of up to three years imprisonment. If caller ID spoofing is used in connection with a fraudulent telemarketing scheme, the spoofer should be prosecuted under 18 U.S.C. § 2326, which would subject him to an enhanced penalty of up to five years imprisonment. If during that scheme he victimized “ten or more persons over the age of 55; or targeted persons over the age of 55,” § 2326 provides “imprison[ment] for a term of up to 10

---

183 § 1029(e)(1).
184 See Plaintiff’s Statement of Material Facts in Support of Motion for Summary Judgment, supra note 49, at 2 (noting that the government had admitted “that on a telephone call the Caller ID string is akin to the identity of the caller”).

185 RAYSMAN & BROWN, supra note 63.

187 Whoever falsely assumes or pretends to be an officer or employee acting under the authority of the United States or any department, agency or officer thereof, and . . . in such pretended character demands or obtains any money, paper, document, or thing of value, shall be fined under this title or imprisoned not more than three years, or both.

Id.

Finally, voicemail hacking is prosecutable under either the Stored Communications Act\textsuperscript{190} or the Wiretap Act\textsuperscript{191} depending on the actions taken by the spoofer once he accesses the voicemail system.\textsuperscript{192} If the spoofer merely accesses the system, then he should be prosecuted under the Stored Communications Act, which provides a maximum penalty of five years in prison.\textsuperscript{193} If however the spoofer also records copies of messages or deletes messages from the system, he has then intercepted communications and should be prosecuted under the Wiretap Act, which provides for penalties of up to five years in prison.\textsuperscript{194}

2. Swatting Prosecution Methods

The TICIDA is the most effective method of prosecuting swatting. Before the TICIDA, there was no federal statute that outlawed spoofed 911 calls.\textsuperscript{195} However, certain states had


\textsuperscript{190} Electronic Communications Privacy Act of 1986, 18 U.S.C.A. § 2701 makes it illegal to “intentionally access without authorization a facility through which an electronic communication service is provided” and “thereby [obtain], [alter], or [prevent] authorized access to a wire or electronic communication.” § 2701(a).

\textsuperscript{191} Omnibus Crime Control and Safe Streets Act of 1968, 18 U.S.C.A. § 2511, outlaws the interception and disclosure of wire, oral, or electronic communications. The section provides that any person who “intentionally intercepts, . . . any wire, oral, or electronic communication . . . shall be fined under this title or imprisoned not more than five years, or both.” Id.

\textsuperscript{192} United States v. Smith, 155 F.3d 1051, 1058 (9th Cir. 1998) (discussing the complexity of both acts and determining that the best approach is one that determines their applicability based upon whether the defendant merely accesses the system or accesses and records or intercepts messages from the system).

\textsuperscript{193} Id.; § 2701(b)(1).

\textsuperscript{194} Smith, 155 F.3d at 1058; § 2511(4)(a).

\textsuperscript{195} See generally Rana Sampson, Guide 19: Misuse and Abuse of 911, PROBLEM ORIENTED GUIDES FOR POLICE 13 (U.S. Dep’t of Justice, Office of
promulgated statutes that specifically outlawed fake, false, or fraudulent 911 calls.\textsuperscript{196} For example, in Oklahoma and Georgia, false 911 calls are first-degree misdemeanors.\textsuperscript{197} Rhode Island imposes a maximum fine of $1,000 dollars and/or a maximum sentence of one-year imprisonment for knowingly making false reports to emergency services.\textsuperscript{198} The offense became a felony in Illinois in response to public outcry after a police officer’s car flipped over as he sped to the scene of a false 911 report of five dead bodies.\textsuperscript{199} However, the monetary penalties under the TICIDA are considerably higher than under state laws. Consequently, prosecutors should charge swatters under the TICIDA.

3. Harassment Prosecution Methods

Congress first enacted legislation on annoying or harassing phone calls in 1968.\textsuperscript{200} These and other federal statutes provide more stringent penalties than the TICIDA, and consequently

\textsuperscript{196} There are also non-specific statutes under which swatters may be prosecuted. Alexis Stevens, \textit{7th Graders Arrested After 911 Prank Calls}, ATLANTA J.-CONST., Mar. 24, 2010, http://www.ajc.com/news/cobb/7th-graders-arrested-after-401061.html. For example, two juveniles were charged with “transmission of false public alarm and disruption of a public school” after making prank 911 calls on school grounds. \textit{Id.} In New York, a teenager was charged with disorderly conduct after placing prank 911 calls. Ben Muessig, \textit{Teen Who Made Prank 911 Call Has History of False Reports}, GOTHAMIST (March 1, 2010, 4:57 PM), http://gothamist.com/2010/03/01/teen_who_made_prank_911_call_has_hi.php. Lastly, a Kentucky man was charged with wanton endangerment for making a false 911 calls. Lampert, \textit{supra} note 89.

\textsuperscript{197} GA. CODE ANN. § 16-11-39.2(b) (West 2010); OKLA. STAT. tit. 63, § 2819 (West 2010).

\textsuperscript{198} R.I. GEN. LAWS § 39-21.1-16 (West 2010).


\textsuperscript{200} SHARON K. BLACK, \textit{TELECOMMUNICATIONS LAW IN THE INTERNET AGE} 290 (Rick Adams ed., 2002). Indeed, many states have also passed laws criminalizing harassment via intrastate phone calls. \textit{Id.}
should be used to prosecute harassment when applicable. Title 47, § 223 of the U.S. Code prohibits “[o]bscene or harassing telephone calls” and violations thereunder are subject to a fine, imprisonment of up to two years, or both.\textsuperscript{201} Section 223(a)(1)(C) prohibits the use of “a telecommunications device, whether or not conversation or communication ensues, without [the disclosure of the caller’s] identity and with intent to annoy, abuse, threaten, or harass any person at the called number or who receives the communications.”\textsuperscript{202} In \textit{United States v. Bowker}, the defendant was found guilty of telephone harassment under § 223(a)(1)(C) after he placed multiple unwanted calls using *67 to block his caller ID information.\textsuperscript{203} The defendant argued, unsuccessfully, that even though the caller ID information was blocked, the victim could recognize his voice, which meant that his identity was not actually concealed.\textsuperscript{204} The court held that the defendant had failed to disclose his real identity when he identified himself by another name during phone calls and voice messages.\textsuperscript{205} Thus, anytime a spoofer uses the technology to harass, annoy, abuse, or threaten, and the spoofer does not state his name during the call, he should be prosecuted under this section.

Congress has also outlawed stalking via telephone calls.\textsuperscript{206} Title 18, § 2261A(2) of the U.S. Code provides that whoever uses “any facility of interstate or foreign commerce to engage in a course of conduct that causes substantial emotional distress to [his victim]” with intent to harass is punishable by up to five years

\begin{footnotesize}
\begin{itemize}
\item[202] § 223(a)(1)(C). In \textit{United States v. Popa}, the U.S. Court of Appeals for the District of Columbia defined these terms: “To annoy means to irritate, to bother, to make someone angry by repeated action; to abuse means to use insulting, coarse or bad language about or to someone; . . . and, fourth, to harass means to trouble, to worry or torment.” United States v. Popa, 187 F.3d 672, 674 (D.C. Cir. 1999).
\item[203] See generally United States v. Bowker, 372 F.3d 365 (6th Cir. 2004).
\item[204] \textit{Id}. at 390.
\item[205] \textit{Id}. (holding that a receiver’s ability to “suspect, or have a very good idea of, the caller’s identity” is irrelevant to the question of whether the defendant disclosed his identity).
\end{itemize}
\end{footnotesize}
imprisonment, a fine, or both.\textsuperscript{207} In \textit{Bowker}, the Sixth Circuit held that harassing and threatening telephone conversations and voicemail messages, coupled with the victim’s testimony that these conversations and messages made her fearful of leaving her house, were sufficient to uphold a conviction under § 2261A.\textsuperscript{208} Thus, a stalker who manipulates caller ID in order to gain access to his victim and causes her emotional distress should be prosecuted under § 2261A.

4. Political Harassment Prosecution Methods

The TICIDA should be charged in political harassment prosecutions rather than other available federal alternatives. Title 18 U.S.C. § 241 provides that “[i]f two or more persons conspire to injure, oppress, threaten, or intimidate any person . . . in the free exercise or enjoyment of any right or privilege secured to him by the Constitution or law of the [United States] . . . [t]hey shall be fined under this title or imprisoned not more than ten years, or both . . . .”\textsuperscript{209} This section requires a specific intent to violate the victim’s constitutional rights; however, this intent need not be the sole intent of the conspiracy.\textsuperscript{210} Nonetheless, it might be difficult to prove that robocalls that merely provided false information through benign language intimidated or oppressed the called party. In these circumstances, § 241 might not be effective in this context.

Alternatively, 42 U.S.C. § 1971(b) provides that “[n]o person, whether acting under color of law or otherwise, shall intimidate, threaten, coerce, or attempt to intimidate, threaten, or coerce any other person for the purpose of interfering with the right of such other person to vote or to vote as he may choose . . . .”\textsuperscript{211} Since political harassment is arguably an attempt to coerce voters, § 1971(b) seems like an easier fit in political harassment prosecutions than § 241. However, § 1971(b) does not provide for

\begin{footnotes}
\item[207] Id.
\item[208] See \textit{Bowker}, 372 F.3d at 370.
\item[210] United States v. Ellis, 595 F.2d 154, 161–62 (3d Cir. 1979).
\item[211] Civil Rights Act of 1957, 42 U.S.C.A. § 1971(b) (West 2010).
\end{footnotes}
Phoney Business

Given the deficiencies of the available alternative methods of federal prosecution, the TICIDA should be utilized to prosecute political harassers who utilize caller ID spoofing.

B. The FCC Should Inform Congress that Text Message Spoofing Will Become a Successor Technology in its Section (e)(4) Report

The TICIDA is underinclusive because it is unclear whether it also prohibits nefarious text message spoofing. Effective legislation must define “call” to include both voice and text calls. Text messages are “the most successful communications medium since e-mail,” and almost 90 percent of Americans use a cell phone. Text message spoofing can accomplish many of the same illegitimate ends as traditional caller ID spoofing, as well as other harms specific to the text message medium. If the TICIDA does not expressly cover text message spoofing, then illegitimate users might simply spoof caller ID through different means. Accordingly, the FCC should notify Congress in its section (e)(4)

---

212 § 1971(c). However, the Attorney General may seek a preventative injunction against any party for which there are reasonable grounds to anticipate a violation. Id.

213 Section (e)(4) of the TICIDA provides that the FCC “shall report to Congress whether additional legislation is necessary to prohibit the provision of inaccurate caller identification information in technologies that are successor or replacement technologies to telecommunications service or IP-enabled voice service.” Truth in Caller ID Act of 2009, 47 U.S.C.A. § 227(e)(4) (West 2010).

214 Pieter Streicher, SMS Is Not to blame, ITWEB ONLINE, July 29, 2009, http://www.itweb.co.za/index.php?option=com_content&view=article&id=24866:sms-is-not-to-blame&catid=143&Itemid=99 (“[Text] messages can be sent to more than three billion people worldwide, and in 2008, a total of six trillion [text] messages were sent globally.”).

215 Press Release, Spoofem.com, Spoofem.com Uses Mobile Media for New Marketing Method (Aug. 4, 2010) (on file with author). In addition, “52,083 text messages are sent every second,” with 83% being read within one hour of receipt. Id. Revenue from text messaging is estimated to reach $110 billion annually by 2013. Streicher, supra note 214.

report that text message spoofing will become a successor technology if it is not included under the TICIDA’s prohibitions.

As with traditional caller ID spoofing, illegitimate users exploit text message caller ID spoofing to harass or defraud others. For example, a Kansas court awarded $7.3 million in a harassment suit where the defendant sent the plaintiff and the plaintiff’s family “profane and defamatory text messages” using a spoofing service to mask her identity. When spoofer phish through text messages, it is known as “smishing.” In one common fraud, a spoofed text message informs victims that her bank account has been suspended. The message states that the victim must call an 800 number in order to unlock her account. Upon dialing this number, a message instructs the victim to enter her debit or credit card account number, pin, and expiration date. There have also been reports of spoofed text messages that fool receivers into downloading costly programs. In one scheme, a cell phone user receives a text message that reads, “Please call the hospital, it’s your mother.” However, when the cell phone user calls the number in the text message, a call is placed to a premium rate service, and the user inadvertently downloads a virus that sends

218 Id.
219 Smishing is a New Cyber Fraud That Uses SMS Messages, NORTHERN STAR (Australia), Nov. 17, 2009, at 15.
221 Krebs, supra note 220.
222 Id.
224 Id.
225 Id. Premium text messages enable cell users to purchase goods billed to
Phoney Business

premium text messages without the user’s knowledge.\footnote{Phoney Business}{226}
The TICIDA defines “caller identification service” as “information provided by a caller identification service regarding the telephone number of, or other information regarding the origination of, a call made using a telecommunications service or IP-enabled voice service.”\footnote{Phoney Business}{227} Unfortunately, the TICIDA does not define the term “call.”\footnote{Phoney Business}{228} Of course, it is possible that the FCC or the courts will interpret the TICIDA to include text calls.\footnote{Phoney Business}{229} Indeed, the FCC has interpreted the Telephone Consumer Protection Act (“TCPA”) to include text messages even though the statute did not expressly cover text messages.\footnote{Phoney Business}{230} The TCPA prohibits any call placed to “any telephone number assigned to a paging service, cellular telephone service, . . . or any service for which the called party is charged” using an automatic dialing system or an artificial or prerecorded voice.\footnote{Phoney Business}{231} The FCC determined that the term “call” under the TCPA includes voice calls and text calls, reasoning that the distinguishing factor was not the type of call placed, but whether the call was placed to a telephone number assigned to a pay service.\footnote{Phoney Business}{232} Thus, if the FCC were to apply similar logic to the TICIDA, perhaps the FCC would find that the determinative factor is whether the caller ID information was spoofed, not whether the communication was made via voice or text call.

However, federal courts are not necessarily bound to incorporate the same interpretation as the FCC.\footnote{Phoney Business}{233} In determining whether to adopt an agency’s interpretation of an ambiguous statute, a court utilizes a two-step process outlined by the Supreme Court in \textit{Chevron, Inc v. Natural Resources Defense Council},

the cell phone bill. \textit{Id.}

\footnote{Phoney Business}{226} \textit{Id.}
\footnote{Phoney Business}{228} \textit{See id.}
\footnote{Phoney Business}{230} \textit{See id.}
\footnote{Phoney Business}{231} \textit{Id.}
\footnote{Phoney Business}{232} \textit{See id.}
First, the court asks, “whether Congress has directly spoken to the precise question at issue.” Second, the court asks if the agency’s interpretation is reasonable. If these two questions are answered in the affirmative, “a court must defer to the federal agency’s interpretation of the statute . . . .” In Satterfield v. Simon & Schuster, Inc., the Ninth Circuit adopted the FCC’s interpretation of the term “call” under the TCPA. In so holding, the court weighed heavily on Congress’ delegation of authority to the FCC to implement the TCPA and the fact that “call” was undefined in the TCPA. Noting that the TCPA was enacted before the availability of text message technology, the court held that Congress could not have clearly spoken about the TCPA’s applicability to text calls. Of course, text messages were available before the TICIDA’s enactment, making Congress’ omission appear more deliberate. Accordingly, even if the FCC were to interpret “call” to include text calls, thereby utilizing its authority to implement the statute, the issue would not simply be settled; courts would nonetheless apply the Chevron analysis before deciding whether to adopt the FCC’s interpretation of the statute. This would render the legal status of text message spoofing under the TICIDA unclear once again.

Since text message spoofing can be used to accomplish that which the TICIDA clearly prohibits, the TICIDA must cover text message spoofing to prevent illegitimate spoofer's from utilizing this successor technology. Therefore, the FCC should notify Congress in its section (e)(4) report that Congress must define the term “call” to include text message calls expressly.

---

234 Id.
235 Id.
236 Id.
238 Id.
239 Id. at 953.
240 Id. at 954.
C. The FCC Should Pass Regulations to Facilitate Law Enforcement Tracing and to Create a Do-Not-Spoof List

During debates on the TICIDA and the PHONE Act, Congress members voiced concerns about law enforcement tracing and the unauthorized substitution of other individuals’ phone numbers during spoofed calls. The TICIDA does not directly address these concerns. Before the passage of the TICIDA, the FCC stated that its jurisdiction over caller ID spoofing companies was unclear. However, now that the TICIDA is law, the FCC must “pass regulations to implement [the TICIDA].” On January 26, 2011, the Department of Justice requested that the FCC promulgate rules of this nature. Moreover, the legislative history shows that Congress intended the FCC to pass regulations “imposing obligations on entities that provide caller ID spoofing services to the public.” This section argues that the FCC should promulgate regulations to address these concerns and any others necessary to maintain the viability of the industry.

1. Tracing

During congressional debates on the TICIDA and the PHONE Act, representatives expressed concern about the difficulty of

---

242 Id.
245 Letter from Lanny A. Breuer to Marlene H. Dortch, supra note 241, at 4 (“[T]he Commission should . . . allow law enforcement to trace such calls to the true originating telephone number with appropriate authority.”); id. at 3 (“The Department of Justice shares Congress’ concern about the ready availability of services that allow users to spoof telephone numbers with which they have no association whatsoever.”).
tracing spoofed calls. When correct caller ID information is reported, civilians and law enforcement agencies are able to dial *57 to implement a tracing service that stores information about the caller for the use of law enforcement. That technology does not render correct information when the number is spoofed, making the tracing process more time-consuming.

Although tracing a spoofed call might be difficult, it is possible. Often, tracking down the caller requires subpoenaing either the commercial spoofing company or the VoIP provider. Teltech reports that it “take[s] a very proactive approach to help law enforcement” when its service is used illegally. In order to help law enforcement trace spoofed phone calls, the FCC should promulgate record-keeping regulations applicable to commercial spoofing companies and VoIP providers. These regulations should specify what information must be kept by spoofing companies, the period of time such records must be kept, and the penalties that should be imposed for failure to keep such records.

2. Do-Not-Spoof List

Another concern voiced during the debates on the PHONE Act was the protection of those whose numbers are used to mask the identity of the spoofer. Often a spoofer will substitute the same

---

247 Sabin Statement, supra note 61 (discussing the difficulty police encountered when trying to locate the source of threatening spoofed phone calls made to a police officer and his family and expressing concern that spoofing could “complicate criminal investigations”).


249 Id.

250 See id.

251 Id.

252 Meyer, supra note 217.

number repeatedly.\textsuperscript{254} This happened to Phil Kiko, Chief of Staff and General Counsel to the House of Representatives, who received upwards of twenty phone calls per day from individuals who believed that he was repeatedly calling them.\textsuperscript{255} Instead, a spoofer had placed those calls spoofing Mr. Kiko’s name.\textsuperscript{256} Currently, those whose numbers are used to spoof have few options but to change their phone numbers.\textsuperscript{257}

Subsection two of the PHONE Act would have criminalized the use of “caller ID information pertaining to an actual person or other entity without that person’s or entity’s consent and with intent to deceive any person or other entity about the identity of the caller” in an apparent attempt to minimize unauthorized use of numbers.\textsuperscript{258} However, it is too demanding to require that everyday users of caller ID spoofing technology determine whether a phone number belongs to another individual. Any such requirement would likely have negative effects on legitimate users, such as domestic violence victims, who might chose a random string of numbers without realizing it is in fact another’s phone number. In addition, the threat of criminal sanction for failure to get approval to spoof a number will force these victims to use their relatives’ or friends’ phone numbers, which might reveal too much about their locations. Thus, the providers of caller ID spoofing should assume the responsibility for determining whether the use of a phone number is appropriate.

On January 26, 2011, the Department of Justice suggested that the FCC “should consider the feasibility of requiring public providers of caller ID spoofing services to make a good-faith effort to verify that a user has the authority to use the substituted number, such as by placing a one-time verification call to that number.”\textsuperscript{259} However, such a requirement is far too burdensome on the

\textsuperscript{254} \textit{Id.}

\textsuperscript{255} \textit{Id.}

\textsuperscript{256} \textit{Id.}

\textsuperscript{257} \textit{Id.}


\textsuperscript{259} Letter from Lanny A. Breuer, supra note 241, at 3.
industry. Just one public provider of caller ID spoofing services reported having over three million customers and these customers might have substituted more than one number each. Clearly, implementing such a system, which placed a phone call to every proposed substitute number, would require a massive undertaking by public providers. Moreover, the system would be inefficient, as repetitive calls would surely occur.

Instead, the FCC should promulgate regulations mandating that commercial spoofing companies maintain a shared Do-Not-Spoof list. Unlike the Department of Justice’s proposed method, this list would prevent spoofers from substituting the listed number or calling listed numbers with spoofed caller ID information. Additionally, all caller ID companies would share the price of maintaining such a list, preventing the duplicate costs incurred under the Department of Justice’s proposed method.

Spoofing companies already maintain lists that operate in a similar manner. SpoofCard reports that it maintains a list of numbers that spoofers cannot call. It created this list in an effort to prevent swatting and SpoofCard continually works to increase the list of police numbers that are not spoofable. In addition, Spoofem.com now offers SpoofAbuse, where for five dollars one can provide up to three numbers which the company will put on its do not spoof list, so its customers will no longer be able to make spoofed calls to that number. Spoofem.com also provides these numbers to other commercial spoofing companies; however, it does not guarantee that other companies will respect the subscriber’s request. Thus, the technology for a Do-Not-Spoof list exists, but could be used more efficiently.

Any Do-Not-Spoof list should include emergency and government numbers so customers cannot engage in swatting and so phishers cannot spoof these numbers to trick others. Individuals

---

260 Thomas, supra note 29 (noting that the number one spoofing company, SpoofCard, has over three million customers).
261 Meyer, supra note 217.
262 Id.
264 Id.
should be able to contact spoofing providers directly and the FCC should determine a method for phone service providers to report their customers’ requests to be added to the list. 265 Lastly, the FCC should determine the appropriate sanctions to impose when companies fail to honor individuals’ requests that their numbers not be used for spoofing.

V. CONCLUSION

Caller ID spoofing provides real societal benefits, but can also deliver dangerous blows. 266 The trust many place in their caller ID service gives spoofer s an advantage when they commit crimes, but spoofing also provides necessary shelter to many legitimate users. 267 The Truth in Caller ID Act of 2009 is a step in the right direction; however, additional federal response is necessary. The Department of Justice should utilize its full array of options to prosecute creatively when appropriate, so that the charge conveys the gravity of the crime. In addition, the TICI DA must clearly include text message caller ID spoofing because it is used to accomplish the same illegitimate ends as traditional caller ID spoofing. Lastly, the FCC should promulgate regulations to help law enforcement trace illegitimate users and create a Do-Not-Spoof list. With a comprehensive approach, Congress will be able to ensure the legality of caller ID spoofing technology for legitimate users, while also minimizing illegitimate uses. Victims of caller ID spoofing, like Doug Bates, should feel safe in their homes again.

265 Monteith Statement, supra note 51; see also Kiko Statement, supra note 253, at 23 (lamenting that his telephone provider informed him that he could not prevent spoofing accomplished with his phone number).
266 See supra Part II.B (explaining legitimate and illegitimate uses of caller ID spoofing).
A PRESCRIPTION FOR CHANGE: CITIZENS UNITED’S IMPLICATIONS FOR REGULATION OF OFF-LABEL PROMOTION OF PRESCRIPTION PHARMACEUTICALS

Kristie LaSalle*

I. INTRODUCTION

Al Caronia, a pharmaceutical sales representative for Orphan Pharmaceuticals, a subsidiary of Jazz Pharmaceuticals, walked into a physician’s office in Great Neck, New York on November 2, 2005.1 He had received several phone calls from the physician asking about Xyrem,2 Orphan’s medication for the treatment of

* J.D. Candidate, Brooklyn Law School, 2012; B.A. Biology, Swarthmore College, 2006. I owe an immeasurable debt of gratitude to Professor Aliza Kaplan, not only for teaching me to write, but for her tireless support and encouragement, and countless hours spent assisting in the writing and editing process. This Note could not have been possible without her love and support. I would like to thank Professors William Araiza and Joel Gora; Paul V. Avelar, James Beck, and Adam Michaels, Esqs.; and Robert Sobelman, J.D. Candidate, Brooklyn Law School, 2012; for their support, guidance, and suggestions during the writing process. I would also like to thank the entire editorial staff of the Journal of Law and Policy for their editorial assistance. Finally, thank you to my family, Richard, Cathy, Kimberlee, and Cadie for their support in all of my endeavors.

1 Brief for Petitioner at 4, United States v. Caronia, No. 09 Cr. 5006 (2d Cir. filed Apr. 15, 2010) [hereinafter Caronia Brief]; Appendix to Caronia Brief at A-32 (transcript of conversation, indicating physicians’ office address).

2 Id. at 3. This physician was an internal medicine specialist who, unbeknownst to Mr. Caronia, had never prescribed a single narcoleptic. Id. Prior to receiving phone calls from the physician, Mr. Caronia had never met him; however, after receiving repeated phone calls, Mr. Caronia decided to add the physician to the list of doctors whom he visited, and met with him for the first time on October 26, 2005. Id. At the October 26 meeting, the physician asked
cataplexy,\(^3\) a muscle disorder associated with narcolepsy,\(^4\) so Mr. Caronia brought Dr. Peter Gleason with him to discuss his experience prescribing the medicine.\(^5\) Mr. Caronia sat silently while Dr. Gleason spoke with the physician, answering questions about Xyrem’s different uses, including the treatment of Excessive Daytime Sleepiness (“EDS”), a new use then under review by the Food and Drug Administration (“FDA”).\(^6\) So promising was that treatment, in fact, that the FDA approved Xyrem for treatment of EDS on November 18, 2005,\(^7\) just sixteen days after Mr. Caronia and Dr. Gleason visited the physician’s office. A year and a half later, however, on July 25, 2007, a federal grand jury indicted Mr. Caronia and Dr. Gleason for illegally promoting a prescription pharmaceutical.\(^8\) The physician, it turned out, was a confidential informant for the FDA, and the Federal Bureau of Investigation (“FBI”) had recorded his meetings with Mr. Caronia and Dr. Gleason.\(^9\) Prior to trial, Dr. Gleason pleaded guilty to a reduced charge, becoming the first physician convicted of misbranding a drug.\(^10\) Mr. Caronia was found guilty of criminally misbranding a

---

\(^3\) XYREM PRODUCT LABEL at 6 (July 17, 2002), available at http://www.accessdata.fda.gov/drugsatfda_docs/label/2002/21196lbl.pdf. Xyrem is a sleep-inducing depressant.

\(^4\) Nat’l Inst. of Neurological Disorders and Stroke, Narcolepsy Fact Sheet, NAT’L INSTS. HEALTH (last updated May 14, 2010), http://www.ninds.nih.gov/disorders/narcolepsy/detail_narcolepsy.htm. Cataplexy is a debilitating symptom of narcolepsy, triggered by strong emotions like anger or fear. It manifests as a loss in muscle control, with the most severe cases resulting in physical collapse and the inability to move, speak, or open one’s eyes. Id.


\(^6\) Caronia Brief, supra note 1, at 4; Appendix to Caronia Brief, supra note 1, at A-41 (transcript of conversation, discussing the use of Xyrem for EDS).

\(^7\) Letter from Russell Katz, Dir., Div. of Neurology Prods., Ctr. for Drug Eval., FDA, to Dr. Reardon, Orphan Med., Inc. (Nov. 18, 2005).


\(^9\) Appendix to Caronia Brief, supra note 1, at A-32.

\(^10\) Judge Strikes Down First Amendment Arguments In Pharmaceutical Sales Representative Federal Trial, FDA ADVER. & PROMOTION MANUAL &
prescription pharmaceutical for promoting Xyrem for EDS,\footnote{Appendix to Caronia Brief, \textit{supra} note 1, at A-85 to A-86 (verdict sheet).} even though the information related to the efficacy of Xyrem in treating EDS was accurate and non-misleading, as evidenced by the FDA’s approval of such a use a mere two weeks later. Why?

scheme surrounding off-label promotion\textsuperscript{15} is murky at best, and leaves manufacturers little absolute guidance.\textsuperscript{16} Policy is delineated largely by the FDA’s own interpretation of its regulatory power, disseminated piecemeal in guidance documents, and often driven by the FDA’s response to private litigation or other emerging issues.\textsuperscript{17}

Until recently, a similar regulatory scheme existed in a seemingly unrelated field of free speech jurisprudence: corporate political speech.\textsuperscript{18} Few recent Supreme Court decisions have garnered as much criticism as the 2010 decision in \textit{Citizens United v. FEC},\textsuperscript{19} in which the Supreme Court held that the Federal

\textsuperscript{15} See 21 U.S.C.A. §§ 301–99; see generally infra Part II.B.

\textsuperscript{16} See generally, e.g., \textsc{Food and Drug Admin.}, \textsc{Good Reprint Practices for the Distribution of Medical Journal Articles and Medical or Scientific Reference Publications on Unapproved New Uses of Approved Drugs and Approved or Cleared Medical Devices}, (Jan. 2009) [hereinafter \textsc{Good Reprint Practices}] available at http://www.fda.gov/oc/op/goodreprint.html.

\textsuperscript{17} For example, in one challenge to the FDA’s prohibition of off-label promotion, the FDA stipulated that no previous guidance document “independently authorizes the FDA to prohibit or sanction speech,” rendering moot an injunction against prohibiting certain forms of off-label promotion imposed at the trial court level. Wash. Legal Found. \textsc{v. Henney (WLF III)}, 202 F.3d 331, 335 (D.C. Cir. 2000). The FDA then promptly issued a guidance document, interpreting the Circuit Court’s ruling as a holding that the “FDA, consistent with its longstanding interpretation of the laws it administers, may proceed, in the context of case-by-case enforcement, to determine from a manufacturer’s written materials and activities how it intends that its products be used.” Notice; Decision in \textsc{Washington Legal Foundation v. Henney}, 65 Fed. Reg. 14,286 (Mar. 16, 2000) [hereinafter \textsc{WLF III Decision Notice}].

\textsuperscript{18} See generally \textsc{Citizens United v. F.E.C.}, 130 S. Ct. 876 (2010).

\textsuperscript{19} Id.; see, e.g., Ronald Dworkin, \textit{The “Devastating” Decision}, \textsc{N.Y. Rev. Books}, Feb. 25, 2010, at 39 (reporting the \textsc{Citizens United} decision arose from the Court’s political preference for the Republican Party and/or its general favoritism of corporate interests); Molly J. Walker Wilson, \textit{Too Much of a Good Thing: Campaign Speech After \textsc{Citizens United}}, 31 \textsc{Cardozo L. Rev.} 2365, 2368–69 (2010); Adam Liptak, \textit{Justices, 5-4 Reject Corporate Spending Limit}, \textsc{N.Y. Times}, Jan. 21, 2010, at A1. Perhaps the most well-known criticism of the holding came from President Obama. Barack Obama, Remarks by the President in the State of the Union Address (Jan. 27, 2010) (“[L]ast week, the Supreme Court reversed a century of law that I believe will open the floodgates for
Elections Commission’s (“FEC”) regulation of corporate political expenditures was unconstitutional.\textsuperscript{20} Much of the criticism focused on the case’s implications for campaign finance;\textsuperscript{21} however, the Court’s language and reasoning\textsuperscript{22} may have dramatic implications for other areas of speech. In particular, \textit{Citizens United} may signal a dramatic increase in the rights of those engaged in heavily-regulated—and, to this point, heavily suppressed—commercial speech.\textsuperscript{23} One area of commercial speech jurisprudence that may change after \textit{Citizens United} is the off-label promotion of pharmaceutical products.

The \textit{Citizens United} holding portrays an inherent mistrust of the use of convoluted regulatory schemes to abridge speech. The Court held that a muddled regulatory scheme essentially functions as a prior restraint: it leaves a potential speaker with the constitutionally unpalatable choice between remaining silent or asking the government whether its speech is lawful.\textsuperscript{24} Justice Kennedy’s declaration that prohibition of corporate independent expenditures constituted an “unprecedented governmental intervention into the realm of speech,”\textsuperscript{25} however, is misplaced: a similar intrusion on speech rights exists in pharmaceutical marketing regulation.\textsuperscript{26} The FDA’s prohibition of off-label promotion applies regardless of whether the marketing message the manufacturer wishes to disseminate is truthful and non-special interests—including foreign companies—to spend without limit in our elections . . . . I don’t think American elections should be bankrolled by America’s most powerful interests, or worse, by foreign entities.

\textsuperscript{20} \textit{Citizens United}, 130 S. Ct. at 913.
\textsuperscript{21} See generally Wilson, \textit{supra} note 19. This Note takes no normative position on the wisdom or failings of \textit{Citizens United} as applied to political speech.
\textsuperscript{22} See \textit{Citizens United}, 130 S. Ct. at 896 (finding that muddled regulatory schemes function as a prior restraint on speech).
\textsuperscript{24} \textit{Citizens United}, 130 S. Ct. at 896.
\textsuperscript{25} \textit{Id}.
misleading. The FDA’s transgressions upon the manufacturers’ speech rights bear many of the marks condemned in *Citizens United*: the regulation constitutes “an outright ban [on speech], backed by criminal sanctions”; the FDA regulates speech by “carving out a limited exemption through an amorphous regulatory interpretation”; and in doing so, the FDA paternalistically “select[s] what . . . speech is safe for public consumption by applying an ambiguous test.” Thus, under *Citizens United*, the FDA’s prohibition of off-label promotion constitutes an unconstitutional abridgement of commercial speech.

Fortunately, the similarities between the infirmities of the FEC’s regulation of independent corporate political expenditures at issue in *Citizens United* and those of the FDA’s regulation of off-label promotion suggest that the solutions posed in *Citizens United* to enable continuing regulation while respecting corporations’ political speech rights will also cure the constitutional failings of the FDA’s off-label regulation. When properly construed, the government’s interest in promoting the public health and safety may be adequately addressed by requiring the speaker—that is, the pharmaceutical manufacturer—to disclose the off-label nature of its promotions.

At the outset, clarification is necessary as to the limited form of off-label promotion for which this Note advocates greater First Amendment protection. That limited category of speech is (1) truthful, accurate, and not misleading; (2) directed only to other

---


28 *Citizens United*, 130 S. Ct. at 897. See also 21 U.S.C.A. § 333 (describing criminal and civil sanctions that attend misbranding).

29 See *Citizens United*, 130 S. Ct. at 889.

30 See id. at 896.

31 See id. at 915.

32 See id.

33 Fraudulent or inherently misleading commercial speech receives no First
prescribing healthcare professionals;\textsuperscript{34} and (3) relates to off-label uses of FDA-approved drugs only. Such speech would benefit the public at large, were it to be permitted.\textsuperscript{35} Physicians may prescribe a medication for any use, including off-label uses.\textsuperscript{36} Indeed, the FDA itself recognizes that off-label prescribing may be safe, and sometimes the standard of care for a particular patient or disease.\textsuperscript{37} Providing the proper constitutional protection for truthful, non-misleading off-label promotion would facilitate the dissemination of medical information to physicians. As explained below, public health will be enhanced by dissemination of such information, resulting in better-informed prescribing decisions by physicians.\textsuperscript{38}

This Note addresses the First Amendment concerns implicated by the FDA’s policies prohibiting off-label promotion of prescription medications, and presents a potential solution. Part II of this Note surveys the current state of the constitutional, statutory, and regulatory frameworks, as well as the litigation landscape, surrounding off-label promotion. Part III demonstrates that \textit{Citizens United}’s rationale extends beyond political speech to heavily-regulated commercial speech. Specifically, the Court’s concern that the FEC’s complex and “amorphous regulatory interpretation”\textsuperscript{39} constitutes a restriction so inscrutable that it functions as a prior restraint finds a ready parallel in the FDA’s regulation of off-label promotion.

Finally, in applying the Court’s commercial speech analysis

---

\textsuperscript{34} This Note does not discuss the constitutional implications of the FDA’s regulation of Direct to Consumer (DTC) marketing.

\textsuperscript{35} \textit{See infra} Part IV.B.

\textsuperscript{36} \textit{Use of Approved Drugs for Unlabeled Indications}, 12 FDA DRUG BULL. 4 (FDA, Washington, D.C., Apr. 1982).

\textsuperscript{37} \textit{Good Reprint Practices}, supra note 16, § III.

\textsuperscript{38} \textit{See infra} Part IV.B.

and the principles informing *Citizens United*, this Note concludes in Part IV that the FDA’s current policies regarding off-label promotion are unconstitutional under the First Amendment. A commercial speech analysis, the test for which was set forth in *Central Hudson Gas and Electric Corporation v. Public Service Commission of New York*, reveals that the constitutional failing is twofold. First, the FDA, and courts who have previously considered the constitutionality of the FDA’s prohibition on speech, have misconstrued the purported government interest: rather, when properly construed, the interest is not advanced by—and in fact may be hindered by—the prohibition of off-label promotion. Second, the current policy is more restrictive than necessary to serve that interest. Using *United States v. Caronia* as an example, this Note advocates the adoption of a disclosure regime, as approved of in *Citizens United*, which would better serve public health while being less restrictive of speech.

II. COMMERCIAL SPEECH CHILLED: THE STATUTORY, REGULATORY, AND LEGAL LANDSCAPE OF OFF-LABEL PROMOTION

Regardless of whether it is classified as pure, scientific, speech or commercial speech, off-label promotion of prescription pharmaceutical products is entitled to some First Amendment

---


41 See United States v. Caputo, 288 F. Supp. 2d 912, 921 (N.D. Ill. 2003) (announcing as a substantial government interest “the government’s interest in subjecting off-label uses to the FDA’s evaluation process as well as the government’s interest in preserving the integrity of the [Food Drug and Cosmetic Act’s] new drug approval process”); see also, generally, infra Part IV.A.


43 Although beyond the scope of this Note, there is good reason to question whether off-label promotion is truly commercial speech or whether it is in fact scientific, and therefore pure, speech entitled to the highest constitutional protection. See generally, e.g., Glenn C. Smith, *Avoiding Awkward Alchemy—In the Off Label Drug Context and Beyond: Fully-Protected Independent Research Should Not Transmogrify into Mere Commercial Speech Just Because Product Manufacturers Distribute It*, 34 WAKE FORREST L. REV. 963 (1999).
protection. Yet under the current regulatory regime created by the FDA, off-label promotion is all but prohibited. Challenges to the constitutionality of this absolute prohibition have thus far done little to vindicate manufacturers’ First Amendment rights.

A. The Commercial Speech Doctrine

Despite the constitutional mandate that “Congress shall make no law . . . abridging the freedom of speech,” commercial speech, or speech which “does no more than propose a commercial transaction,” has historically received diminished protection under the First Amendment. Until the 1970s, commercial speech was entitled to no constitutional protection whatsoever. In 1976, the Supreme Court afforded limited constitutional protection to commercial speech. In *Virginia Board of Pharmacy v. Virginia Citizens Consumer Council*, the Court struck down a state ban on

---

44 See Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, 425 U.S. 748, 761 (1976); see generally *Central Hudson*, 447 U.S. 557 (extending, for the first time, qualified First Amendment protection to commercial speech); see generally infra Part II.A.


47 U.S. CONST. amend. I.


49 See, e.g., Valentine v. Chrestensen, 316 U.S. 52, 54 (1942) (finding “the Constitution imposes no . . . restraint on government as respects purely commercial advertising”).

50 *Va. State Bd. of Pharmacy*, 425 U.S. at 773 (“What is at issue is whether a State may completely suppress the dissemination of concededly truthful [commercial] information about entirely lawful activity, fearful of that information’s effect upon its disseminators and its recipients . . . [W]e conclude that the answer to this [question] is in the negative.”).
advertising of drug prices by pharmacies.\textsuperscript{51} Eschewing the paternalistic argument that the petitioner had an interest in protecting “unwitting customers” from opting for “low-cost, low-quality service,”\textsuperscript{52} the Court noted that the consumer, seller, and society at large had an interest in the “free flow of commercial information” that “may be as keen, if not keener by far, than [the] interest in the day’s most urgent political debate.”\textsuperscript{53}

On the heels of \textit{Virginia Board of Pharmacy}, the Supreme Court developed a test for determining whether commercial speech received First Amendment protection.\textsuperscript{54} In \textit{Central Hudson Gas \& Electric Corporation v. Public Service Commission of New York}, the Court struck down a ban on advertisements by electric utility services,\textsuperscript{55} in part because respondent’s reasons for the ban reflected its fear that the advertisements would needlessly increase energy consumption and costs.\textsuperscript{56} Echoing its sentiments in \textit{Virginia Board of Pharmacy}, the Court rejected the government’s inherently paternalistic suppression of commercial speech.\textsuperscript{57} The Court set forth a four-prong test for assessing the constitutionality of governmental restrictions on commercial speech: the speech (1) “must concern lawful activity and not be misleading.”\textsuperscript{58} However, the government may ban truthful, non-misleading speech if: (2) the asserted governmental interest in suppressing the

\textsuperscript{51} \textit{Id.} at 762.  
\textsuperscript{52} \textit{Id.} at 769.  
\textsuperscript{53} \textit{Id.} at 763.  
\textsuperscript{55} \textit{Id.} at 561.  
\textsuperscript{56} \textit{Id.} at 560–61.  
\textsuperscript{57} \textit{See, e.g., id.} at 562; \textit{see also id.} at 574–75 (Brennan, J., concurring) (noting that the restriction at issue was “a covert attempt [by the state] to manipulate the choices of its citizens, not by persuasion or direct regulation, but by depriving the public of the information needed to make a free choice. . . . If the First Amendment guarantee means anything, it means that, absent clear and present danger, government has no power to restrict expression because of the effect its message is likely to have on the public”).  
\textsuperscript{58} \textit{See id.} at 566.  
\textsuperscript{59} \textit{Id.}
speech is substantial; (3) the regulation directly advances the substantial governmental interest; and (4) the restriction on speech is not “more extensive than is necessary to serve that interest.”

The *Central Hudson* test was refined and fortified in *44 Liquormart v. Rhode Island*, in which an unanimous Court struck down a Rhode Island prohibition on advertising liquor prices because the regulation failed to satisfy the fourth prong of the *Central Hudson* test. Although the Court was divided in its reasoning, Justice Thomas, concurring, expressed clear disapprobation for the regulation’s attempts to control consumer choice through the suppression of truthful, non-misleading commercial speech; he argued that if paternalism provides the sole support for a challenged regulation, application of the *Central Hudson* test should be unnecessary, as the law is “per se illegitimate.” It is clear that “the Supreme Court looks askance at restrictions on commercial speech imposed for paternalistic purposes.”

**B. Statutory and Regulatory Framework of Off-Label Promotion**

The FDA exercises authority over the production, sale, and marketing of pharmaceutical products, pursuant to the Federal

---

60 Id.
61 Id.
63 Id. at 518 (Thomas, J., concurring) (“In cases such as this, in which the government’s asserted interest is to keep legal users of a product or service ignorant in order to manipulate their choices in the marketplace, the balancing test adopted in [Central Hudson] should not be applied, in my view. Rather, such an ‘interest’ is per se illegitimate and can no more justify regulation of ‘commercial’ speech than it can justify regulation of ‘noncommercial’ speech.”).
Food, Drug, and Cosmetic Act (“FDCA”). The FDCA mandates that all new drugs must be approved by the FDA as safe and effective, based on extensive clinical and pre-clinical testing for each intended use before introduction to the market.

The FDCA does not, however, directly constrain physicians’ decisions to prescribe drugs off-label. Congress has not endowed the FDA with the authority to prohibit physicians from prescribing medications off-label. Rather, the FDA has recognized that Congress did not intend FDA to interfere with the practice of medicine. Thus, once a product is approved for marketing for a specific use, FDA generally does not regulate how, or for what uses, physicians prescribe it. A licensed physician may prescribe a drug for other uses, or in treatments, regimens, or patient populations, that are not listed on the FDA-approved labeling.

Nevertheless, by manipulating the information disseminated to physicians, the FDA exerts substantial control over their prescribing habits.

---

66 See id. §§ 321, 355, 360m.
67 See id. § 396 (reflecting Congress’s amendment of the FDCA to clarify that “[n]othing in this chapter shall be construed to limit or interfere with the authority of a health care practitioner to prescribe or administer any legally marketed [drug] to a patient for any condition or disease within a legitimate health care practitioner-patient relationship”).
68 See id. This particular provision of FDAMA was enacted in response to the FDA’s assertion that “The [FDCA] provides [the] FDA with explicit regulatory authority over the use of [drugs].” James M. Beck & Elizabeth D. Azari, FDA, Off-Label Use, and Informed Consent: Debunking Myths and Misconceptions, 53 Food & Drug L.J. 71, 78 (1998) (quoting Attachment to Letter from FDA to Hon. Joseph Barton, Chairman, Subcomm. on Oversight and Investigation, House Comm. on Commerce (Apr. 14, 1995)).
69 Beck & Azari, supra note 68, at 78 n.64 (quoting Michael Friedman, Deputy Comm’r for Operations, FDA, Prepared Statement Before the Subcommittee on Human Resources and Intergovernmental Relations of the House Committee on Government Reform and Oversight (Sept. 12, 1996)).
A Prescription for Change

Within the drug approval process, the FDA retains authority over the labeling of the drug. A drug’s label must detail its risks, benefits, and adequate instructions for use. The FDA only approves a drug for sale if its labeling conforms to the FDA’s specifications. “Labeling” is a term of art, which, in addition to physical labels affixed to the drug’s packaging, encompasses all written, printed, or graphic material “accompanying” the drug. “Accompanying” includes not only materials physically associated with the distribution of the drug itself, but all materials supplementing or explaining the product. According to the FDA’s own interpretation of the scope of its power, the term “labeling” encompasses nearly every form of manufacturer communication. The FDA has also adopted an expansive definition of promotion to include nearly every interpersonal contact between representatives of pharmaceutical manufacturers and prescribers or the public at large.

It is unlawful for the manufacturer to introduce a “misbranded” or “adulterated” drug into interstate commerce. A drug is on physicians’ prescribing habits, its ability to limit the dissemination of information regarding off-label uses to prescribing physicians represents a significant indirect control over how physicians prescribe medications.”).

72 See, e.g., id.
73 Id. §§ 331(a), (d).
74 Id. §§ 321(k–m).
76 21 C.F.R. § 202.1(l)(1–2) (2010) (defining “labeling” as “[b]rochures, booklets, mailing pieces, detailing pieces, file cards, bulletins, calendars, price lists, catalogs, house organs, letters, motion picture films, film strips, lantern slides, sound recordings, exhibits, literature, and reprints, and similar pieces of printed, audio, or visual matter descriptive of a drug and references published (for example, the Physicians’ Desk Reference)” and promotional materials).
77 See Joseph Leghorn, Elizabeth Brophy & Peter V. Rother, The First Amendment and FDA Restrictions on Off-Label Uses: The Call for a New Approach, 63 FOOD & DRUG L.J. 391, 394 (2008) (identifying “certain company-supported scientific or educational activities,” “initiation of person-to-person contact between sales representatives and prescribers,” “direct-to-consumer advertisements,” and “improper dissemination of information about an investigational drug during a clinical trial” as promotional activities).
misbranded or adulterated if its labeling is “false or misleading in any particular” or if it includes information regarding a use not approved by the FDA. Promotion of a drug in a manner inconsistent with its labeling results in the drug being “misbranded.” Because the FDA approves drugs for specific uses, promotional activity that covers uses which are not included in the label, or “off-label” promotion, is a form of misbranding and is prohibited. The FDA is empowered to enjoin manufacturers who promote drugs for off-label uses, to seize those drugs, and in some cases, to seek criminal sanctions.

The prohibition is absolute, banning not only fraudulent and untruthful marketing activity, but truthful, non-misleading information as well. The FDA’s sweeping prohibition of off-label promotion has drawn significant criticism on constitutional grounds. First, it constitutes a restriction on who may speak,
A Prescription for Change

because physicians may freely discuss amongst themselves off-label uses, and may prescribe a medication for any purpose, yet a manufacturer may not. Second, the restriction on commercial speech is more expansive than necessary to ensure public health and safety.

C. Legal Challenges to the FDA’s Prohibition of Off-Label Promotion

In the past two decades, several legal challenges to the FDA’s regulatory scheme have been brought in federal court. These lawsuits, while challenging various provisions within the FDA’s regulatory framework, all attack the prohibition of off-label promotion as an impermissible abridgement of commercial speech in violation of the First Amendment. Constitutional challenges to this commercial speech restriction have met with mixed results,


89 See infra Part III.B.1.

90 Viagra, for example, is best known for its use in treating erectile dysfunction (ED), but, it was originally approved for use in treating angina—chest pain associated with cardiac disease. Its discovery and success as an ED drug hails from what was originally an off-label use. See Fritch, supra note 70, at 319 n.9 (citations omitted). Efficacy has also been demonstrated for other off-label uses, including treating premature babies with under-developed lungs. See id. Another example can be found in verapamil, a calcium-channel blocker FDA-approved for treatment of heart disease. Physicians find them to be efficacious for treating headaches, including migraines, yet the FDA has not approved them for treatment of headaches. See Daniel B. Klein & Alexander Tabarrok, Do Off-Label Drug Practices Argue Against FDA Efficacy Requirements?, 67 AM. J. ECON. & SOCIOL. 743, 756 (2008).


92 See infra Part III.D.2.


94 See supra note 93.
leaving the status of the FDA’s prohibition of off-label promotion very much unsettled.

1. Washington Legal Foundation

The first challenge to the FDA’s policies came in 1997, and concerned two guidance documents restricting (1) dissemination of peer-reviewed scientific journal articles and textbooks by manufacturers and (2) manufacturer-sponsored continuing medical education programs (“CMEs”). The Washington Legal Foundation (“WLF”) sought to enjoin the FDA from enforcing these policies on free speech grounds. After determining that off-label promotion constituted commercial speech, the district court applied the Central Hudson test. The court found that the speech at issue was truthful and not inherently misleading. Although the court recognized that the government had a substantial interest in protecting public health and safety by incentivizing manufacturers to seek FDA approval for off-label uses, which was directly advanced by the challenged regulations, it nevertheless declared the FDA’s policies unconstitutional because it found the restriction on speech to be more extensive than necessary. The court found that there were less restrictive alternatives available, including a

---


96 The Washington Legal Foundation is “a nonprofit public interest law and policy center that defends ‘the rights of individuals and businesses to go about their affairs without undue influence from government regulators.’” WLF I, 13 F. Supp. 2d at 54 (internal citations omitted); Wash. Legal Found., WLF Mission, WASH. LEGAL FOUND., http://www.wlf.org/org/mission.asp (last visited Apr. 8, 2011).

97 WLF I, 13 F. Supp. 2d at 54.


100 Id. at 69–71.

101 Id.

102 Id. at 72–74.
“full, complete, and unambiguous disclosure by the manufacturer” that the advocated use is off-label and not FDA approved. Faced with such disclosure, “[a] physician would be immediately alerted to the fact that the ‘substantial evidence standard’ had not been satisfied, and would evaluate the communicated message accordingly.”

After the decision in *WLF I*, the Food and Drug Administration Modernization Act of 1997 (“FDAMA”) took effect. Within FDAMA, the very policies declared unconstitutional in *WLF I* were codified in section 401. In a separate decision, *WLF II*, the district court *sua sponte* extended its *WLF I* ruling to apply to section 401 of FDAMA, declared that section unconstitutional, and enjoined the FDA from enforcing it.

On appeal, the FDA disclaimed its original argument that FDAMA authorized the FDA to regulate or prohibit speech, stating instead that “in its view, neither the FDAMA nor the CME guidance independently authorizes the FDA to prohibit or sanction speech.” At oral argument the FDA elected to interpret section 401 as providing a “safe harbor,” that ensures that certain types of conduct would not be used against manufacturers in criminal actions for misbranding. It insisted, however, that FDAMA permitted criminal sanctions against any manufacturer who completely disregarded section 401’s conditions on off-label promotion “provides that a manufacturer who disregards [section...
401]’s conditions . . . might be liable in some fashion.”¹¹² The FDA’s new interpretation brought its position in line with WLF’s, eliminating the controversy.¹¹³ The D.C. Circuit lamented the lost opportunity to resolve “a difficult constitutional question of considerable practical importance” and noted that, because of the FDA’s concession, “the dispute between the parties has disappeared before our eyes.”¹¹⁴ It vacated the WLF I and WLF II rulings for want of controversy.¹¹⁵ In the wake of WLF III, the FDA interpreted the Circuit’s ruling to mean that the “FDA, consistent with its longstanding interpretation of the laws it administers, may proceed, on a case-by-case basis, to determine from a manufacturer’s written materials and activities how it intends that its products be used.”¹¹⁶ Although the WLF decisions addressed only a narrow subset of off-label speech—distribution of scientific literature and sponsorship of CMEs¹¹⁷—the district court’s vacated opinions influenced subsequent decisions addressing the constitutionality of restrictions on off-label promotion.¹¹⁸

2. United States v. Caputo

In 2003, the FDA’s prohibition of off-label promotion faced another constitutional challenge.¹¹⁹ In United States v. Caputo, the defendants challenged the FDA’s prohibition of off-label promotion of a medical device.¹²⁰ The district court largely
A Prescription for Change

adopted *WLF I*’s reasoning, but departed from *WLF I* in holding that the prohibition was constitutional. In addressing the FDA’s interest in restricting speech, the court reasoned that “[p]ermitting manufacturers to promote off-label uses would completely undermine the government’s interest in subjecting off-label uses to the FDA’s evaluation process as well as the government’s interest in preserving the FDCA’s new drug approval process.”

Distinguishing the promotion at issue in *Caputo* from the dissemination of academic writing and sponsorship of CMEs in *WLF I*, *Caputo* found prohibition of all off-label promotion—even that which is truthful and non-misleading—to be constitutional because “permitting Defendants to engage in all forms of truthful, non-misleading promotion of off-label use would severely frustrate the FDA’s ability to evaluate the effectiveness of off-label uses.” Because the court could not envision a less restrictive means of achieving this interest, it preserved the FDA’s authority to prohibit off-label commercial speech. On appeal, the Seventh Circuit did not resolve the First Amendment quandary, and declared that it “[f]ortunately . . . need not decide today whether a seller of drugs . . . has a constitutional right to promote off label uses,” affirming *Caputo* on more narrow grounds.

3. Allergan

More recently, the FDA commenced a barrage of civil and criminal sanctions against the manufacturer Allergan, Inc. for alleged off-label marketing of its pharmaceutical product, Botox. Allergan fought back, filing a civil suit challenging the...
constitutionality of the prohibition on First Amendment grounds.\textsuperscript{128} With the threat of massive civil and criminal penalties looming, Allergan was permitted to plead guilty and settle the civil complaints, conditioned upon abandonment of its First Amendment suit.\textsuperscript{129} The settlement left the question of the constitutionality of the FDA’s prohibitions unanswered.

4. Caronia’s First Amendment Challenge

The FDA investigated alleged off-label promotion of Xyrem by Jazz Pharmaceuticals, resulting in the indictment of Caronia on two misdemeanor counts:\textsuperscript{130} (1) conspiracy to misbrand Xyrem by marketing it to the undercover informant-physician for off-label uses;\textsuperscript{131} and (2) misbranding a drug held for sale in interstate commerce.\textsuperscript{132} Caronia moved to dismiss the charges because, inter
"A Prescription for Change"

alia, the misbranding provisions violated his free speech rights.\textsuperscript{133} The court found that the off-label promotion constituted commercial speech,\textsuperscript{134} and that the provisions did not violate the First Amendment.\textsuperscript{135}

Under \textit{Central Hudson}, the threshold question is not whether the \textit{speech} itself is unlawful: such a test would present a closed tautology;\textsuperscript{136} the question is whether the \textit{conduct urged by the speech} is unlawful.\textsuperscript{137} Relying on \textit{WLF I}, \textit{Caronia} held that, because physicians may prescribe Xyrem for off-label uses, Caronia’s speech did not concern unlawful activity,\textsuperscript{138} and that the statements were not inherently misleading.\textsuperscript{139} Adopting the reasoning of \textit{WLF I} and \textit{Caputo}, \textit{Caronia} found that physicians receiving Caronia’s speech were generally aware of the FDA-approval process and its implications, and could adequately evaluate the validity of their claims: the \textit{Caronia} court wrote “[g]iven the sophistication of the audience to whom the off-label uses were promoted, this Court cannot conclude . . . that

\begin{flushleft}
\textsuperscript{133} \textit{Caronia}, 576 F. Supp. 2d at 393.
\textsuperscript{134} \textit{Id.} at 396. The court noted that:

regardless what else might have been covered in his discussions, Caronia’s alleged speech was made on behalf of the manufacturer and clearly (1) encouraged physicians to prescribe Xyrem, (2) referred to a specific product, and (3) was economically motivated. Any such promotion by Caronia to physicians on behalf of Xyrem’s manufacturer of the drug’s off-label uses would be commercial speech and be “entitled to the qualified but nonetheless substantial protection accorded to commercial speech.”

\textit{Id.} (citing Bolger v. Young Drug Prods. Co., 463 U.S. 60, 68 (1983)).
\textsuperscript{135} \textit{Id.} at 402.
\textsuperscript{137} 44 Liquormart v. Rhode Island, 517 U.S. 484, 497 n.7 (1996) (“[T]he First Amendment does not protect speech about unlawful activities.”) (emphasis added)).
\textsuperscript{138} \textit{Caronia}, 576 F. Supp. 2d at 397–98.
\textsuperscript{139} \textit{Id.}
[Caronia’s] speech was inherently misleading.”

Turning to the second prong of the Central Hudson test, Caronia found, relying on WLF I, that restricting off-label promotion served a substantial government interest. Caronia acknowledged the substantial government interest in protecting the health and the public and that, in order to attain that objective, the government had “a substantial interest in compelling manufacturers to get off-label treatments on-label.” In addressing the third prong of the Central Hudson test, and relying yet again on WLF I and Caputo, the Caronia court ruled that prohibition of off-label promotion directly advanced the “substantial government interest in requiring manufacturers to submit supplemental applications to obtain FDA approval for new uses of previously approved drugs.”

Finally, addressing the fourth prong of the Central Hudson test—whether the prohibition is only as extensive as is necessary to further the government’s interest—the court distinguished WLF I, which concerned only a limited form of off-label speech. The Caronia court failed to recognize the WLF I finding that a less restrictive alternative to absolute prohibition existed: namely, “full, complete, and unambiguous disclosure by the manufacturer of its involvement in the subject activities and the fact that the uses discussed were off-label,” and relied instead on WLF I dicta that

---

140 Id. (internal quotation marks and citations omitted) (citing United States v. Caputo, 288 F. Supp. 2d 912, 921 (N.D. Ill. 2003)).
141 Id. at 398.
142 Id.
143 Id.; accord Caputo, 288 F. Supp. 2d at 921; WLF I, 13 F. Supp. 2d 51, 72 (D.D.C. 1998), vacated sub nom. WLF III, 202 F.3d 331 (D.C. Cir. 2000) (“[O]ne of the few mechanisms available to the FDA to compel manufacturer behavior [with regard to ensuring the safety of new uses for drugs] is to constrain marketing options; i.e. control labeling, advertising, and marketing.”).
144 Caronia, 576 F. Supp. 2d at 398.
146 See WLF I, 13 F. Supp. 2d at 54; see also Final Guidance Regulation, supra note 95; Advertising and Promotion, supra note 95.
“[w]ere manufacturers permitted to engage in all forms of marketing of off-label treatments, a different result might be compelled.”¹⁴⁸ Instead, it relied on Caputo’s finding that a First Amendment challenge to the off-label prohibition threatened the FDA’s ability to control and prohibit manufacturer’s off-label promotion.¹⁴⁹ Noting that no less restrictive means of advancing the government’s interest was identified in Caputo,¹⁵⁰ Caronia concluded that “the prohibitions . . . pass constitutional muster under the fourth prong of Central Hudson.”¹⁵¹

III. CITIZENS UNITED: FROM POLITICAL SPEECH TO COMMERCIAL SPEECH

The Supreme Court’s holding in Citizens United v. F.E.C. has garnered significant criticism for its implications for corporate participation in political discourse.¹⁵² However, the Court’s reasoning in striking down the FEC’s regulatory scheme¹⁵³ may impact other areas of heavily regulated speech, including some forms of commercial speech, such as the off-label promotion of prescription pharmaceuticals.

¹⁴⁸ Id. (second emphasis added).
¹⁵¹ Id. at 401–02 (“[T]his Court is unable to identify non-speech restrictions that would likely constrain in any effective way manufacturers from circumventing [the FDA’s] approval process.” (citing Caputo, 288 F. Supp. 2d at 922)).
¹⁵² See, e.g., Dworkin, supra note 19, at 39 (reporting the Citizens United decision arose from the Court’s political preference for the Republican Party and/or its general favoritism of corporate interests); Walker Wilson, supra note 19, at 2368–69; Liptak, supra note 19, at A1; Barack Obama, Remarks by the President in the State of the Union Address (Jan. 27, 2010) (“[L]ast week, the Supreme Court reversed a century of law to open the floodgates for special interests—including foreign companies—to spend without limit in our elections . . . . Well, I don’t think American elections should be bankrolled by America’s most powerful interests, and worse, by foreign entities.”).
A. The Citizens United Decision

In January 2008, Citizens United, a non-profit organization, aired *Hillary: the Movie* that criticized then-Senator Hillary Clinton, and sought to dissuade voters from electing her as the Democratic candidate for President. Citizens United wanted to make the film available via online video-on-demand in the days leading to the 2008 primary. It feared, however, it might violate section 441b of the Bipartisan Campaign Reform Act, which prohibited corporations and unions from using general treasury funds to finance advocacy for the election or defeat of a candidate. Citizens United sought a declaratory judgment that it could air *Hillary*, as well as injunctive relief against the FEC in a District of Columbia trial court. That court granted summary judgment in favor of the FEC, finding the law constitutional, both facially and as applied to Citizens United. The Supreme Court heard the appeal directly, and held section 441b’s prohibition of corporate electioneering unconstitutional. Writing for the Court, Justice Kennedy called the FEC’s ambiguous regulatory scheme an “unprecedented governmental intervention into the realm of speech.”

---

154 *Id.* at 886–87.
155 *Id.* at 887 (“[T]o promote the film, [Citizens United] produced two 10-second ads and one 30-second ad for Hillary. Each ad includes a short (and in our view, pejorative) statement about Senator Clinton.”).
156 *Id.* at 888.
157 *Id.* at 887–88.
158 *Id.* at 887.
159 See Bipartisan Campaign Reform Act, 2 U.S.C.A. § 441b(b)(2) (West 2010).
160 *Citizens United*, 130 S. Ct. at 888.
161 *Id.*
162 *Id.* at 913; see 2 U.S.C. § 441b.
163 *Citizens United*, 130 S. Ct. at 896.
A Prescription for Change

B. But Is It Unprecedented? Citizens United and Off-Label Promotion

A regulatory scheme analogous to the regulation of corporate political speech found unconstitutional in Citizens United exists in the prohibition of off-label promotion. An examination of the parallels between the FDA’s regulation of off-label promotion and the FEC’s regulation of corporate political speech reveals that the FDA’s regulation should fail under the principles set forth in Citizens United.

1. The FDA’s Prohibition of Off-Label Promotion By Manufacturers, Like the Prohibitions in Citizens United, Constitutes a Restriction On Who May Speak

Citizens United recognized that the ban on corporate independent expenditures constituted a restriction on who may speak. In terms reaching more broadly than political speech alone, the Court noted that:

[p]remised on mistrust of governmental power, the First Amendment stands against attempts to disfavor certain subjects or viewpoints. Prohibited too are restrictions distinguishing among different speakers, allowing speech by some but not others. As instruments to censor, these categories are interrelated: speech restrictions based on the identity of the speaker are all too often simply a means to control content.

The Court continued, finding that if the regulation applied to individuals, not corporations, engaged in political advocacy, “no one would believe that it is merely a time, place, or manner restriction on speech. Its purpose and effect are to silence entities whose voices the Government deems to be suspect.”

Likewise, the FDA’s prohibition of off-label promotion is a

---

164 See supra Part II.B.
166 Id. (emphasis added).
167 Id. at 898.
restraint based upon government mistrust of a particular speaker. Believing, rightly or wrongly, that a manufacturer may attempt to persuade physicians to prescribe their drug, and believing, rightly or wrongly, that such a speech is harmful, the FDA severely limits the ways in which manufacturers may discuss with physicians truthful, non-misleading, peer-reviewed scientific or medical journal articles related to an off-label use. Yet that discussion between two physicians, when each is unaffiliated with a pharmaceutical company, is fully protected scientific speech entitled to the highest First Amendment protection. When the manufacturer or its agent speaks, however, the same speech magically transforms into commercial speech—the red-headed stepchild of First Amendment jurisprudence—and the manufacturer may face criminal liability. The facts of Caronia make the disparity clear: Dr. Gleason became the first physician ever prosecuted under the misbranding statutes because he spoke about an off-label use of a drug on behalf of its manufacturer.

---

168 But see infra Parts VI.A & B.
172 See Ohralik v. Ohio State Bar Ass’n, 436 U.S. 447, 456 (1978) (noting that commercial speech receives “a limited measure of protection, commensurate with its subordinate position in the scale of First Amendment values” (emphasis added)).
175 Id.
Had he instead discussed the same topic, made the same recommendations, and said the same exact words while at a casual dinner with a colleague, no criminal sanctions would attach.\textsuperscript{176} That speech regarding off-label uses of a drug becomes criminal only when a manufacturer speaks makes the constitutional infirmity of the FDA’s regulation pellucid.

2. Under Citizens United, the FDA’s Ambiguous Regulatory Scheme Functions Like a Prior Restraint on Speech

\textit{Citizens United} identified section 441b as “an outright ban [on speech], backed by criminal sanctions.”\textsuperscript{177} Acknowledging that the ban does not constitute a prior restraint on speech in the traditional sense,\textsuperscript{178} the Court noted that “[a]s a practical matter, however, given the complexity of the regulations and the deference courts show to administrative determinations, a speaker who wants to avoid the threats of criminal liability and the heavy costs of defending against FEC must ask a governmental agency for prior permission to speak.”\textsuperscript{179} That is, a corporation could not discern whether its proposed speech was lawful from the perplexing regulatory scheme cobbled together by the FEC, and was therefore forced either to remain silent or to ask the government to pass upon the acceptability of its speech prior to speaking.\textsuperscript{180} The FEC, in essence, created a regulatory regime where it could “select what political speech is safe for public consumption by applying ambiguous tests”\textsuperscript{181} which the Court roundly criticized.\textsuperscript{182}

\begin{itemize}
\item \textsuperscript{176} In \textit{Caronia}, the District court was even hesitant to apply the ban to Dr. Gleason—noting his favored status as a physician. See \textit{Caronia}, 576 F. Supp. 2d at 395–96.
\item \textsuperscript{177} \textit{Citizens United v. F.E.C.}, 130 S. Ct. 876, 897 (2010).
\item \textsuperscript{178} \textit{Id.} at 895 (noting that “prospective speakers are not compelled by law to seek an advisory opinion from the FEC before the speech takes place” (emphasis added)).
\item \textsuperscript{179} \textit{Id.} at 896 (citing Bipartisan Campaign Reform Act 2 U.S.C.A. § 437(f) (West 2010)) (equating the regulation to “licensing laws implemented in the 16\textsuperscript{th} and 17\textsuperscript{th} century England, laws and governmental practices of the sort that the First Amendment was drawn to prohibit”).
\item \textsuperscript{180} See \textit{id.} (citing 2 U.S.C. § 437(f)).
\item \textsuperscript{181} \textit{Id.} at 896. The Supreme Court held that “[t]he Government may
The complexity of the FDA’s ban of off-label promotion has clear parallels to the regulatory scheme struck down in *Citizens United*. The FDA’s regulatory scheme is hopelessly muddled, so as to fail to provide clear guidance as to what promotional activities are permitted. Despite the codification of the FDA’s authority in the FDCA\(^\text{183}\) and FDAMA,\(^\text{184}\) the FDA continues to disseminate guidance documents containing “amorphous regulatory interpretation”\(^\text{185}\) and in so doing it essentially “select[s] what . . . speech is safe for public consumption by applying an ambiguous test.”\(^\text{186}\) As in *Citizens United*, a regulatory agency—this time the FDA—has constructed an abstruse regulatory scheme that essentially functions as a prior restraint on speech.\(^\text{187}\)

regulate corporate political speech through disclaimer and disclosure requirements, but it may not suppress that speech altogether.” *Id.* at 886.

\(^{182}\) *Id.* at 895–96 (comparing the FEC’s regulatory scheme to the English licensing laws of the 16th and 17th centuries that prompted the ratification of the First Amendment, and describing the ill-effects of such a regulatory scheme on modern free speech).


\(^{185}\) *Citizens United*, 130 S. Ct. at 889.

\(^{186}\) *Id.* at 896.

\(^{187}\) See *id.* at 895–96. Although commercial speech receives some First Amendment protection, see supra Part II.A, the Supreme Court has intimated from the inception of the commercial speech doctrine that some protections available to other speech might be unavailable to commercial speech. In *Virginia State Board of Pharmacy*, Justice Blackmun noted that because “commercial speech may be more durable than other kinds [of speech] . . . there is little likelihood of its being chilled by proper regulation and forgone entirely.” *Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council*, 425 U.S. 748, 772 n.24 (1976) (emphasis added). As a result, “the greater objectivity and hardness of commercial speech, may make it less necessary to tolerate inaccurate statements for fear of silencing the speaker,” *id.* (emphasis added), and “may also make inapplicable the prohibition against prior restraints.” *Id.* (emphasis added). Relying on this equivocal, footnoted statement in *Virginia State Board of Pharmacy*, the Central Hudson Court found it would be appropriate to require “a system of previewing advertising campaigns to insure that they will not defeat” governmental objectives. *Cent. Hudson Gas & Elec. Corp. v. Public Serv. Comm’n N.Y.*, 447 U.S. 557, 571 n.13 (1980). The Supreme Court, however, has never definitively held that the doctrine of prior
In the FDA’s most recent interpretation of its own authority to regulate the dissemination of truthful, non-misleading, peer-reviewed scientific journal articles, the infirmity of a de facto prior restraint is pellucid. The provision of FDAMA challenged in *WLF II*, section 401, sunsets in 2006, leaving no guidance as to how, if at all, manufacturers could disseminate journal articles to physicians. After three years without guidance, the FDA at last published its “Good Reprint Practices.” Riddled with subjective criteria for appropriate dissemination of journal reprints, it does little to provide adequate guidance. First, the Good Reprint Practices dictate that journal articles must be “distributed separately from information that is promotional in nature,” yet nowhere do they define promotional material, and the FDA has previously interpreted promotional material to include “literature [and] reprints.” Second, the guidance document states, without explanation, that “[t]he information must not . . . pose a significant risk to public health, if relied upon.” Third, it omits the “safe harbor” contained within section 401, as interpreted by the FDA in *WLF III*. It perpetuates, however, the FDA’s “interpretation” of *WLF III*’s holding, notwithstanding the fact that the case was mooted for lack of controversy based on the existence of restraint absolutely does not apply to commercial speech.

---

188 See GOOD REPRINT PRACTICES, supra note 16.


190 GOOD REPRINT PRACTICES, supra note 16 § II.

191 See, e.g., Robert I. Field, *The FDA’s New Guidance for Off-Label Promotion Is Only a Start*, 33 P&T 220, 220 (2008) (noting that “FDAMA’s limitations on off-label promotion expired on September 30, 2006, and Congress has yet to reauthorize them” and that what was, at the time, a draft guidance was “an attempt to fill the void”).

192 See generally GOOD REPRINT PRACTICES, supra note 16.

193 See generally id.

194 See id. § IV. B.


196 GOOD REPRINT PRACTICES, supra note 16 § IV.A.

of this safe harbor.\textsuperscript{198} the guidance document states that the “FDA’s legal authority to determine whether distribution of medical or scientific information constitutes promotion of an unapproved ‘new use’ or whether such activities cause a product to violate the [FDCA] has not changed.”\textsuperscript{199} Because the FDA enabled itself, in the wake of \textit{WLF III}, to “proceed, in the context of case-by-case enforcement, to determine whether in the manufacturer’s written materials and activities . . .”\textsuperscript{200} off-label promotion has occurred, the Good Reprint Practices leaves a manufacturer with no guidance whatsoever as to what it may or may not lawfully do. The manufacturer must either remain silent\textsuperscript{201} or seek clarification—essentially permission—from the FDA to engage in promotional speech.\textsuperscript{202}

One consideration that led the \textit{Citizens United} Court to declare the FEC’s ban on speech unconstitutional has particular relevance to the FDA’s regulatory scheme. In \textit{Citizens United}, the Court noted that “[w]hen the FEC issues advisory opinions that prohibit speech, ‘[m]any persons, rather than undertake the considerable burden (and sometimes risk) of vindicating their rights through case-by-case litigation, will choose simply to abstain from protected speech.’”\textsuperscript{203} The risk of vindicating a corporation’s speech rights through litigation is especially acute in the off-label promotion realm: if a manufacturer is found guilty of criminally misbranding its drug,\textsuperscript{204} the Health and Human Services Office of the Inspector General (OIG) may preclude the manufacturer from

\begin{flushright}
198 \textit{WLF III} Decision Notice, \textit{supra} note 17, at 14, 287. \\
199 \textit{GOOD REPRINT PRACTICES}, \textit{supra} note 16 § III. \\
200 \textit{See WLF III} Decision Notice, \textit{supra} note 17, at 14, 287. \\
201 \textit{Cf.} \textit{Citizens United v. F.E.C.}, 130 S. Ct. 876, 896 (2010) (noting that the FEC’s regulation of corporate political speech left corporations to choose to remain silent or request permission to speak). \\
203 \textit{Id. at 896} (citing Virginia v. Hicks, 539 U.S. 113, 119 (2003)). \\
\end{flushright}
receiving reimbursement from Medicaid and Medicare for prescriptions of any drug it manufactures. Manufacturers “cannot realistically challenge the government in court [on] . . . whether the charges alleged are compatible with the Constitution . . . The risk/reward calculus is skewed dramatically in favor of settlement when a loss would jeopardize the [manufacturer’s] viability by forfeiting government reimbursement for its products.”

While the similarities between a system of regulatory patchwork and the odious, traditional, form of prior restraint in which “prospective speakers are not compelled by law” to seek permission from the government may not be obvious, both systems chill speech in a similar way. The Supreme Court has previously declared unconstitutional regulations that, while not traditionally a prior restraint, function collectively in much the same way. In Bantam Books, Inc. v. Sullivan, the Court struck down a Rhode Island law designed to shield youth from books that the state considered to be obscene, immoral, or impure. The statute enabled the Rhode Island Commission to Encourage Morality in Youth to “investigat[e] situations which may cause . . . undesirable behavior of juveniles . . . [and] recommend legislation, prosecution, and/or treatment that would ameliorate or eliminate said causes.”

---

207 Citizens United, 130 S. Ct. at 895.
209 Id.
particular book was obscene,\textsuperscript{211} sent notices to publishers and bookstores, announcing that the condemned books could not be sold, displayed, or distributed to customers under the age of eighteen and threatening that “[t]he Attorney General will act for us in the case of non-compliance.”\textsuperscript{212} Upon receiving such notice, the petitioners ceased to publish or offer for sale the banned books.\textsuperscript{213} The Supreme Court found that Rhode Island’s scheme constituted a system of prior restraints\textsuperscript{214} in part because “[t]he distributor [was] left to speculate whether the Commission considers [a] publication obscene or simply harmful to juvenile morality,” and “the ‘cooperation’ [the Commission sought] from distributors invariably entail[ed] the complete suppression of the listed publication.”\textsuperscript{215}

IV. FROM CARONIA ONWARD: A ROUTE FORWARD

In conducting its \textit{Central Hudson} analysis, \textit{Caronia} erred in two ways, resulting ultimately in an incorrect ruling on three of the four \textit{Central Hudson} prongs. First, the court too willingly accepted as “substantial” the FDA’s interest in subjecting new uses of medications to FDA approval,\textsuperscript{216} and therefore did not consider the actual interest the FDA is charged with protecting.\textsuperscript{217} Second, this error resulted in \textit{Caronia}’s finding that the regulation served a substantial government interest under \textit{Central Hudson}’s third prong.\textsuperscript{218} Finally, despite acknowledging the consumer savvy of the physicians to whom off-label promotion is directed,\textsuperscript{219} the court

\textsuperscript{211} See \textit{Bantam Books}, 372 U.S. at 64. The commission defined “obscene” more broadly than the meaning of obscene under the First Amendment. \textit{Id.}
\textsuperscript{212} \textit{Id.} at 62 n.5.
\textsuperscript{213} \textit{Id.} at 63.
\textsuperscript{214} \textit{Id.} at 70.
\textsuperscript{215} \textit{Id.} at 71.
\textsuperscript{217} See \textit{infra} Part VI.A.
\textsuperscript{218} See \textit{infra} Part IV.B.
\textsuperscript{219} See \textit{Caronia}, 576 F. Supp. 2d at 397–98.
errantly followed the flawed reasoning of Caputo,\(^{220}\) failing to find a less speech-restrictive alternative to absolute prohibition of an entire class of commercial speech.\(^{221}\) This section addresses each of these errors in turn, and proposes an alternative that, while protecting the FDA’s mission, would also protect the First Amendment rights of manufacturers as speakers.

### A. The Government’s Substantial Interest Redefined

Courts considering the constitutionality of the FDA’s prohibition on off-label promotion thus far\(^{222}\) have erred in their conception of the government’s substantial interest served by regulating off-label promotion. The root of the FDA’s regulatory authority is the FDCA.\(^{223}\) In the FDCA, Congress announced the purpose of the agency:

The [Food and Drug] Administration shall—

(1) Promote the public health by promptly and efficiently reviewing clinical research and taking appropriate action on the marketing of regulated products in a timely manner;

(2) . . . protect the public health by ensuring that . . . (b) human . . . drugs are safe and effective . . . .\(^{224}\)

Yet despite the clear Congressional mandate that the FDA promote public health, the FDA’s purpose has been distorted or misconceived repeatedly by both the FDA and courts in litigation surrounding the constitutionality of the FDA’s regulation of off-label promotion.\(^{225}\) 

WLF I recognized that Supreme Court precedent has repeatedly held the protection of the public’s health and safety to be a

---


\(^{221}\) See Caronia, 576 F. Supp. 2d at 401; see also infra Part IV.C.


\(^{224}\) Id. §§ 393(b)(1–2) (emphasis added).

\(^{225}\) See, e.g., Caronia, 576 F. Supp. 2d at 398; Caputo, 288 F. Supp. 2d at 921; WLF I, 13 F. Supp. 2d at 70–71.
substantial government interest, and admonished that “[a]ny claim that the government’s general interest is insufficient under Central Hudson is frivolous.”226 Certainly, that interest is substantial, if not compelling; but the court went further to assess the government’s purported interests.227 It properly rejected the government’s contention that the FDA could restrict speech out of fear that the information will be misused228 because “[i]f there is one fixed principle in the commercial speech arena, it is that ‘a State’s paternalistic assumption that the public will use truthful, nonmisleading commercial information unwisely cannot justify a decision to suppress it.’”229 It accepted, however, the argument that the FDA had a substantial interest in compelling manufacturers to subject uses of a drug to the FDA approval process.230 WLF I noted, apparently with approval, that conditioning a manufacturer’s ability to disseminate any information about a particular use of its drug on FDA approval would “encourage[], if not compel[]” the manufacturer to seek approval for off-label uses of its drugs.231 The Caputo court adopted the WLF I reasoning regarding substantial government interest232 and held that permitting off-label promotion would “completely undermine the government’s interest in subjecting off-label uses to the FDA’s evaluation process as well as the government’s interest in preserving the integrity of the FDCA’s new drug approval process.”233

Caronia adopted the reasoning from WLF I and Caputo, recognizing “the government’s substantial interest in subjecting off-label uses of a drug . . . to the FDA’s evaluation process.”234

---

226 WLF I, 13 F. Supp. 2d at 69.
227 See id. at 69–70.
228 Id.
229 Id. (citing 44 Liquormart v. Rhode Island, 517 U.S. 484, 497 (1996)).
230 Id. at 70.
231 Id.
233 Id.
Although it acknowledged the government’s substantial interest in ensuring the health and safety of its citizens, identified in *WLF I*, it erred in accepting the FDA’s interest in compelling manufacturers to seek additional indications via the FDA approval process as an additional substantial interest. Perhaps even more obviously, its adoption of *Caputo’s* holding that the success of the FDCA is a substantial government interest* betrays a fundamental misconception of the policies and purposes served by the FDA’s approval process. “Preserving the integrity”* of a statutory scheme such as the FDCA is no more a substantial government interest than is preserving the integrity of Rhode Island’s law forbidding advertisement of retail alcoholic beverages.* Ensuring the success of a law must be subordinated to constitutional concerns.* Subjecting new uses of drugs to the FDA’s approval process and ensuring the rigor of the FDCA are means to an end—namely, promoting the public health and safety—but neither is an end unto itself.* The substantial—indeed compelling—interest served by the FDA’s regulatory scheme is, and should be conceived of as, ensuring public health

---

235 *See WLF I*, 13 F. Supp. 2d at 69.
237 *Caputo*, 288 F. Supp. 2d at 921 (citing Thompson v. Western States Medical Center, 535 U.S. 357, 369 (2002)).
238 *See Ball, Duffy & Russakoff, supra note 88, at 7–9* (implying that courts applying *Western States* have improperly balanced this interest against free speech rights).
239 *Caputo*, 288 F. Supp. 2d at 921.
241 *See Marbury v. Madison*, 5 U.S. 137, 177 (1803) (“Certainly all those who have framed written constitutions contemplate them as forming the fundamental and paramount law of the nation, and consequently the theory of every such government must be, that an act of the legislature, repugnant to the constitution, is void.” (emphasis added)).
242 Even assuming, *arguendo*, however, that the government may have a substantial interest in subjecting new drug uses to FDA scrutiny, less restrictive means of achieving this goal are available. The proposed disclosure requirement, *see infra Part IV.C*, would advance this goal by incentivizing manufacturers to seek to bolster their promotional claims with the FDA’s approval.
and safety: this comports with Congress’ intent in forming the FDA, with the history of drug regulation generally, and with the FDA’s conception of its own mandate.

**B. Medical Marketplace, Meet the Marketplace of Ideas: Off-Label Speech Directly Advances Public Health and Safety**

Promotion of public health and safety is best served by the free exchange of truthful, non-misleading information regarding potential drug therapies, regardless of whether a particular use has been approved by the FDA. The WLF I Court recognized this fact, writing:

> the open dissemination of scientific and medical information regarding [off-label] treatments is of great import. The FDA acknowledges that physicians need reliable and up-to-date information concerning off-label...

---


244 See generally JAMES T. O’REILLY, FOOD AND DRUG ADMINISTRATION §§ 3.1–3.12 (3d ed. 2005). The precursor to the FDA was the Department of Chemistry. The Pure Food and Drugs Act was passed in 1906, see Pure Food and Drug Act, 34 Stat. 768 (1906), in response to public outcry over the abhorrent conditions in meat-packing facilities. See generally UPTON SINCLAIR, THE JUNGLE (1905). The Act announced as the Department of Chemistry’s purpose “preventing the manufacture, sale, or transportation of . . . misbranded . . . or deleterious foods, drugs, medicines, and liquors . . .” Pure Food and Drug Act, 34 Stat. 768 (1906). In 1912, through the Sherley Amendment, the category of misbranding offenses was expanded to include false or fraudulent statements about drugs. 37 Stat. 416 (1911). The Department of Chemistry was reorganized and dissolved in 1927, leading to the formation of the FDA. O’REILLY, supra, § 3.3 (3d ed. 2005).

245 What We Do, FDA.gov (last updated Nov. 18, 2010) http://www.fda.gov/AboutFDA/WhatWeDo/default.htm (“protecting the public health by assuring the safety, efficacy, and security of human . . . drugs” (emphasis added)).

246 See Ball, Duffy & Russakoff, supra note 88, at 7 (“[S]urely there is . . . a substantial interest in providing open access to available data about unapproved uses for drugs because it results in more informed and therefore safer medical decision making.”).
A Prescription for Change

uses . . . . The need for reliable information is particularly acute in the off-label treatment area because the primary source of information usually available to physicians – the FDA approved label – is absent. 247

Under the current regulatory scheme, manufacturers are prohibited from distributing truthful, non-misleading information that is not in the current, FDA-approved labeling. 248 This state of affairs is anathema to the promotion of public health and safety repeatedly recognized to be a substantial or even compelling interest: “the FDA’s public protection mandate should lead it . . . to welcome the wide circulation of contemporaneous and accurate scientific data on off-label pros and cons.” 249 Indeed, the FDA itself envisions its function with regard to drugs as “protecting the public health by assuring the safety, efficacy, and security of human . . . drugs[,] advancing the public health by helping to speed innovations . . . [;] and by helping the public get the accurate, science-based information they need to use medicines and foods to maintain and improve their health.” 250 Recently, the FDA appears to have recognized that its prohibition of off-label speech may diminish public health. 251 The government’s interest is served, not

247 WLF I, 13 F. Supp. 2d 51, 56 (D.D.C. 1998), vacated sub nom. WLF III, 202 F.3d 331 (D.C. Cir. 2000) (emphasis added). See also id. at 55 (noting that an inherent contradiction in the FDA’s regulation of off-label promotion is that “what a manufacturer may lawfully claim that a drug does under the statutory and regulatory scheme, and what a physician may prescribe a drug for, do not match”).

248 See Osborn, supra note 88, at 328.

249 Smith, supra note 43, at 971; see also Fritch, supra note 70, at 364 (advocating mandatory disclosure by pharmaceutical manufacturers of all clinical data regarding the drugs they promote).

250 What We Do, FDA.gov, supra note 245 (emphasis added).

251 See GOOD REPRINT PRACTICES, supra note 16 § III (“[The] FDA does recognize, however, the important public health and policy justification supporting dissemination of truthful and non-misleading medical journal articles and medical or scientific reference publications on [off-label] uses of approved drugs . . . to healthcare professionals and healthcare entities . . . These off-label uses or treatment regimens may be important and may even constitute a medically recognized standard of care. Accordingly, the public health may be advanced by healthcare professionals’ receipt of medical [literature] on
Critics of this position point to the potential for manufacturers to manipulate or mislead physicians by sharing only favorable information regarding their products. This complaint is unfounded for three reasons. First, as noted above, off-label promotion is not inherently misleading. Second, even assuming a manufacturer did attempt to mislead physicians, market competition would drive the dissemination of not only favorable information, but also information revealing any risks or inefficacies of the drugs. Competing manufacturers, insurance companies, pharmacy benefit managers, and states often “counterdetail” by disseminating information about either the superior efficacy of their own product or the lack of efficacy in their competitors’ drugs.

Thus, while the manufacturer of drug unapproved new uses of approved or cleared medical products that are truthful and not misleading.” (emphasis added)).

See Peter E. Kalb & Paul E. Greenberg, Legal and Economic Perspectives Concerning US Government Investigations of Alleged Off-Label Promotion by Drug Manufacturers, 27 PHARMACOECONOMICS 623, 624 (2009) (“[I]t may follow that suppressing untruthful or misleading information advances public health, but there is no reason to believe that suppressing the dissemination of truthful, non-misleading information will have that same effect.”); Byron Stier, Promotion of Off-Label Use: In Favor of a Regulatory Retreat, 19 ALB. L. J. SCI. & TECH. 609, 609 (2009) (arguing that “the FDA’s background prohibition on off-label promotion should recede and just go away” in part because the prohibition “is counterproductive because it curtails the dissemination of useful information that doctors need to make informed judgments”).

See Marc J. Scheineson & M. Lynn Sykes, Major New Initiatives Require Increased Disclosure of Clinical Trial Information, 60 FOOD & DRUG L.J. 525, 543 (2005) (noting that “there have been a number of allegations of pharmaceutical and device companies selectively disclosing favorable clinical trials and/or failing to disclose unfavorable clinical trial results”). But see infra Part IV.C (discussing a potential disclosure solution).


See, e.g., IMS Health, Inc. v. Sorrell, 630 F.3d 263, 280 (2d Cir. 2010) cert granted 131 S.Ct. 857 (2011) (indicating Vermont has a “counter-speech”
A Prescription for Change

X may trumpet to physicians a new study showing a promising new use for drug X, the manufacturer of competing drug Y has an economic incentive to inform those same physicians of five studies showing that drug X is inefficient for that use. The result is the dissemination of a broad range of scientific information, leading to a richer marketplace of ideas and a better-informed medical community.

Third, the paternalistic view that the government may suppress speech where the recipients of the speech may misuse the information provided has been repeatedly rejected in the context of off-label promotion, and beyond. The Supreme Court has noted that “people will perceive their own best interests if only they are well enough informed, and . . . the best means to that end is to open the channels of communication rather than to close them.” Furthermore, the Second Circuit recently struck down a paternalistic Vermont statute that limited the availability to doctors of information regarding medications. Vermont prohibited manufacturers from using information about physicians’ prescribing habits, compiled and sold by data mining companies, in order to tailor promotional speech to each physician’s need.

program in place to disseminate information to physicians).

See Caronia, 576 F. Supp. 2d at 398 (citing 44 Liquormart v. Rhode Island, 517 U.S. 484, 497 (1996)); WLF I, 13 F. Supp. 2d at 69–70 (“To the extent that the FDA is endeavoring to keep information from physicians out of concern that they will misuse that information, the regulation is wholly and completely unsupportable.”).


IMS Health, 630 F.3d at 267 (“We conclude that because [the challenged statute] is a commercial speech restriction that does not advance the substantial state interests asserted by Vermont, and is not narrowly tailored to advance those interests, the statute cannot survive intermediate scrutiny under Central Hudson.”).

The Vermont legislature gave several justifications for this statute, including: (1) the aims of manufacturers in marketing their drugs “often . . . conflict with the goals of the state”\(^{262}\) (2) the “marketplace for ideas on medicine safety and effectiveness is frequently one-sided” resulting in prescribing based upon “incomplete and biased information”\(^{263}\) and (3) “[p]ublic health is ill served by the massive imbalance in information presented to doctors.”\(^{264}\) Finding that the Vermont legislature “inten[ded] to interfere with the marketplace of ideas to promote the interests of the state,”\(^{265}\) the Second Circuit struck down Vermont’s statute, and noted that, even if Vermont succeeded in manipulating physicians’ prescribing habits, “the Supreme Court reminds us that ‘[i]t is precisely this kind of choice, between the dangers of suppressing information, and the dangers of its misuse if it is freely available, that the First Amendment makes for us.’”\(^{266}\)

Physicians’ prescribing decisions, and therefore the public health, will be improved if physicians are well-informed about potential off-label uses. When prescribing medications, physicians need to know whether a medication is a safe and effective treatment: the research of other physicians, scientists, and academics constitutes an essential source of this information.\(^{267}\) Manufacturers are in a unique position as the most efficient aggregators of information about their products, to disseminate such information to physicians.\(^{268}\) The FDA’s policy of absolute suppression of manufacturer dissemination of this information

\(^{262}\) Vt. Act No. 80 § 1(3) (2007).
\(^{263}\) Id. § 1(4).
\(^{264}\) Id. § 1(6).
\(^{265}\) IMS Health, 630 F.3d at 270 (citing Vt. Act No. 80 § 1).
\(^{266}\) Id. at 270 (citing Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, 425 U.S. 748, 770 (1976)).
\(^{268}\) See Stier, supra note 252, at 610–11; Smith, supra note 43, at 971–72. See also Va. State Bd. of Pharmacy, 425 U.S. at 772 n.24 (“[O]rdinarily the advertiser seeks to disseminate information about a specific product or service that he himself provides and presumably knows more about than anyone else.” (emphasis added)).
frustrates, rather than serves, the interest of enhancing and protecting public health.

C. Less Speech-Restrictive Means of Enhancing Public Health: Citizens United’s Approval of Disclosure and Disclaimer and the Benefit of Sophisticated Consumers

Given the clarification of the government’s substantial interest in regulating off-label promotion—enhancing public health and safety— and the failing of a complete ban on off-label speech to achieve that interest, it follows that the current regulatory scheme is more restrictive than necessary. Despite Caronia’s inability to identify any less speech-restrictive means, such means do exist and have been proposed. The most promising—and least speech-restrictive—means of regulating off-label promotion while remaining respectful of the speakers’ constitutional rights comes from what is, at first blush, an unlikely source.

Despite the deep divide between political and commercial speech, much of the reasoning in Citizens United applies just as well to heavily-regulated commercial speech. The parallel infirmities of the regulatory schemes at issue in Citizens United and Caronia suggest that the less speech-restrictive means

269 See supra Part IV.A.
270 See supra Part IV.B.
272 Blackwell & Beck, supra note 88, at 456 (“Other alternatives include: 1) permitting off-label promotion to physicians but not consumers; 2) permitting promotion through any means other than direct to consumer advertising; 3) permitting speech promoting off-label uses except on product labels; and 4) simply clarifying the boundaries between dissemination of off-label information that is considered promotional and, this, prohibited and dissemination that is considered nonpromotional and, thus, permitted.”); see also Ball, Duffy & Russakoff, supra note 88, at 1 (proposing alternative means of encouraging manufacturers to seek additional indications for their medication while respecting free speech rights).
273 See supra Part III.B.
proposed in *Citizens United* would be equally applicable in the off-label promotion context. In striking down section 441’s ban on corporate independent expenditures, *Citizens United* upheld the requirement that corporations disclose their sponsorship of the messages. Relying on precedent that found the government’s interest in the prevention of real or apparent corruption inadequate to justify a ban on independent expenditures, the Court held that disclosure and disclaimer is a “less restrictive alternative to more comprehensive regulations of speech” that avoided the constitutional infirmities of an outright ban.

Although First Amendment jurisprudence traditionally frowns on compelled speech, compelled disclosures in commercial speech are favored over the alternative here: complete suppression of a particular type of commercial speech. Thus, if a required disclosure would be a constitutional means of “permit[ting] citizens and shareholders to react to the [political] speech of corporate entities in a proper way,” certainly, it would be constitutional in regard to commercial speech, where compelled speech is already less problematic, in order to facilitate the

---


276 *Citizens United*, 130 S. Ct. at 901–02 (citing *Buckley v. Valeo*, 424 U.S. 1, 26 (1976)).

277 *Id.* at 915.

278 See, e.g., *West Virginia Bd. of Ed. v. Barnette*, 319 U.S. 624, 634 (1943) (refusing to sustain a statute mandating a compulsory flag salute in public schools); *United States v. United Foods, Inc.*, 533 U.S. 405, 410 (2001) (“Just as the First Amendment may prevent the government from prohibiting speech, the Amendment may prevent the government from compelling individuals to express certain views.”).


280 See *Citizens United*, 130 S. Ct. at 915.

281 See *WLF I*, 13 F. Supp. 2d at 70.
dissemination of accurate information regarding the off-label use of drugs.

Given the fact that a physician is the targeted recipient of the off-label speech, disclosure would likely be an even more effective regulatory strategy in the off-label context than in the political speech context. Physicians are sophisticated consumers of speech regarding off-label drug uses: they are well aware of the implications for safety and efficacy inherent in a prescription drug’s off-label status. A disclosure to physicians that the advocated use (1) is being advanced by a speaker with a commercial interest in the product, and (2) has not been approved as safe and effective by the FDA, would signal the need for a discerning approach to the information provided and caution in the use of the drug for that purpose.

Concerns have been raised about the potential for a manufacturer engaging in off-label promotion to skew the information provided to physicians by sharing only favorable information about a drug’s efficacy for an off-label use, while downplaying or omitting information that suggests inefficacy. Consistent with the First Amendment, the FDA could require manufacturers to disseminate a bibliography, complete with abstracts, of all scientific literature, favorable and unfavorable, discussing the off-label use which the manufacturer proposes. In

---

282 See id. at 63.
283 See id. (noting that “despite the FDA’s occasional statements in its briefs to the contrary, physicians are a highly educated, professionally-trained and sophisticated audience”).
284 See id. at 73 (Faced with such a disclosure, “[a] physician would be immediately alerted to the fact that the ‘substantial evidence standard’ had not been satisfied, and would evaluate the communicated message accordingly.”).
285 See Fritch, supra note 70, 357 (“For any given prescription drug therapy, there may be a variety of positive and negative studies available, yet drug manufacturers are motivated only to promote studies that reflect positively on their drug.”).
286 See, e.g., WLF I, 13 F. Supp. 2d at 73. Here, although compelled speech is often suspect under First Amendment jurisprudence, it is the lesser of two evils. See Bates v. State Bar Ass’n of Arizona, 433 U.S. 350, 375 (1970). For example, under such a regime, Mr. Caronia could have discussed with the physician the use of Xyrem in fibromyalgia patients, one of the off-label uses he
fact, under the FDA’s most recent guidance for the dissemination of scientific literature that discusses off-label uses of medications, similar requirements exist. Given the existence of such a requirement, concerns about misleading dissemination of off-label information are unfounded.

Critics of a disclosure-based approach to regulating off-label promotion may be concerned that allowing off-label promotion, even with disclosure, may circumvent the FDA’s approval process; however ample opportunities to implement incentives to seek on-label status for drugs remain available. The government would be free to, inter alia

preempt product liability cases for products that receive FDA approval, but preserve product liability theories against uses which have not received FDA approval[;] . . . provide several economic incentives to encourage companies to seek FDA approval for new uses[;] . . . [or] establish a streamlined approval process for an already-approved drugs’ additional widespread uses.

was originally accused of promoting. If he chose to do so, he would be required to provide the physician with a bibliography containing the peer-reviewed, medical journal articles discussing this use for Xyrem. See PUBMED, http://www.pubmed.gov (last visited Apr. 8, 2010) (search for “sodium oxybate fibromyalgia”).

See GOOD REPRINT PRACTICES, supra note 16 § 4.B (mandating that manufacturers must disseminate “a comprehensive bibliography of publications discussing adequate and well-controlled clinical studies published in medical journals or medical or scientific texts about the use of the drug” and “a representative publication, when such information exists, that reaches contrary or different conclusions regarding the unapproved use”).

See, e.g., United States v. Caputo, 288 F. Supp. 2d 912, 921 (N.D. Ill. 2003) (“Manufacturers, knowing that they could promote off-label uses, would have an incentive only to seek FDA approval for uses that would be approved easily and inexpensively. Thus, the Court holds that subjecting off-label uses to the FDA’s evaluation process is a substantial governmental interest.”) But see supra, Part IV.A (noting paternalistic fears of misuse of speech does not justify suppression of the speech).

Blackwell & Beck, supra note 88, at 456 (listing additional, less speech restrictive safeguards against circumvention of the approval process).

Ball, Duffy & Russakoff, supra note 88, at 1. See also Blackwell & Beck, supra note 88, at 456.
Additionally, the very fact that a particular use is off-label may give some physicians pause. Even under a disclosure regime, there is a clear economic incentive to “get off-label treatments on-label,” obviating the perceived need for the FDA’s current oppressive regulatory scheme.

V. CONCLUSION

Caronia presents a concrete example of the ramifications for free speech inherent in the FDA’s current regulatory regime. Fortunately, the constitutional infirmities of the FDA’s scheme can be remedied by a straightforward application of First Amendment law, including, in particular, the recent decision in Citizens United. The complexity of the FDA’s scheme, like the FEC’s in Citizens United, is so abstruse that it essentially functions as a prior restraint. This de facto prior restraint is indefensible, however, upon a careful examination of the purpose and fit of the FDA’s stranglehold on off-label speech. This stranglehold may ensure the sanctity of the FDA’s regulatory scheme, but this, in itself, is not a substantial governmental interest: the FDA’s only substantial interest in regulating off-label promotion is the promotion of public health and safety. Dissemination, not suppression, of scientific information regarding off-label uses of drugs serves this interest. To the extent that concern may linger as to a potential conflict of interest between the promotion of public health and the manufacturer’s commercial interests, the Supreme Court, in Citizens United, has approved of a less restrictive means of policing this conflict: candid disclosure of the manufacturer’s financial interest in the off-label speech should replace the current prohibition. An appeal of Caronia is currently pending: therefore the Second Circuit, and potentially the Supreme Court,

294 See Unofficial Oral Argument Transcript, United States v. Caronia, No. 09 Cr. 5006 (2d Cir. Dec. 2, 2010) (on file with author); Caronia Brief, supra note 1.
will soon have the opportunity to correct the constitutional deprivations currently worked by the FDA and to extend protection to a subset of commercial speech essential to the public health.
FRACKING AND FEDERALISM: SUPPORT FOR AN ADAPTIVE APPROACH THAT AVOIDS THE TRAGEDY OF THE REGULATORY COMMONS

Emily C. Powers*

INTRODUCTION

New York State is currently engaged in the process of crafting a regime to regulate a controversial gas extraction technique called hydrofracking.1 The breadth and scale of hydrofracking’s potential impacts present state and local officials with novel and uncertain environmental and regulatory challenges. These challenges provide a unique opportunity to test some of the assumptions underlying academic discussions of environmental federalism.2

Congress exempted hydrofracking and its roughly thirty affiliated and component processes3 from key portions of federal

* J.D. Candidate, Brooklyn Law School, 2012; M.L.A., Cornell University, 2007. I extend many thanks to Professor Gregg P. Macey for his invaluable guidance, insight, and support; to the staff and Executive Board of the Journal of Law & Policy; and to Pamela Hamilton for her excellent copyediting. I also thank Dick Le Count, Director, Emergency Management Office, Tioga County, NY; Andrew Fagan, Director, Cornell Cooperative Extension for Chemung and Tioga Counties, NY; Rod Howe, Cornell Community and Regional Development Institute; Elaine Jardine, Director of Planning, Tioga County, NY; David Kay, Cornell Community and Regional Development Institute; Judith Quigley, County Attorney, Tioga County, NY; Wendy Walsh, Director, Soil and Water Conservation District, Tioga County, NY.

1 See infra Part I.


3 See 3D Rig Animation, ENERGY IN DEPTH, http://www.energyindepth.
environmental laws, leaving regulation largely to the States. As a result, policymakers in gas-rich states like New York are under unusually high pressure to make difficult trade-offs between significant economic benefits and uncertain harms to public health and the environment, some of them potentially catastrophic and long lasting. The difficult choices hydrofracking poses and the nature of its potential harms illustrate the character of federalism concerns within the context of environmental problems.

Although New York generally permits local governments to regulate local activities, it is hardly surprising that in such a high-
Fracking and Federalism

stakes atmosphere the State has removed almost all regulatory authority from municipalities.\(^8\) This contraction of regulatory authority may lead to suboptimal results. On the one hand, strict state primacy with diminished local input threatens to result in inadequate environmental protection.\(^9\) On the other hand, concerns about the adequacy of New York’s proposed regulatory regime have led to sharp and vocal criticism of hydrofracking in general, which has in turn generated intense public opposition to the practice.\(^10\) This opposition itself threatens to curtail gas production.\(^11\) New York’s regulatory primacy could lead to both underprotection and underdevelopment of natural gas resources as public and political fears dominate regulators’ decision making processes.\(^12\)

The daunting task of addressing complex policy problems, like hydrofracking, has driven academic debate over how federal, state, and local governments can interact within our federalist system to most effectively protect environmental quality without unduly sacrificing economic growth.\(^13\) Legal scholars drawing from economics and political science make efficiency-maximizing arguments based on assumptions about the way markets and

---

\(^8\) Note that this restriction of local authority is not unique to hydrofracking, and has long been applied to oil and gas extraction in the state. See N.Y. ENVTL. CONSERVATION LAW § 23-0303 (McKinney 2010).

\(^9\) See infra Parts III–IV.


\(^11\) See supra note 10; infra Part IV.

\(^12\) See infra Parts III–IV.

\(^13\) See infra Part II.B.1.
rational actors function.\textsuperscript{14} However, recent theories take a cue from ecology and attempt to describe the dynamic interplay among the levels of government.\textsuperscript{15} These theories apply observations about structural federalism in order to craft flexible frameworks for generating policy responses to environmental problems.

Two relatively new approaches, “adaptive federalism,”\textsuperscript{16} and “the regulatory commons,”\textsuperscript{17} counter the efficiency-focused approaches to environmental protection. Each approach highlights how a limited focus on matching environmental problems with the “right” tier of government is likely to lead to underprotection.\textsuperscript{18} Adaptive federalism advocates for flexible roles for the three levels of government, based on the observation that overlapping jurisdiction provides a system of vertical checks and balances.\textsuperscript{19} The regulatory commons analysis outlines how confusion over

\textsuperscript{14} See infra Part II.B.1.
\textsuperscript{15} See infra Part II.B.2.
\textsuperscript{17} William W. Buzbee, Recognizing the Regulatory Commons: A Theory of Regulatory Gaps, 89 IOWA L. REV. 1 (2003) [hereinafter Recognizing the Regulatory Commons]. Buzbee has expounded upon the mechanisms at work in the regulatory commons in several articles since. See, e.g., Contextual Environmental Federalism, supra note 16; William W. Buzbee, Asymmetrical Regulation: Risk, Preemption, and the Floor/Ceiling Distinction, 82 N.Y.U. L. REV. 1547 (2007) [hereinafter Asymmetrical Regulation].
\textsuperscript{19} See generally Adaptive Federalism, supra note 16. The “checks and balances” discussion was developed more fully by Engel in an earlier article. See Kirsten H. Engel, Harnessing the Benefits of Dynamic Federalism in Environmental Law, 56 EMORY L.J. 159, 179 (2006) [hereinafter Harnessing the Benefits of Dynamic Federalism]. Similarly, Buzbee’s discussion of contextual environmental federalism challenges matching arguments that presume state primacy is preferable to a robust, three-tiered scheme. See Contextual Environmental Federalism, supra note 16.
Fracking and Federalism

jurisdictional boundaries can lead to gaps in protection even where an apparently vigorous overlapping regulatory scheme is in place. Together, these approaches provide a hopeful perspective on the future of environmental federalism while remaining realistic about how protective goals can be frustrated.

This Note asserts that the kinds of challenges hydrofracking presents, as well as its current regulatory status, lend support to theories that advocate for an adaptive approach to federalism. In forming its conclusions, this Note draws heavily on the experiences of some county-level officials in New York State, which illustrate the kinds of pressures localities face and the compromises they are compelled to make in the high stakes atmosphere surrounding hydrofracking. These officials’ accounts and New York’s overall experience suggest that not only is underprotection likely when complex policy choices are left solely to the States, but that federal regulatory involvement can prevent the development of market inefficiencies that arise when massive uncertainty over environmental impacts stimulates popular opposition to a new technology.

In Part I, this Note describes what hydrofracking is, and lays out its harms and history. Part II introduces the relevant approaches to environmental federalism and some of their weaknesses. Part III explains the existing regulatory framework for hydrofracking, presents the approaches some officials in gas-rich Tioga County, NY take as they prepare for the commencement of

---

20 See Recognizing the Regulatory Commons, supra note 17.
21 See infra Part IV.
22 County-level officials were selected for their capacity to communicate both detail and generality about how a number of towns and villages are dealing with the same or similar sets of issues.
23 Tioga County was chosen as a focal point for this Note during discussions with David Kay at the Community and Regional Development Institute at Cornell University (CARDI), whom I contacted because CARDI has been actively involved in trying to help central New York communities prepare for hydrofracking. Tioga County’s geologic features and proximity to the Millennium Pipeline make it particularly amenable to hydrofracking, thus a lot of activity is likely to occur there. In addition, as a primarily rural area, Tioga County shares landscape and political similarity with other counties likely to experience hydrofracking activity, thus is fairly representative of the kinds of
a high level of hydrofracking activity, and describes their perspectives on the sufficiency of the available regulatory tools. Part IV considers the policy implications of these officials’ and New York State’s experience with hydrofracking within the discussion of environmental federalism. Part IV also describes how the State’s experience provides insight into incentive structures that result in regulatory commons problems. This Note concludes by asserting that a conception of federalism that includes all three levels of authority—local, state, and federal—is more likely to forestall commons problems than reliance on strict state primacy.

I. THE HIGH STAKES OF HYDRAULIC FRACTURING IN NEW YORK STATE

Recent interest in developing energy alternatives to oil, and advancements in natural gas extraction technologies, have led to controversy over the proper way to regulate a drilling and production process called hydrofracking. The potential harms from hydrofracking are serious, and some would be irreversible. At the same time, the economic benefits to be derived from hydrofracking are substantial. Thus, public debate has focused on

issues that will be faced throughout the state. After identifying Tioga County as a good subject area for study, I contacted Andrew Fagan at the Cornell Cooperative Extension for Tioga and Chemung Counties, also a member of Tioga County Investigates Natural Gas, a working group formed in order to anticipate hydrofracking issues, who helped me identify and select other officials whose duties are likely to be significantly implicated by hydrofracking. I visited Owego, NY, in October of 2010 in order to conduct interviews with Andrew Fagan and Elaine Jardine, Director of Planning for Tioga County. Jardine suggested I contact Dick Le Count, Director of Emergency Management, and Judith Quigley, County Attorney. Interviews with Wendy Walsh, Director of the Soil and Water Conservation District, Le Count, and Quigley were conducted by telephone. During the interviews, I asked the officials to describe their experiences with preparing for hydrofracking, the kinds of concerns they have, how they are addressing these concerns, and how they feel their responsibilities are impacted by the current and proposed regulatory scheme.

24 See infra Part I.A.
25 See infra Part I.C.
26 See infra Part I.C.
the value, safety, and wisdom of allowing hydrofracking, and has divided many communities along economic lines.27

A. Hydraulic Fracturing: Description and Background

Hydrofracking is a term derived from the name of a gas drilling and extraction process called hydraulic fracturing. Hydraulic fracturing involves the injection of fluid into a well to cause subsurface formations to fracture and release natural gas.28 First developed in the 1940s, hydraulic fracturing has been used throughout the country for about sixty years and in New York State since the 1950s.29 Beginning in the 1990s, the oil and gas industry began to use hydraulic fracturing in horizontal wells—so called because they are drilled on an angle to run horizontally within target formations below the earth’s surface.30 The first horizontal well in the East was drilled in Pennsylvania in 2003 to reach the gas-rich Marcellus shale formation, which underlies much of the Appalachian region.31 At present, “hydrofracking” is a term generically used to refer to a process that employs hydraulic fracturing, horizontal drilling, and high-volume fluid injection.32

Hydrofracking is an intensive industrial activity that involves significant environmental disturbance.33 First, an access road is

27 See infra Part I.C.
28 Marcellus Shale, supra note 2.
29 N.Y. ST. DEP’T OF ENVTL. CONSERVATION, DRAFT SUPPLEMENTAL GENERIC ENVIRONMENTAL IMPACT STATEMENT ON THE OIL, GAS, AND SOLUTION MINING REGULATORY PROGRAM 5–32 (2009) [hereinafter SGEIS].
33 SGEIS, supra note 29, at ch. 5.
built, after which a two- to five-acre site is cleared and prepared for drilling and pumping operations, and a drilling rig and other storage and processing structures are installed.\textsuperscript{34} Next, a well is drilled deep into bedrock, and acid is injected to clean the resulting “wellbore.”\textsuperscript{35} The wellbore is then fitted with a steel and concrete casing, which is perforated within the profile of the target formation to allow fluid to enter and break it up.\textsuperscript{36} Next, millions of gallons of water are trucked to the wellpad,\textsuperscript{37} where it is then mixed with chemical agents with anti-corrosive and anti-bacterial functions, many of which are highly toxic.\textsuperscript{38} In addition, materials like sand, which prop open fractures, and emulsifiers are added to the mix.\textsuperscript{39} The resulting mixture, or fracking fluid, is pumped into the wellbore at high pressure.\textsuperscript{40} The fluid fractures the target formation and escaping gas flows out of the wellbore to gathering lines, which carry it to larger pipelines. An estimated 9 to 35 percent of the fracking, or “flowback,” fluid flows back up the wellbore over a period of about two weeks.\textsuperscript{41} The rest remains below the earth’s surface and has the potential to move through cracks in well casings or the target substrate into surrounding rock and eventually to migrate into and contaminate groundwater sources for waterways and drinking supplies.\textsuperscript{42}

\textsuperscript{34} Id. at 5-5 to -12.
\textsuperscript{35} Id. at 5-93. More than one well will often be drilled on a single wellpad. Id. at 5-20.
\textsuperscript{36} See 3D Rig Animation, supra note 3. SGEIS, supra note 29, at 5-91, 5-93.
\textsuperscript{37} The DEC has estimated that each frack job will require 400–600 tanker trucks of water. SGEIS, supra note 29, at 6-137.
\textsuperscript{38} See Hydraulic Fracturing 101, EARTHWORKS ACTION, http://www.earthworksaction.org/FracingDetails.cfm (last visited Feb. 27, 2011); see also SGEIS, supra note 29, at 5-32.
\textsuperscript{39} See Marcellus Shale, supra note 2.
\textsuperscript{41} See also SGEIS, supra note 29, at 5-97.
\textsuperscript{42} Id. at 6-37. DEC considers this a low probability, considering that fluid will be injected into the dense shale layer thousands of feet below water supply aquifers, thus treats leaks that occur during drilling or from surface spills as
Fracking and Federalism

The combination of horizontal drilling and hydraulic fracturing presents concerns because it greatly increases the extent of drilling and the quantity of fracking fluid required. Wellbores can extend as far down as 12,000 feet, and can radiate as far as 4,500 feet from a wellsite. Used together, these techniques have the potential to reduce surface disturbance, because horizontal drilling can access subsurface deposits of gas via fewer surface perforations that conventional drilling allows. However, less surface disturbance comes at the expense of subsurface integrity, which can be cause for concern given the potential for both chemical-laced fracking fluid and natural gas to migrate into and contaminate aquifers and groundwater supplies. In addition, the sudden increase in profitability due to new access to gas deposits has incentivized a rush to states with hydrocarbon-rich formations. Many fear that the rapid onset of a high level of fracking will generate novel problems at a rate and scale that will make them difficult to detect or address adequately.


45 SGEIS, supra note 29, at 5-22. However, even though there may be fewer wellsites, each wellpad disturbs more surface area than smaller wellpads for conventional gas production. Id. at 5-21.

46 Id. at 5-20. See also id. at 6-153.

47 Id. at 6-34 to -37; see, e.g., Delen Goldberg, Tioga County Man Blames Nearby Gas Drilling for Polluting His Well, SYRACUSE.COM (Jan. 2, 2010, 6:00 AM), http://www.syracuse.com/news/index.ssf/2010/01/tioga_county_man_blames_natura.html.

48 Harper, supra note 31, at 5.

49 SUSAN RIHA, DIR., N.Y. WATER RESOURCES INST. & CHARLES L. PECK, PROF. DEP’T OF EARTH & ATMOSPHERIC SCI., CORNELL UNIV. ET AL., COMMENTS ON DRAFT SGEIS ON THE OIL, GAS AND SOLUTION MINING REGULATORY PROGRAM [hereinafter Riha et al., Comments], available at http://cce.cornell.edu/EnergyClimateChange/NaturalGasDev/Documents/PDFs/
B. Natural Gas in New York State

Due to the State’s geologic and hydrologic features, natural gas has been mined in New York since the 1820s, long before the oil and gas industry pioneered methods to extract gas from hard shales like the Marcellus.\(^{50}\) Prior to the development of hydrofracking and horizontal drilling methods, gas extraction in New York State was on the decline due to the inability of producers to access gas trapped in shale.\(^{51}\) Growing energy demands over the past fifty years have focused industry and government attention on developing domestic shale gas resources.\(^{52}\) In addition, the construction of a major natural gas pipeline through New York’s Southern Tier\(^{53}\) and the proximity of nearby gas markets in the rest of the state, as well as in New Jersey and Pennsylvania, have directed industry efforts towards New York and other Marcellus states.\(^{54}\)

The Marcellus formation is attractive because it has an extensive range, covering nearly thirty-four million acres from New York to Tennessee.\(^{55}\) A 2008 report estimates that there are as

---

\(^{50}\) Native Americans introduced early explorers of western New York to naturally occurring natural gas springs in Ontario County and near the present day town of Canandaigua, in Allegany County. The first gas well was drilled in Devonian shale near the town of Fredonia, Chataqua county, in 1821, and gas extraction has occurred in the state ever since. \textit{John P. Herrick, Empire Oil: The Story of Oil in New York State} 316–25 (1949); SGEIS, supra note 29, at 4-1.

\(^{51}\) \textit{Herrick, supra} note 50, at 316–76; SGEIS, \textit{supra} note 29, at 4-2 to -3; Harper, \textit{supra} note 31, at 3–5.

\(^{52}\) \textit{Marcellus Shale, supra} note 2.

\(^{53}\) The Millennium Pipeline is a natural gas pipeline that was completed in 2008 and extends across the southern tier of New York from Corning to Ramapo supplying customers including National Grid, Con Edison, Central Hudson Gas & Electric, and Orange & Rockland Utilities. \textit{Millennium Pipeline, NiSource Gas Transmission & Storage}, http://www.ngts.com/about-ngts/millennium-pipeline/ (last visited Feb. 10, 2011).

\(^{54}\) \textit{Marcellus Shale, supra} note 2.

many as five hundred trillion cubic feet of natural gas in place, with fifty trillion cubic feet recoverable.\textsuperscript{56} The New York State Department of Environmental Conservation (DEC) estimates that roughly eight to ten trillion cubic feet of natural gas could be recovered over time from the Marcellus in New York,\textsuperscript{57} where statewide annual consumption is over one trillion cubic yards.\textsuperscript{58} Thus, the Marcellus presents New Yorkers with an intrastate source of energy, a valued prospect during a period when the nation seeks alternative bridge fuels.\textsuperscript{59} In addition, below the Marcellus lie other shale formations that most likely will be subject to exploration and extraction in the future.\textsuperscript{60}

\textbf{C. Hydraulic Fracturing: Potential Harms and Benefits}

The hydrofracking process\textsuperscript{61} has already been used in other states, providing a sense of what New York faces in the near future. The technique was pioneered in the 1990s as a means to force the last remaining gas from old wells in the Barnett Shale in Texas.\textsuperscript{62} Fracking has also been used to varying extents throughout the country, most extensively in the West.\textsuperscript{63} In the East, the first
high-volume fracking operations are in Pennsylvania and West Virginia.64

There are a number of documented or suspected harms from hydrofracking operations. These range from quality of life issues, such as persistent noise and vibrations from drilling and underground injection, to health impacts from exposure to air and water pollutants, to property value destruction, to social disruption.65 Specific concerns include: the threat that gas or fracking fluid can pollute groundwater;66 toxic air emissions from gas leaks, processing, gas “flaring” and truck exhaust;67 erosion from construction and pipeline siting;68 degradation of surface waterways from leaks, accidental chemical spills, and stormwater runoff;69 noise and light pollution;70 increased truck traffic and roadway deterioration;71 and destruction of ecologically sensitive habitat and the landscape.72 Other potential problems include


65 See SGEIS, supra note 29, at ch. 6.


67 SGEIS, supra note 29, at 6-109 to -123.

68 Id. at 6-16; Telephone Interview with Wendy Walsh, Dir., Soil and Water Conservation District, Tioga County, NY (Oct. 4, 2010).

69 SGEIS, supra note 29, at 6-15 to -16.

70 Telephone Interview with Judith Quigley, County Att’y, Tioga County, NY (Oct. 20, 2010).

71 SGEIS, supra note 29, at 6-138 to -139; Telephone Interview with Dick Le Count, Director, Emergency Management Office, Tioga County, NY (Oct. 19, 2010).

72 SGEIS, supra note 29, at 6-44 to -47, 6-139 to -140 (discussing potential for invasive species introductions and increased access to remote areas and disturbance, which are the kinds of activities that pose threats to ecologically sensitive areas). Interview with Elaine Jardine, Dir. of Planning, Tioga County, NY in Owego, NY (Oct. 15, 2010); Telephone Interview with Judith Quigley,
chemical fires and gas explosions\textsuperscript{73} and hydrofracking fluid and solid waste disposal,\textsuperscript{74} including disposal of waste high in naturally occurring radioactive elements (“NORMS”).\textsuperscript{75} In addition, subsurface interference could cause seismic disturbance and trigger earthquakes.\textsuperscript{76} More likely, seismic disruption could induce fractures in subsurface faults, causing gas to migrate into aquifers and thus contaminate creeks, wells, and other waterways.\textsuperscript{77} Migration of gas into water sources and aquifers is of particular concern in a water-rich state like New York, where surface and groundwater connections are extensive.\textsuperscript{78} Hydrofracking in the Catskills poses a potentially grave threat to New York City’s water supply, which is fed by surface streams and groundwater. Furthermore, some are concerned that drillers and consumers will be forced to compete for water, especially during summer months when water sources are low.\textsuperscript{79} A heavy burden on water bodies is also likely to result in diminished water quality and higher concentrations of pollutants.\textsuperscript{80} Finally, the potential for social disruption is considerable as municipalities face new injections of

\textsuperscript{73} Telephone Interview with Dick Le Count, \textit{supra} note 71.
\textsuperscript{75} SGEIS, \textit{supra} note 29, at 6-40. \textit{See also Drilling Down, supra} note 74.
\textsuperscript{78} \textit{Id.} at 13-20.
\textsuperscript{79} \textit{Riha et al., Comments, supra} note 49.
\textsuperscript{80} SGEIS, \textit{supra} note 29, at 6-3 to -14.
wealth, sudden land use changes, and large influxes of temporary workers. The81
Hydrofracking primarily occurs on private or state-owned lands leased by extractors, also known as landmen, who compensate landowners with lease and royalty payments.82 Conflicts between industry and landowners have arisen over leases in other parts of the country and often have been resolved to the disadvantage of landowners.83 In addition, landowners who do not wish to enter lease agreements can see their interests compromised due to common law interpretations of rights to subsurface migratory resources, such as the Rule of Capture.84 Drilling in New

81 Telephone Interview with Judith Quigley, supra note 70. See also SGEIS, supra note 29, at 6-139 to -140 (discussing some of the kinds of disruption that occur before and during drilling).
83 Courts have interpreted royalty and compensation disputes in favor of oil and gas industry interests even though the industry enjoys informational and strategic advantages such as greater familiarity with contracts. See Thomas A. Mitchell, The Future of Oil and Gas Jurisprudence: Past as Prologue, 49 WASHBURN L.J. 379, 406 (2010). In some instances, gas companies have seemed to take advantage of informational disparities. Telephone Interview with Judith Quigley, supra note 70. As an example, some gas companies have been pursuing leases in New York since the early 1980s, and some residents have been displeased to see their neighbors, who have held out longer or begun to organize, achieve significantly better lease terms and higher compensation arrangements. Id. In some cases, compensation has increased from $20/acre up to $3000/acre over period of a few years. Id.; Telephone Interview with Wendy Walsh, supra note 68. See also Marcellus Shale – Appalachian Basin Gas Play, supra note 55; Michael Rubinkam, The Marcellus Dilemma: One Family’s Struggle with a Giant Natural Gas Company, ITHACA J. (Nov. 25, 2010, 6:05 PM), http://www.theithacajournal.com/article/20101125/NEWS01/11250342/The+Marcellus+dilemma++One+family+s+struggle+with+a+giant+natural+gas+company.
84 See Mitchell, supra note 83, at 406 (noting that natural gas from the Marcellus “will be leased, developed, and produced . . . solely under the ‘pure’ rule of capture as articulated at the end of the nineteenth century.”).
Fracking and Federalism

York State is also permitted on public lands, although with limitation; drilling is not allowed within the Adirondack State Park or in parts of the Catskills.\(^{85}\)

Despite the myriad potential harms from hydrofracking, many New York politicians and residents wish to encourage it and support permissive regulation.\(^{86}\) Gas production is enormously profitable and brings hard-to-resist economic benefits to landowners and depressed areas of the state.\(^{87}\) At the time of writing, royalties from just one acre of leased land can total about $180,000 a year, in addition to a signing bonus of several hundred to several thousand dollars.\(^{88}\) Multiply that amount by the many acres that a lessor might own and it becomes clear why, for many individual landowners, the incentives to permit drilling outweigh the costs.\(^{89}\) Relatedly, increased gas production brings job opportunities to economically depressed areas of the state, so landowners are not the only ones who benefit from hydrofracking.\(^{90}\) In addition, natural gas is characterized by many as a fuel critical to bridge the transition from oil to renewable and

---

\(^{85}\) Telephone Interview with Judith Quigley, supra note 70. See also SGEIS, supra note 29, at 1-2 to -3.


\(^{88}\) This is based on a typical royalty rate of 12.5%, a production average of one million cubic feet per day, and the last available published natural gas rate per cubic foot, $3.96 (rate in December, 2010), for a one acre unit. See Natural Gas Royalty Estimate, GEOLOGY.COM, http://geology.com/royalty/ (last visited Mar. 18, 2011), for a “royalty calculator.” Note that production averages range from .2 to 2 million cubic feet per day. Id. Last available gas rate was taken from Natural Gas Navigator, U.S. Natural Gas Wellhead Price, U.S. ENERGY INFO. ADMIN., U.S. DEP’T OF ENERGY (last updated Feb. 28, 2011), http://tonto.eia.doe.gov/dnav/ng/hist/n9190us3M.htm.

\(^{89}\) See Krauss & Zeller, supra note 87.

\(^{90}\) Id.
other domestic energy sources.91 Yet, despite the potential economic benefits, the sudden influx of enormous profits to previously depressed areas of New York State may create another series of long-lasting and destructive effects. Existing socioeconomic divisions are likely to be exacerbated, creating environmental justice concerns.92 Those who benefit financially from leasing their land will be able to afford to relocate away from any environmental or public health hazards caused by hydrofracking, with their departures diminishing the local tax base. Meanwhile neighbors without financial resources may have no choice but to remain.93 Moreover, some who lease their land may find it difficult or impossible to sell once hydrofracking has begun, or realize that their property has significantly diminished in value.94 Hydrofracking operations and impacts may also make land undesirable for other uses—especially if contamination has occurred. In addition, some farmers may cease operations in light of fracking windfalls, which could alter the landscape significantly.95 However, despite its risks, the

91 See SGEIS, supra note 29, at 2-2.
92 Telephone Interview with Judith Quigley, supra note 70.
93 Id.
95 This is already happening, as some farmers who formerly struggled have ceased farming due to windfalls. However, it is important to note that “gas drilling is allowing farmers to stay on their land, which is environmentally superior to selling it off piecemeal for suburbanization.” Peter Applebome, Will New York Rebel Against Fracking?, GREEN BLOG, N.Y. TIMES (June 10, 2010, 2:17 PM), http://green.blogs.nytimes.com/2010/06/10/will-new-york-rebel-against-fracking/?scp=2&sq=hydraulic+fracturing+farms&st=nyt; see also Krauss & Zeller, supra note 87.
Fracking and Federalism

immediate and short-term benefits make hydrofracking difficult for many in the state to reject and make its regulation a subject of considerable controversy.96

II. THE DEBATE OVER ENVIRONMENTAL FEDERALISM

Congress has enacted a number of environmental laws with the broadly ambitious goals of protecting human health and the environment.97 These statutes have engendered considerable discussion and disagreement over how to balance federal, state, and local roles in the execution of the laws while achieving protection adequate to satisfy congressional mandates.98 The debate over environmental federalism concerns how best to conceive of environmental problems, how to characterize the achievement of statutory goals, and how to distribute regulatory authority among the three levels of government.99 Analysis of the various perspectives on environmental federalism reveals the difficulties inherent in solving complex environmental problems.100 However, it is the complex nature of these problems that suggests that adaptive regulatory approaches may be more suitable than matching approaches at meeting these challenges.101

96 See supra note 10; Krauss & Zeller, supra note 87; infra Parts III.C.1–2.
97 See ROBERT L. GLICKSMAN ET AL., ENVIRONMENTAL PROTECTION: LAW AND POLICY 73–74 (2007). See, e.g., CAA (with the purpose “to protect and enhance the quality of the Nation’s air resources so as to promote the public health and welfare and the productive capacity of its population.” 42 U.S.C.A. § 7401(b)(1) (West 2010)); CWA (with objective to “restore and maintain the chemical, physical, and biological integrity of the Nation’s waters.” 33 U.S.C.A. § 1251(a) (West 2010)); and NEPA (with purposes “to promote efforts which will prevent or eliminate damage to the environment and biosphere and stimulate the health and welfare of man; to enrich the understanding of the ecological systems and natural resources important to the Nation.” 42 U.S.C.A. § 4321 (West (2010)).
98 See infra Part II.B.
99 See infra Part II.B.
100 See infra Part II.B.
101 See infra Part II.B, Parts IV.B–C.
A. Federal Environmental Law

Beginning in 1969, and throughout the 1980s, Congress enacted a host of federal environmental statutes that together articulate a broad goal to protect public resources, including air, water, human health, and ecological integrity. These statutes enjoyed wide political support, which is frequently attributed to a popular recognition of threats to the environment after several highly publicized disasters. Some of these new laws expanded existing schemes, but as a whole they embody an unprecedented and comprehensive approach to protection that acknowledges the irreplaceable value of our environment.

Federal environmental laws use a variety of means to achieve their ends, but one hallmark is their reliance on federal and state implementation and enforcement roles. In the more “cooperative” schemes, the federal government, generally through the U.S. Environmental Protection Agency (EPA), sets minimum standards that industry or states must meet before being subject to sanctions. Laws like the Clean Air Act (CAA), Clean Water Act (CWA), Safe Drinking Water Act (SDWA), and the Emergency Planning and Community Right to Know Act (EPCRA) require states to devise and implement comprehensive plans to meet...
federal goals. Depending on the set of issues at hand, the federal government sets a regulatory floor, which empowers states to formulate regulations for activities within their borders that are more stringent than the federal standards but prevents any state from allowing suboptimal standards. These schemes also have permitting or enforcement components, where actors engaging in certain kinds of activity are required to comply with federal standards directly. Other schemes, like the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), the Resource Conservation and Recovery Act (RCRA), and EPCRA’s Toxics Release Inventory, are directly administered by federal authorities, while still providing some roles for the states. State law and police powers are included in these schemes via savings clauses or where state law does not conflict with, or prevent implementation or enforcement of, federal law—except to the extent that state law has been expressly or impliedly preempted.

Several models are used to explain the need for federal regulation in the face of issues that concern resources held in common. These concepts not only describe the kinds of comprehensive schemes that can be desirable, but also try to predict human behavior by focusing on the logic and incentives that lead to collective action problems. One such explanation is the “tragedy of the commons,” which describes how individual actors are driven by short-term self-interest to pollute a commonly held resource even where each individual knows the group’s collective actions will eventually destroy or seriously damage it. The tragedy is that even where the specter of damage is apparent to all, the lack of capacity to overcome coordination barriers leads to an

107 Id.
108 Id.
109 Id.
110 Id.
111 Id.; see Asymmetrical Regulation, supra note 17, at 1554 (noting role of state standards where a federal floor has been set).
ostensibly inevitable and disastrous outcome. The federal government can be seen as capable or necessary to overcome collective action problems with the resources and information to identify the nation’s needs and the authority to force compliance with standards designed to protect both present and future interests.

The tragedy of the commons can be pervasive even after regulation is implemented; short-term interests often continue to weigh more heavily in decision making than long term or future interests, and additional incentives and political motivations come to bear. As is highlighted in the hydrofracking context, one of the drivers of commons problems even after extensive regulation is the fact that the benefits and costs of regulation are often conceived of as reciprocal. Benefits to industry in the form of less regulation can be correlated with costs of protecting public health and the environment from harm or increased risk of harm.

Thus, the operative inquiry for policymakers concerns who should bear the cost of regulation, which can be discussed in terms of efficiency or may require reaching normative conclusions about who ought to bear the burdens environmental harms present. The idea that harms are reciprocal is the building block of arguments that posit that regulation becomes inefficient when it is overly cautious and leads to unnecessary costs, as in the form of lost jobs.

The “race-to-the-bottom” paradigm has also been used to

---

114 Id.
115 Id.; GLICKSMAN ET AL., supra note 97, at 8–11 (discussing the role of the influence of the “tragedy of the commons” on environmental federalism scholarship).
116 See generally Recognizing the Regulatory Commons, supra note 17 (introducing discussion of “regulatory commons” problems that persist after a regulatory regime is instituted).
118 Id.
Defend federal intervention in environmental issues. Proponents of this argument allege that absent federal involvement, state competition for mobile industry resources will lead state regulators to lower environmental standards to suboptimal levels. The race-to-the-bottom has been the subject of much discussion within the field of environmental law and although contested, there is at least some empirical evidence to suggest that it does occur. More consequentially, study of whether or not states engage in a race-to-the-bottom has revealed that state regulators may feel compelled to lower standards to attract industry even where there is no actual need to do so. Thus, as a threshold matter, state regulators appear to tend towards making underprotective policy choices.

B. The Debate About Environmental Federalism

Disagreement over how best to implement the existing federal environmental laws has given rise to debate about environmental

120 Richard B. Stewart, Pyramids of Sacrifice? Problems of Federalism in Mandating State Implementation of National Environmental Policy, 86 YALE L.J. 1196, 1213 (1977); see also GLICKSMAN ET AL., supra note 97, at 15–27.

121 The “race to the bottom” has its origins in labor law and policy. See, e.g., Liggett v. Lee, 288 U.S. 517, 558–60 (1933) (Brandeis, J., dissenting).


123 Engel and Saleska & Engel presented empirical evidence to show that even if state policymakers do not actually “need” to lower standards to attract industry, they continue to believe that standard-lowering is necessary. See State Environmental Standard-Setting, supra note 122; Facts Are Stubborn Things, supra note 122. In fact, the race to the bottom might be a misnomer if applied in the context of natural gas extraction. With gas, the concern is not with mobile industry in the typical sense. Rather, gas companies will likely seek to access the extent of gas wherever it is available, limited only by technological feasibility, so interstate competition is less operative.
federalism. Discussion revolves around how to balance federal, state, and local interests. Some scholars argue for economics-based approaches to determine what level of government should have regulatory primacy, while others attempt to respond to the complex systems that environmental laws govern and the behavioral forces that appear to motivate regulators. However, given the complexity of environmental problems and federal laws that, in many cases, require regulation regardless of cost, adaptive approaches that seek to respond to the dynamic nature of environmental issues are more likely than market based approaches to result in effective policy.

1. Economics and The “Matching Principle”

For the past three decades, academics have debated how to apportion authority over environmental problems most effectively, without encouraging overregulation. A large part of the discussion has turned to economics arguments, which seek efficiency-based means of determining whether the federal government or the states should have primacy. One argument posits that, contrary to conventional wisdom concerning the race-to-the-bottom, states are more likely to maximize all social goods because federal intervention impedes accurate expression of state preferences through state policy choices and market mechanisms. Building on this idea, and the related assumption that inducing internalization of costs leads industry and regulators alike to make more efficient choices, some academics have

124 See infra Parts II.B.1–2.
125 See infra Part II.B.1.
126 See infra Part II.B.2.
127 See infra Parts II.B.2 & IV.
128 See, e.g., Revesz, supra note 122; Jonathan H. Adler, Jurisdictional Mismatch in Environmental Federalism, 14 N.Y.U. ENVT. L.J. 130, 131–33 (2005); Stewart, supra note 120, at 1210; Butler & Macey, supra note 18, at 25.
130 Compare Revesz, supra note 122, with State Environmental Standard-Setting, supra note 122, and Facts Are Stubborn Things, supra note 122.
Fracking and Federalism

espoused the “matching principle,” which involves identifying the best match between a level of government and the geography of an environmental problem.\textsuperscript{131} Under matching theories, the federal government may be the most efficient regulator in some instances, but state or local regulation is most appropriate much of the time.\textsuperscript{132} Thus, geographic correlation between problems and regulatory authorities should guide any attempts to regulate.\textsuperscript{133} A corollary of this argument is that one-size-fits-all federal regulations are ineffective due to highly disparate ecological and social conditions across the states.\textsuperscript{134}

Despite their elegance, economics-based or efficiency-focused models may not be well suited to addressing environmental issues. Environmental laws confront problems that can appear to have clear geographic boundaries but in fact defy clear delineation.\textsuperscript{135} Theories that rely on market or regulatory preferences to measure social goods also run the risk of under-representing non-economic or hard-to-price values, pervasive in environmental issues,\textsuperscript{136} given widespread reliance on cost-benefit analysis in orienting such preferences. Moreover, these theories fail to appreciate fully that social goods and ills, even if reciprocal, are rarely cleanly symmetrical. These theories may also ignore the role uncertainty plays in policymaking or the wide variety of interests and “hidden” considerations that can skew regulatory incentives.\textsuperscript{137} Furthermore,

\begin{itemize}
  \item \textsuperscript{131} See Adler, supra note 128; Butler & Macey, supra note 18, at 25.
  \item \textsuperscript{132} See Adler, supra note 128; Butler & Macey, supra note 18, at 25.
  \item \textsuperscript{133} See, e.g., Adler, supra note 128, at 137–39 (discussing the need for decentralization and the variable needs of different regions).
  \item \textsuperscript{134} Stewart, supra note 120, at 1210; see generally Colloquium, State Roles in U.S. Environmental Law and Policy, 14 N.Y.U. ENVT'L. L.J. (2005); Dan Esty, Revitalizing Environmental Federalism, 95 Mich. L. Rev. 570 (1996).
  \item \textsuperscript{135} See, for example, the difficulty the Supreme Court has had in trying to reconcile agency interpretation manifesting a “scientific” understanding of what delineates a wetland with the statutory meaning of “waters of the United States” in United States v. Riverside Bayview Homes, Inc., 474 U.S. 121 (1985); Solid Waste Agency of Northern Cook County v. Army Corps of Engineers, 531 U.S. 159 (2001); and Rapanos v. United States, 547 U.S. 715 (2006).
  \item \textsuperscript{136} See, e.g., GLICKSMAN ET AL., supra note 97, at 15–27.
  \item \textsuperscript{137} See Facts Are Stubborn Things, supra note 122, at 74–84 (discussing political, behavioral, and risk-based incentive structures that may come to bear
\end{itemize}
tautological approaches that rely on preferences to arrange themselves into a balance between goods and ills and then define that equilibrium as achievement of maximum efficiency lack normative direction. Finally, efficiency-focused arguments do not fully address the iterative possibilities created by a three-tiered system or the potential for exploitation of institutional differences across levels of government to increase efficiency of the existing environmental law framework. The weaknesses in efficiency arguments are particularly relevant where information about systems and risk is both largely incomplete and contested and where economic pressures distort public choice.

2. Adaptive Federalism and the Regulatory Commons

One theory has recently emerged that describes an “ecological” approach to environmental federalism. Critiquing economics-based theories as ill equipped to address the complexity of environmental issues, not to mention the environmental law framework, Kirsten H. Engel and David E. Adelman describe “adaptive federalism” as a form of federalism that embraces flexibility and overlap, features that make ecological systems more durable. According to Engel and Adelman, adaptive federalism is likely to be more responsive to the complexities and variation inherent in environmental problems and to result in higher levels of protection than the “classical” or “static” conceptions, like the matching principle, which they argue assume away critical on regulatory decision making). See also Gregg P. Macey, Coasean Blind Spots: Charting the New Institutionalism, 98 GEO. L. J. 863 (2010) (discussing complex nature of transaction costs and how these create “Coasean blind spots” that complicate decision making in environmental disputes).

140 See generally Contextual Environmental Federalism, supra note 16.
141 See KYSAR, supra note 138, at 71–75.
143 Id. at 1832; see also Harnessing the Benefits of Dynamic Federalism, supra note 19, at 179.
Fracking and Federalism

variables. By contrast, adaptive federalism relies on the institutional stability of our existing environmental law while encouraging flexibility that allows regulators to react to an ever-evolving body of information.

William Buzbee’s discussion of the regulatory commons supplements adaptive federalism by focusing on the operative concerns about overregulation that motivate proponents of matching approaches. Buzbee points out that matching jurisdiction to environmental problems can be difficult because many issues are cross-jurisdictional. Buzbee argues that even where there appears to be too much regulation, as in apparently robust regulatory frameworks, gaps develop due to perceptions of jurisdictional inadequacy, paucity of incentives, and political machinations. Regulators become inattentive to regulatory opportunities because, for instance, multiple regulators share jurisdiction, or causes and effects of an activity make it difficult to identify the regulatory body with controlling jurisdiction. Buzbee’s discussion suggests that the etiology of commons problems is structural and behavioral, and may be pervasive even where a state has sole regulatory authority. Moreover, despite the potential for jurisdictional confusion that overlapping vertical jurisdiction presents, one can conclude that regulatory commons problems are more likely to be prevented by clarifying roles and granting a variety of regulators increased responsibility for problems than by contracting jurisdiction and reducing available resources.

144 Adaptive Federalism, supra note 16, at 1799.
145 See generally id.
146 See generally Recognizing the Regulatory Commons, supra note 17.
147 Id. at 22–28.
148 See id.
149 See id.
150 See id.; Contextual Environmental Federalism, supra note 16, at 126; Asymmetrical Regulation, supra note 17.
151 See Recognizing the Regulatory Commons, supra note 17.
III. HYDROFRACKING: THE REGULATORY FRAMEWORK

Several holes in federal environmental laws allow hydrofracking to escape federal oversight. Some exemptions have been explicitly placed in statutes. Other aspects of hydrofracking slip through loopholes in the laws or simply do not trigger the existing scheme. Therefore, regulatory authority has been handed to state governments. New York’s proposed regime may not provide adequate protection from hydrofracking’s harms. In addition, New York prevents local governments from exercising direct regulatory authority over hydrofracking processes, leaving localities vulnerable to potential environmental and public health harms.

A. Federal Regulation

Like any activity with an impact on the environment, federal environmental laws touch upon aspects of hydrofracking. However, the oil and gas industry successfully lobbied for exemptions for hydrofracking from several major federal environmental laws, many of which went into effect with the enactment of the Energy Policy Act of 2005 (“the Act”). Apparently, the view that exemption from federal statutes and

---

152 See infra Part III.A.; see also note 4.
153 See infra Part III.A.; see also note 4.
154 See infra Part III.A.
155 See infra Parts III.B–C.
156 See infra Parts III.B–C.
157 See infra notes 168–69 and accompanying text.
Fracking and Federalism

reduced federal oversight of oil and gas development would lead to increased energy independence and development of so-called bridge fuels, like natural gas, prevailed in Congress. However, some critics are suspicious of the motives behind what skeptics have termed the “Halliburton loophole.” Whatever its intent, Congress removed federal oversight of most aspects of hydrofracking and its component practices.

Section 322 of the Act exempts hydraulic fracturing from the SDWA, which protects public and municipal water supplies from underground injection and disposal of hazardous substances through imposition of water quality standards. Further, the Act effectively exempted wellpad construction activities associated with hydrofracking from the National Pollutant Discharge Elimination System (NPDES) under the CWA. In addition, because Congress rolled hydrofracking-related practices into its

160 Congress’ characterization of the Act is aligned with the views of the Energy Policy Development Group, as expressed in NAT’L ENERGY POL’Y, supra note 159, at viii-xv; the conference report states that the Act is meant “to ensure jobs for our future with secure, affordable, and reliable energy.” H.R. REP. NO. 109–190, at 1 (2005), reprinted in 2005 U.S.C.C.A.N. 448, 448. See also infra notes 159–67 and accompanying text.

161 The exemptions have been termed the “Halliburton Loophole” because Halliburton pioneered much of the hydrofracking technology, and Vice President Cheney, formerly a Chief Executive Officer of Halliburton, has been widely criticized for permitting industry representatives and lobbyists to participate in a secret task force to create energy policy early in the tenure of the Bush Administration. Some believe the Energy Policy Act of 2005 is a legislative expression of many of the policies hashed out by that task force. See Editorial, The Halliburton Loophole, N.Y. TIMES, Nov. 2, 2009, at A28, available at http://www.nytimes.com/2009/11/03/opinion/03tue3.html?scp=1&sq=halliburton+loophole&st=nyt.

162 See Energy Policy Act, §§ 322–23, 119 Stat. 594, 694, amending the SDWA, 42 U.S.C. § 300h(d), to exclude underground injection from hydraulic fracturing. Previously, in 1997, the 11th Circuit had ruled that that, under the SDWA, the EPA had to regulate “underground injection.” Legal Environmental Assistance Foundation v. EPA, 118 F.3d 1467 (11th Cir. 1997).

exemption language, it potentially expanded existing oil and gas exemptions in CERCLA to aspects of site construction, drilling, and postfracking production. The Act also weakened review under the National Environmental Policy Act (NEPA) by presuming that certain categorical exclusions apply for oil and gas extraction. Hydrofracking is also exempt from RCRA, which provides for federal oversight of storage and disposal of hazardous materials, and from toxic substance reporting requirements under EPCRA.

Hydrofracking is not entirely beyond the scope of federal oversight, yet significant federal involvement is unlikely given the structure of potentially applicable laws. States must still meet water quality standards under the CWA and the CAA’s national ambient air quality standards via existing state-formulated plans. However, current interpretations of “navigable waterways” make it unlikely that the federal government has jurisdiction under the CWA to regulate emissions unless a “significant nexus” exists between an impacted groundwater connection and navigable waters. Establishing a “significant nexus” is likely a difficult showing in the hydrofracking context, as most impacts will be on groundwater sources that are hard if not impossible to trace to navigable waters. In addition, the EPA will not aggregate air

---

164 What’s the Hydro-Fracking Rush, supra note 4, at 12.
165 Id.
166 See Wiseman, supra note 4, at 251 n.125.
168 Rapanos v. United States, 547 U.S. 715, 787 (2006). Justice Kennedy articulated the “significant nexus” test in his concurrence; although not commanding a majority, his position essentially embodies the nexus between the majority and the dissent in Rapanos, suggesting to lower courts that his formulation should govern their decisions. See id.
169 Proving hydrologic connections is likely to be difficult because it may be impossible to determine where and how much fracking fluid will travel subsurface after injection. Furthermore, Rapanos has proven difficult for lower courts to apply and has led to erratic results. See Gregory H. Morrison, Note, A Nexus of Confusion: Why the Agencies Responsible for Clean Water Act Enforcement Should Promulgate a New Set of Rules Governing the Act’s
emissions from the various operations that occur on a wellpad, and the agency has exempted pollutants emitted by surface waste, like fracking fluid, from stationary source regulation under the CAA.\footnote{Wiseman, \textit{supra} note 4, at 251 n.125.}

Courts are also unlikely to hold that the CAA applies to increased emissions from truck traffic.\footnote{Village of Oconomowoc Lake v. Dayton Hudson Corp., 24 F.3d 962, 963 (7th Cir. 1994) (holding that the CAA does not apply to emissions from increased truck traffic).} Finally, although savings clauses in federal laws preserve state police powers and common law authority,\footnote{Asymmetrical Regulation, \textit{supra} note 17, at 1550 (discussing savings clauses in the context of preemption doctrine).} including tort liability for harm after the fact, standing and evidentiary hurdles typically prevent recovery in suits brought over environmental harms.\footnote{See ROBERT V. PERCIVAL ET AL. \textit{ENVIRONMENTAL REGULATION: LAW, SCIENCE, AND POLICY} 70–75 (5th ed. 2006).} Thus, the federal government has effectively vacated the field, and regulation of hydrofracking is achieved via a patchwork of state policies.\footnote{See Wiseman, \textit{supra} note 4, for a comprehensive comparison of state policies.} Although industry often welcomes federal standard setting when faced with the burden of meeting a proliferation of state schemes,\footnote{Asymmetrical Regulation, \textit{supra} note 17, at 1569–70.} it is apparent that in the hydrofracking context, industry supporters have preferred a state-led approach.\footnote{See Letter from John Hoeven, Gov. of N.D., Chairman of Interstate Gas and Oil Compact Commission, to U.S. Reps. Tauzin and Dingell (Mar. 18, 2003), \textit{available at} http://www.energyindepth.org/PDF/Tauzin-Dingell.pdf.}

\begin{center}
\textit{Fracking and Federalism}
\end{center}

\textbf{B. State Regulatory Scheme}

State police power includes the authority to regulate activity that impacts natural resources and human health, and New York State has exercised this power to propose comparatively stringent
environmental regulations on hydrofracking. Article 23 of the New York Environmental Conservation Law (‘‘Article 23’’) establishes the DEC’s broad jurisdiction to regulate oil and gas extraction via its Division of Mineral Resources, with dual regulatory goals of encouragement of natural gas development and protection of correlative ownership. In addition, the Department of Transportation has jurisdiction over transportation of hazardous materials, and the Public Service Commission regulates siting of gas gathering lines, which is not subject to public review under the State Environmental Quality Review Act (SEQRA), New York’s NEPA corollary.

Pursuant to SEQRA, the DEC has devised a land use focused regulatory strategy over hydrofracking implemented largely via permitting and reporting requirements. The DEC prepared its draft supplemental Generic Environmental Impact Statement (SGEIS), released in 2009, after receiving applications for permits to drill using high-volume hydrofracking methods. The SGEIS supplements the DEC’s Generic Environmental Impact Statement, which outlines the agency’s approach to conventional

177 See generally Wiseman, supra note 4.
178 N.Y. ENVTL. CONSERVATION LAW § 23-0301 (McKinney 2010).
181 See SGEIS, supra note 29, at 5-5 to -6. DOT regulations specify that fracking fluid components should be separately trucked and mixed only onsite. Id.
182 N.Y. COMP. CODES R. & REGS. tit. 6 § 617.5(c)(35) (2010); see also SGEIS, supra note 29, at 5-129.
183 SGEIS, supra note 29, at §§ 7.1.1, 7.1.1.1, 7.1.1.3–7.4. The DEC Plan relies on the Susquehanna River Basin Commission and the Delaware River Basin Commission to regulate water withdrawals from those bodies of water.
184 Id. at 1-1.
185 See generally GENERIC ENVIRONMENTAL IMPACT STATEMENT ON THE OIL, GAS, AND SOLUTION MINING REGULATORY PROGRAM, DIVISION OF MIN. RESOURCES, N.Y. ST. DEP’T OF ENVTL. CONSERVATION (1992), available at
Fracking and Federalism

extraction methods and which proved inadequate to address the significantly greater environmental impacts from high volume hydrofracking. 186

The regulatory strategy the DEC has presented in the SGEIS concerns many New Yorkers. The most significant concern that critics share is that the agency is inadequately funded or staffed to ensure compliance with state regulations and policies. 187 For instance, as of 2009, the Division of Mineral Resources had only sixteen enforcement staff members to oversee more than thirteen thousand conventional wells. 188 Even drillers are concerned that their permits will be held up by administrative delays because DEC’s staff is inadequate to process the large number of forthcoming requests. 189


Critics are also concerned that the DEC’s plan does not address the cumulative impacts of even routine aspects of hydrofracking.190 The SGEIS deals with these effects in a cursory fashion and asserts that too much uncertainty exists to be able to assess them with accuracy.191 Many critics feel the DEC’s failure to address these impacts is unsatisfactory, for it is the uncertainty of these effects that frustrates attempts to prepare for them and compounds the risk of harm.192 For example, small-scale chemical spills, accidents, and incremental burdens on surface waters and infrastructure are difficult for localities to anticipate without more information about how extensive drilling will be.193 Some also suggest that the state has not recognized the extent of hidden economic costs associated with environmental contamination and the potential loss of ecosystem services.194 Similarly, there are concerns that emergency management plans are lacking and that worst case scenarios have not been sufficiently elaborated, in light of federal exemptions.195 Moreover, many hold that the DEC inadequately considered the findings and conclusions of regulators from other states that have experienced harms from horizontal drilling and high-volume hydraulic fracturing.196 There is widespread concern that the DEC has not ensured there will be full, public disclosure of chemical components in fracking fluid, and some urge public reporting requirements for frack fluid components, all locations of


190 See Riverkeeper Comments, supra note 187; Riha et al., Comments, supra note 49; DEP Comments, supra note 187.

191 SGEIS, supra note 29, at 6-141 to -143.

192 Riha et al., supra note 49. See also DEP Comments, supra note 187, at 7; Riverkeeper Comments, supra note 187.

193 See Riverkeeper Comments, supra note 187; Riha et al., Comments, supra note 49; DEP Comments, supra note 187, at 9.

194 See Riverkeeper Comments, supra note 187; Riha et al., Comments, supra note 49; DEP Comments, supra note 187, at 39.

195 DEP Comments, supra note 187, at 48.

196 Id. at 1; see also Riverkeeper Comments, supra note 187.
drilling operations, and any spills or contamination.\footnote{Riha et al., Comments, supra note 49; see also About the DEC Supplemental Generic Environmental Impact Statement (SGEIS), N.Y. CITY DEP’T OF ENVTL. PROT., http://www.nyc.gov/html/dep/html/news/natural_gas_drilling_SGEIS.shtml (last visited Feb. 27, 2011) (“[T]he City urged DEC to rescind the draft SGEIS given the serious omissions and the grave consequences of the proposed action.”).} Overall, the SGEIS leaves concern that once hydrofracking kicks into high gear, the State will not be poised to address problems that arise. Addressing emergencies and incidental effects will be largely left to localities, which bear the most risk of immediate harms from hydrofracking.

C. Local Roles

As a home rule state, New York typically allows municipalities a degree of latitude to govern local activities.\footnote{N.Y. MUN. HOME RULE LAW § 1 et. seq. (McKinney 2010).} However, Article 23 stipulates that the State’s regulatory program “supersede[s] all local laws or ordinances relating to the regulation of the oil, gas and solution mining industries” although local primacy over road use and property taxes are retained.\footnote{N.Y. ENVTL. CONSERVATION LAW § 23-0303(2) (McKinney 2010); Michael E. Kenneally & Todd M. Mathes, Natural Gas Production and Municipal Home Rule in New York, 10 N.Y. ZONING L. & PRACTICE REP., no. 4, 2010 at 2, available at http://counties.cce.cornell.edu/yates/documents/NaturalGasProduction.pdf.} For instance, the State will not require local approval of permits under its State Pollution Discharge Elimination System (SPDES) and corollary Stormwater Pollution Prevention Plan (SWPPP) program.\footnote{Telephone Interview with Wendy Walsh, supra note 68.} Thus, localities will not have control over such critical decisions as wellpad siting, stormwater planning, erosion control, or pipeline placement.\footnote{See generally Stormwater, N.Y. ST. DEP’T OF ENVTL. CONSERVATION, http://www.dec.ny.gov/chemical/8468.html (last visited Apr. 7, 2011); SGEIS, supra note 29, at 1-2; Telephone Interview with Wendy Walsh, supra note 68.} The State has issued nonbinding directives to industry to consult with local planning documents and procedures in siting decisions, and operators are expected to comply with local floodplain permitting...
requirements that establish broadly applicable siting and setback guidelines. However, localities cannot set any laws or regulations that specifically refer to or are clearly directed at hydrofracking activities. Thus, hydrofracking is not subject to zoning restrictions, although zoning has long been considered a valid exercise of local authority. Localities are nonetheless expected to investigate water quality complaints and provide emergency response services. Waste disposal and sanitation are also typically local responsibilities, which will be dramatically impacted by the high volume of flowback water, waste, and drill...
cuttings from hydrofracking operations. Despite the significant emergency responsibilities and infrastructural demands hydrofracking poses, localities have few authoritative tools to help them prepare.

1. Local Responses

Many local officials in communities where hydrofracking activity is likely to occur have been attempting to prepare for its probable impacts. Several officials in gas-rich Tioga County, an area likely to see extensive hydrofracking activity, have expressed frustration with their limited ability to influence the course of drilling or its effects. In particular, some are frustrated by the fact that local discretion has been removed from activities typically subject to local input, such as permitting for construction and industrial activities. Others are less concerned about the State’s primacy, do not feel lack of federal oversight is an issue of concern, and downplay fears over some of hydrofracking’s risks. While some are concerned to an extent about threats to the environment and public health, few feel it is within his or her authority to take a position on whether hydrofracking should be encouraged or prevented. Most see the value in compromise,

207 See VANDEMARK, supra note 206, at 41 (setting forth and describing basic services provided at the county level); Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70.

208 See infra Part III.C.1–2.

209 See infra Part III.C.1–2.

210 Telephone Interview with Wendy Walsh, supra note 68.

211 Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70. In particular, Wendy Walsh is concerned about the potential for erosion from wellpad activities and gathering line siting and notes that although she typically has input in the permitting process for construction activities, she does not in the hydrofracking context. She hopes that gas companies will be responsive to local requests about siting. Telephone Interview with Wendy Walsh, supra note 68.

212 Telephone Interview with Dick Le Count, supra note 71 (stating his “biggest worry is that we’ll run out of freshwater” and opining that truck traffic will be the most formidable problem localities will face).

213 Id.; Interview with Andrew Fagan, Dir., Cornell Cooperative Extension
given the contentious nature of hydrofracking and its perceived inevitability.\textsuperscript{214} Understandably, local officials are interested in achieving pragmatic solutions with what little regulatory authority they have.\textsuperscript{215} Several seem persuaded that, as heavy users of natural gas, New Yorkers have an obligation to permit drilling.\textsuperscript{216} Furthermore, local officials recognize the State’s interest in encouraging gas production and understand the rationale behind cutting off local input and control.\textsuperscript{217} Opposition at the local level could allow a few individuals to stall or prevent fracking, which would be undesirable and unacceptable\textsuperscript{218} given the many interests in favor of drilling and the benefits that would come from gas production—even if local opposition could provide a bulwark against hydrofracking’s harms.

\textsuperscript{214} Telephone Interview with Dick Le Count, \textit{supra} note 71; Interview with Andrew Fagan, \textit{supra} note 213; Interview with Elaine Jardine, \textit{supra} note 72 (stating “I can’t say whether it’s good or bad. That’s not my role. My role is to prepare the municipalities for the impact that it will have with the constraints they’re given by New York State government.”); Telephone Interview with Judith Quigley, \textit{supra} note 70 (expressing the common sentiment that people in Tioga County “don’t want to scare companies away” and that “the DEC is already perceived as “overregulating”); Telephone Interview with Wendy Walsh, \textit{supra} note 68.

\textsuperscript{215} Telephone Interview with Dick Le Count, \textit{supra} note 71; Interview with Andrew Fagan, \textit{supra} note 213; Interview with Elaine Jardine, \textit{supra} note 72; Telephone Interview with Judith Quigley, \textit{supra} note 70; Telephone Interview with Wendy Walsh, \textit{supra} note 68.

\textsuperscript{216} Interview with Andrew Fagan, \textit{supra} note 213; Interview with Elaine Jardine, \textit{supra} note 72; Telephone Interview with Wendy Walsh, \textit{supra} note 68.

\textsuperscript{217} Interview with Elaine Jardine, \textit{supra} note 72; Telephone Interview with Wendy Walsh, \textit{supra} note 68.

\textsuperscript{218} Interview with Elaine Jardine, \textit{supra} note 72 (explaining “[i]n defense of the state, they don’t want every gas drill to have to go through local permitting. They know that local permitting is a difficult process here if it’s a controversial use. It’s really hard to get anything through.”); Telephone Interview with Wendy Walsh, \textit{supra} note 68.
Fracking and Federalism

2. Mitigating Impacts from Hydrofracking

Some officials in Tioga County seek to mitigate impacts from hydrofracking through traditional means of municipal control. While a few towns and villages in New York State are testing their zoning discretion, this power is less relevant in Tioga County, where most towns and villages do not have comprehensive zoning plans in place. Therefore, concerned county officials have focused instead on tweaking traffic rules and on implementing light, noise, and wellhead protection ordinances, which restrict uses and set water quality standards within municipal borders. Localities may also require ancillary service providers or businesses, such as pipe yards and chemical storage facilities, to comply with standards for lighting, traffic flow, and signage under site plan review ordinances—although site plan review does not apply to drilling or hydraulic fracturing processes themselves.

It is important to note the limitations of these regulatory tools. In Tioga County, for example, while twelve of the fifteen towns and villages do now have site plan review ordinances in place, only four have enacted zoning regulations. The majority of towns have not enacted new ordinances in anticipation of hydrofracking besides site plan review, and enacting these measures would not ultimately be feasible. Zoning plans require staff and expertise to formulate, while noise and light ordinances and traffic plans require prohibitively expensive environmental testing and

219 Telephone Interview with Dick Le Count, supra note 71; Interview with Andrew Fagan, supra note 213; Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70; Telephone Interview with Wendy Walsh, supra note 68.
220 See Lawyer, supra note 204.
221 Interview with Elaine Jardine, supra note 72.
222 Wellhead protection ordinances set forth general land-use guidelines and setbacks that are intended to protect a municipal water supply from contamination. Id.; Telephone Interview with Judith Quigley, supra note 70.
223 Id.
224 Interview with Elaine Jardine, supra note 72.
225 Id.
engineering consultants. In addition, wellhead protection plans do not apply where a town lacks a public water supply and are not applicable to wellpad activity.

There are also political obstacles to implementing new measures. Many towns and villages have only part-time or volunteer officials without the institutional capacity or political will to enact new ordinances. Furthermore, there is frequently opposition to measures that appear to increase governmental interference with private property rights, even in protective ways. In addition, county legislators have proven to be unwilling to address fears about hydrofracking given uncertainty over practical issues, such as whether groundwater contamination is a serious concern. The county is also suffering from “personnel drain,” as some of the best-trained and most knowledgeable workers are hired by industry in anticipation of drilling. Furthermore, the State has not provided supplemental resources to help localities prepare, nor has it indicated it will do so once fracking begins in earnest.

Perhaps the most crucial tool localities lack is adequate

---

226 Telephone Interview with Judith Quigley, supra note 70. One local engineering firm has put together a package that would help municipalities get reimbursed for wear and tear on roads. Interview with Elaine Jardine, supra note 72.

227 Id.

228 Id.; Telephone Interview with Judith Quigley, supra note 70.

229 Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70.

230 Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70.

231 For instance, one looming issue concerns how long it will take the State to finalize its regulatory scheme. Interview with Elaine Jardine, supra note 72. Note too that many politicians, as landowners, do or will have personal economic interest in encouraging hydrofracking. Telephone Interview with Judith Quigley, supra note 70.

232 Interview with Andrew Fagan, supra note 213; Interview with Elaine Jardine, supra note 72; Telephone Interview with Wendy Walsh, supra note 68.

233 Interview with Elaine Jardine, supra note 72; Telephone Interview with Wendy Walsh, supra note 68.
Fracking and Federalism

enforcement ability. Towns and villages do not have the necessary financial resources to conduct adequate oversight or ensure that violations of local laws are addressed. By contrast, energy companies generally have more than sufficient resources to pay small fines, which are therefore not likely to deter behavior that results in harm. Energy companies also tend to have significant legal resources readily available to challenge town enforcement attempts or to counter opposition.

Aware of their disproportionate resources, localities have facilitated relationships with industry and sought voluntary agreements in which they seek commitments for infrastructure investments. For instance, officials seek promises that industry will purchase firefighting equipment, build, repave, and

234 Telephone Interview with Judith Quigley, supra note 70.
235 Id.
236 See, e.g., Christopher Helman, Range Resources is King of the Marcellus Shale, FORBES.COM (July 22, 2010, 7:00 PM), http://www.forbes.com/forbes/2010/0809/companies-energy-range-resources-bp-gas-blowout-beneficiary.html (describing the $500 million dollar operating income of just one “small” gas company active in the East, Range Resources); see also Donald Gilliland, Marcellus Shale Gas Drillers Committed 1,435 Violations in 2.5 Years, Report Says, PENNLIVE.COM (Aug. 2, 2010, 5:37 PM), http://www.pennlive.com/midstate/index.ssf/2010/08/marcellus_shale_gas_drillers_c.html (reporting that gas industry averaged one and a half regulatory violations a day over two and a half year period); Drilling Down, supra note 74 (describing how energy companies in Pennsylvania have engaged in illegal dumping of fracking wastes because it is cheaper than proper disposal where resulting fines are dwarfed by the daily profits reaped from gas production).
238 Telephone Interview with Dick Le Count, supra note 71; Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70; Telephone Interview with Wendy Walsh, supra note 68.
239 Telephone Interview with Dick Le Count, supra note 71; Interview with Elaine Jardine, supra note 72.
maintain roads;240 use closed systems to store fracking fluid;241 and disclose the chemicals used in fracking mixtures.242 In one striking example, the Tioga County Emergency Management Office will rely solely on industry to extinguish any wellpad fires.243 In addition to realizing that basic needs will be better met if they work with private industry, local officials are interested in fostering positive and cooperative relationships with firms operating in and among their communities.244

IV. HYDROFRACKING AND ENVIRONMENTAL FEDERALISM

New York’s experience with hydrofracking illustrates how an adaptive approach to regulation is more likely to result in sufficient environmental protection than an approach that attempts to match potential problems with a level of authority based on geography. While an essentially localized activity, hydrofracking nonetheless presents a regulatory challenge to state and local governments.245 Deciding whether to encourage or limit hydrofracking requires a highly subjective analysis that relies on uncertain and incomplete

240 Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70; Telephone Interview with Wendy Walsh, supra note 68.

241 Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70; Telephone Interview with Wendy Walsh, supra note 68.

242 Telephone Interview with Dick Le Count, supra note 71; Telephone Interview with Wendy Walsh, supra note 68.

243 Telephone Interview with Dick Le Count, supra note 71.

244 Id.; Interview with Andrew Fagan, supra note 213; Interview with Elaine Jardine, supra note 72; Telephone Interview with Judith Quigley, supra note 70; Telephone Interview with Wendy Walsh, supra note 68. Of course, there are concerns about whether these agreements will satisfactorily address localities’ needs in actuality, and these agreements can give rise to disagreements between energy companies and localities over how best to implement them. For example, although energy companies and towns share an interest in having roads maintained in good quality, thus operators may agree to provide resources to accomplish this task, they may resist doing so in a manner that meets local desires or conventions—or even labor laws. Telephone Interview with Judith Quigley, supra note 70.

245 See supra Part III.
Fracking and Federalism

information about risk.246 Meanwhile, state and local decisionmakers are incentivized to accept risk of harms they lack resources to prevent or mitigate.247 Applying federal laws to hydrofracking would help relieve some of the pressures on state and local authorities by placing the burden of precaution onto energy companies.248 Moreover, there may be appreciable benefits to fostering a flexible regime that includes responsive interaction among all three levels of government.249 Finally, an active federal role in regulating new technologies like hydrofracking can give states and localities a better chance to formulate policies aligned with their resources and expertise, leading to increased political accountability, jurisdictional confidence, and fewer regulatory commons problems.250

A. Strict State Primacy: A Poor Match for Hydrofracking

In many respects, the existing federal approach to regulation of hydrofracking would likely meet the approval of matching principal proponents. A land-based activity without any obvious interstate impacts, hydrofracking seems a good candidate for a state-led approach, and the federal government essentially has mandated state primacy.251 Federal regulation would increase costs and slow fracking efforts.252 If federal laws were operative,

246 See supra Part I.C.
247 An example is contamination of an entire county’s drinking water source, as in Tioga County, where the Clinton Street Ballpark Aquifer supplies both Tioga and neighboring Broome County. See Clinton Street-Ballpark Valley Aquifer System Broome and Tioga County Areas, NY, 50 Fed. Reg. 2025 (Envtl. Prot. Agency Jan. 14, 1985) (final determination). See also infra Part III.C.
248 See infra Part IV.B–C.
249 See infra Part IV.B–C.
250 See infra Part IV.C.
251 See supra Part III.C; see also Wiseman, supra note 4, at 251 n.125.
252 Adler, supra note 128, at 157–60 (arguing that SDWA is an example of inefficient regulation because it makes a poor jurisdictional match); see also Butler & Macey, supra note 18, at 62 (arguing that “[b]ecause land pollution and the potential for groundwater contamination are very localized phenomena, our federalism model leads to the argument that these externalities should be
industry, states, and localities would bear additional SDWA compliance burdens. Energy companies would have to comply with and obtain permits under the CWA, which are costly and delay development. Aggregation of emissions from wellpad activity or truck traffic under the CAA would increase permitting and cause companies to incur pollution control costs. Removal of exemption from CERCLA would increase risk of liability that would create disincentives to drill, and strengthened NEPA review would also add time and cost to preproduction planning. Regulation under RCRA would also add time and expense to disposal of fracking byproducts. Moreover, and especially given the extensive private property interests involved in hydrofracking, local geographic, socioeconomic, geological, and hydrological differences make state regulation arguably more appropriate than federal regulation to meet unique state preferences.254

However, an analysis of New York’s experience with hydrofracking to date suggests that state primacy may well result in underprotection and even hamper production activity.255 Notably, New York’s proposed regulations are more protective than those in many other states,256 yet its plan nonetheless has demonstrated weaknesses.257 New York’s inadequate enforcement and oversight capacity and its failure to anticipate cumulative impacts may mean that state primacy will result in unintended and undesirable outcomes.258

First, it is important to note that New York’s regulations may not result in protection of some baseline standards that EPA has established.259 For instance, concentrations of toxic pollutants in flowback hydrofracking fluids have measured in excess of amounts regulated exclusively by state and local jurisdictions”).

253 Adler, supra note 128, at 157–60.
254 See supra Part I.C.
255 See infra note 285 and accompanying text.
256 See generally Wiseman, supra note 4 (comparing multiple state regimes and describing New York’s proposed regulation as comparatively more protective than other states).
257 See supra Part III.B.
258 See supra Part III.B.
259 Wiseman, supra note 4, at 277.
Fracking and Federalism

that would be permissible under the SDWA, and flowback fluids can be high in pollutants ruled hazardous under RCRA.\textsuperscript{260} Because New York State has not updated its wetland map after pivotal Supreme Court decisions that altered federal jurisdiction over wetlands,\textsuperscript{261} there may be pollutant releases onto what should be federally regulated land.\textsuperscript{262} New York’s ability to achieve water quality standards under the CWA may be seriously overestimated in light of criticism that the State has not accurately estimated the extent of cumulative impacts on water quality.\textsuperscript{263} Air emissions from wellpad activities, if aggregated, might far exceed the minimum requirements that trigger the CAA.\textsuperscript{264} These examples demonstrate that New York’s scheme might result in failure to meet federal standards that would apply absent exemptions or do apply to the same or similar harms caused by other industries and suggest that as a threshold matter, state primacy over hydrofracking will create results inconsistent with existing law.

Furthermore, the New York scheme’s reliance on local implementation of emergency planning, public health, waste disposal, and road regulation enhances the probability that unintended harms may occur. The level of risk that localities are expected to bear is disproportionate to the resources they have to handle that risk.\textsuperscript{265} As a result, the State’s regulatory regime is

\textsuperscript{260} See id at 277–78.
\textsuperscript{261} See Solid Waste Agency of Northern Cook County v. Army Corps of Engineers, 531 U.S. 159 (2001) (invalidating the “migratory bird rule” and holding that the Army Corps does not have jurisdiction over wetlands not directly adjacent to navigable waters); Rapanos v. United States, 547 U.S. 715 (2006) (plurality holding that federal jurisdiction extends only to navigable waterways and water bodies with continuous surface connection with navigable waterways, as well as to wetlands adjacent to these waters). See also supra note 168 and accompanying text.
\textsuperscript{262} See Riha et al., Comments, supra note 49.
\textsuperscript{263} See Drilling Down, supra note 74 (examining how processed wastewater from hydrofracking may result in water body quality below that required by federal standards).
\textsuperscript{265} See supra Part III.C.
effectively dependent on voluntary industry action. Local officials hope energy companies will agree to provide necessary infrastructure—like roads—and emergency response support—like basic firefighting equipment—even though the industry has incentives to downplay and minimize concerns to the detriment of preparedness. Moreover, experiences of local officials show that localities do not feel well equipped to handle even routine incidental, let alone catastrophic, impacts from fracking and that they lack reliable information to help them bargain with energy companies optimally.268

Lack of oversight and enforcement powers at state and local levels may lead to lax, inconsistent, or insufficient compliance with existing state and local regulations. Even where officials are dedicated to proactive prevention and oversight efforts, local

266 See supra Part III.C; see infra note 267.
267 As Elaine Jardine explained:

I’d really like to get into meeting with more [gas companies]. Because they don’t necessarily have to comply with [wellhead protection ordinances] . . . . We really want to establish some kind of relationship with the gas companies so they may decide to comply with a municipal wellhead protection ordinance [and] we would like to see some of these critical things that are needed in the community followed, too—like if a fire station needs to have a ladder truck, none of our fire stations have ladder trucks—so you know have the gas company pay for that, and community benefit type things.

Interview with Elaine Jardine, supra note 72.

268 According to Elaine Jardine:

I don’t get much support from the legislature. It’s a combination of funding, it’s not believing in planning—and they just don’t want to tackle the issue of natural gas right now . . . . [And the planning department] will be the main point of contact for gas industries’ dealing with the municipalities and we’re trying to set that up now . . . The associate planner I used to have went to Chesapeake [Energy] to be a municipal relations person.

Id. As Andrew Fagan put it: “[n]ational and international companies—they’re so excited about this play and they’re saying it’s the second largest in the world, but are we ready for it? Are we ready for these companies to descend upon us, and can we trust them?” Interview with Andrew Fagan, supra note 213. See supra Parts III.C–D.

269 See supra Part III.C.
Fracking and Federalism

staffs are inadequate to conduct inspections that would ensure that companies—who may be unfamiliar with desired or mandated local practices, given variation from town to town—are heeding regulations and ordinances. If noncompliance is detected, local enforcement power does not appear sufficient to induce adherence to laws. Furthermore, at the state level, an insufficient number of wellpad inspectors can lead to severely reduced checks on drilling and production activities, which may make it more likely that harm-generating errors will occur. If harm does occur, lack of oversight may make identification of the responsible party difficult, given the number of component processes that make up hydrofracking.

New York’s experience also illustrates some of the difficulty of drawing jurisdictional lines around complicated environmental problems that have corollary economic benefits. New York is proceeding to set its jurisdictional lines by declaring what hydrofracking’s impacts are likely to be, by establishing permitting and policy standards in its SGEIS to respond to these impacts, and by restricting local authority. Yet the defects in New York’s

270 Telephone Interview with Wendy Walsh, supra note 68.
271 As Judith Quigley explained, “if there is a disagreement over something, gas companies are the ones with deep pockets ... and of course companies will battle if [an issue] goes to court,” while towns and villages simply do not have money in the coffers or legal staff to fight any opposition. Telephone Interview with Judith Quigley, supra note 70. Furthermore, as Andrew Fagan explained, “[i]f it is just a gentlemen’s agreement, where [companies] just agree to a checklist, and there’s no punitive actions that will happen, then we have to be ready for the consequences.” Interview with Andrew Fagan, supra note 213.
273 The presence of a series of subcontractor relationships can complicate the process of identifying the responsible party in a spill or disaster. See, e.g., the controversy over whether British Petroleum, TransOcean, or Halliburton could be held accountable for the oil leak in the Gulf of Mexico of the spring and summer of 2010. See Erika Boldstad, BP, Transocean, Halliburton Will Blame One Another for Spill, MCCLATCHY (May 10, 2010), http://www.mcclatchydc.com/2010/05/10/93868/bp-transocean-halliburton-will.html.
274 See supra Parts III.B–C.
proposed scheme suggest that state regulators have downplayed or failed to anticipate fully the gravity of uncertain impacts, which has led to inaccurate characterization of problems and poor allocation of authority.\textsuperscript{275} Furthermore, because DEC regulators are not responsible for the kinds of issues that first responders at the local level face, such as waste disposal, road degradation, emergency response, or even local air quality, problems within the jurisdiction of localities are likely to prove beyond any local regulator’s control or resources to address.\textsuperscript{276} Moreover, the uncertainty over whether hydrofracking should be viewed as presenting problems rather than opportunities has fostered ambivalence that colors local elected officials’ willingness to cast hydrofracking as a policy priority.\textsuperscript{277} Thus, many of the attempts to address potential harms have been initiated by unelected officials who are concerned about harm and feel they should prepare now in order to help forestall foreseeable problems.\textsuperscript{278} Unelected local officials effectively are left with a Hobson’s choice—they can either curry favor with energy companies and establish voluntary agreements and risk considerable harm, or push the boundaries of what might be permitted by law, such as attempting to zone out hydrofracking, only to find they have overstepped their power and in the process have alienated and lost the ability to work with and extract concessions from gas producers.\textsuperscript{279} In addition, confusion about jurisdiction with regard to foreseeable harms may lead local decisionmakers to be even more reluctant to grapple with

\textsuperscript{275} See supra Part III.B.

\textsuperscript{276} See supra Part III.C; Drilling Down, supra note 74.

\textsuperscript{277} See supra Part III.C. Local legislators have not been focused on devising plans to prepare for fracking “saying it’s not here yet and we don’t know how long it’s going to take.” Interview with Elaine Jardine, supra note 72.

\textsuperscript{278} See supra Part III.C.

\textsuperscript{279} As Andrew Fagan explained:

I think the tricky part, for anyone looking at [hydrofracking], is - okay who has jurisdiction over what piece of this [issue], and what is this piece anyway? How are [we] supposed to stay on top of this? What’s the responsibility of towns? What’s the responsibility of counties? What’s the responsibility of individuals?

Interview with Andrew Fagan, supra note 213.
unforeseeable and novel problems that arise. Drawing strict jurisdictional boundaries around an environmental issue without fully addressing the scope of possible problems or recognizing the asymmetry between local and industry resources creates a danger that the drawn boundaries will prove arbitrary or fail to account for the broad scope of consequent harms. Thus, setting strict jurisdictional lines—especially where, as here, there is pressure to underregulate—might leave designated regulators unprepared, while responsibility for handling many intractable problems will be pushed off onto those who are not politically accountable.

Besides the potential for environmental and public health harms, New York’s experience with hydrofracking suggests that even apparently well-matched jurisdiction can hamper economic growth. Fear that New York’s regulations are inadequate in light of federal exemptions, combined with widely publicized accounts of hydrofracking’s harms, has led to a high level of opposition from some groups. Vocal opposition has delayed the release of final

---

280 Id.

281 Both from the energy industry and interested property owners. See supra Part II.C. In addition, we might presume the pressures that typically have been seen to drive state regulators are operative here. For discussion of these forces, see Contextual Environmental Federalism, supra note 16, at 120–22; Facts Are Stubborn Things, supra note 122; State Environmental Standard-Setting, supra note 122.

282 Elaine Jardine expressed this point:

We need that natural gas as a nation and as consumers, so I can see both sides. It’s a two-edged sword. The impact is [going to be major] because of all of the incidental impacts [but if] farmers get good payment so the farm will most likely stay in farming . . . I see it from both sides of the coin. It’s just a wait and see game I guess. We can’t do much about it.

Interview with Elaine Jardine, supra note 72; see also Baca, supra note 272 (quoting former DEC Commissioner Grannis as understanding “both sides” of the issue and discussing the DEC’s dual mandate of environmental protection and promoting gas and oil exploitation).

283 Some of the most vocal opposition to hydrofracking has coalesced around the documentary Gasland, which won an award for best documentary at the Sundance Film Festival and was nominated for an Academy Award, and chronicles the experience of landowners and neighbors of hydrofracking activity. See About the Film, GASLAND: A FILM BY JOSH FOX, http://gasland
regulations in New York\textsuperscript{284} and may result in an outright ban if sufficient public opposition is sustained or continues to grow.\textsuperscript{285} Even if New York proceeds to set regulations and allows hydrofracking to commence, its insufficient oversight capacity is likely to limit the extent of production, as DEC’s current strategy is to issue permits only in proportion to oversight capacity.\textsuperscript{286} Thus, hydrofracking shows that the matching principle can lead to


\textsuperscript{284} See Sickle, \textit{supra} note 189.

\textsuperscript{285} See Comments by N.Y. Gov. David Paterson (WAMC radio broadcast Nov. 24, 2010) (“At this point, I would say that the hydrofracking opponents have raised enough of an argument to thwart us going forward at this time.”). Former Gov. Paterson also issued an executive order suspending the approval of some permits until after July 2011, although the efficacy of this order is doubtful given that permits will not be issued before DEC completes its review of hydrofracking, not expected to occur before the order ceases to have effect. Sasha Chavkin, \textit{Executive Order Suspending Fracking Brings Little Change}, PROPUBLICA (Dec. 17, 2010, 1:23 PM), http://www.propublica.org/article/executive-order-suspending-fracking-brings-little-change. Of course, New York’s experience may not be representative of how the hydrofracking debate will play out in the rest of the country. Downstate residents are concerned primarily with protecting their water supply and can be out of touch with the kinds of problems communities upstate face that make them amenable to hydrofracking. See, e.g., Eric Engquist, \textit{The New Gold Rush}, CRAIN’S, (Nov. 9, 2009, 5:59 PM), http://www.crainsnewyork.com/article/20091109/SMALLBIZ/311019959# (contrasting opposition to hydrofracking by Mayor Bloomberg and organized NYC residents with benefits from hydrofracking accrued to upstate residents). It may be unlikely that public opposition will grow significant enough to stanch hydrofracking activity in other states that do not have such a polarized makeup.

\textsuperscript{286} In an interview, former Commissioner of the DEC, Pete Grannis explained:

\textbf{Baca, \textit{supra} note 272.}
Fracking and Federalism

apparent inefficiency, by its own terms.

B. An Adaptive Approach to Environmental Federalism

Hydrofracking demonstrates the desirability of an adaptive approach to environmental federalism that includes an active role for the federal government and regulatory flexibility both within applicable federal agencies and between state and local actors to respond to changing information and circumstances. The network of federal environmental laws is critical to prevent and check harms from environmental problems that might immediately seem to have only localized effects. Each environmental law is designed to allocate burdens of risk and precaution differently and to place those burdens onto industry, localities, states, and the federal government in nuanced ways. Layers of accountability, liability, and penalty institute procedural nodes that incentivize industry actors to act with greater precaution where they propose to engage in risky activity. Greater precaution would undoubtedly lead to a reduced risk of harm, and could mollify opposition that stands as an obstacle to production in the face of uncertainty. Thus, while federal laws might impose upfront and ongoing compliance costs for energy companies seeking to exploit the Marcellus Shale in New York, their application would likely lead to more satisfactory outcomes for a larger number of stakeholders.

One criticism of federal environmental laws that adaptive federalism addresses is that federal regulation suffers from path dependence and poor responsiveness to local concerns and conditions. To some extent, hydrofracking proves these criticisms have merit. For example, the way the EPA currently

---

287 See supra Part III.A.
288 See GLICKSMAN ET AL., supra note 97, at 72–78 (describing the design of environmental laws and how functions, enforcement, and risk and burden allocation vary); Harnessing the Benefits of Dynamic Federalism, supra note 19, at 178–83 (arguing that federal regulation provides valuable checks that provide a better safety net within the context of environmental problems).
289 Asymmetrical Regulation, supra note 17, at 1595 (noting that “the first significant regulatory action on an issue may sit unrevised for years, long after the state of the art has changed”).
defines a “stationary source” under the CAA leaves hydrofracking operations just out of regulatory range. Of course, the EPA has been vested with the discretion to exercise considerable regulatory flexibility to change how it defines terms, such as what constitutes a “stationary source,” in order to respond to evolving problems. However, under an adaptive approach to federalism, some path dependence is not detrimental to the objective of environmental protection; rather, a degree of overarching rigidity can provide structure. Applying federal laws in a consistent manner across industries—rather than based on arbitrary exemptions—would provide the regulatory system with greater stability, even as regulators strive to increase their capacity to respond to evolving circumstances.

Hydrofracking illustrates how regulators may use information to increase their responsiveness to local conditions and improve the adaptability of the federal environmental law framework. Because the EPA has been essentially directed to leave the field, community groups, localities, and states have been compelled to gather and generate information about harm and problem-solving strategies and to disseminate this information across state lines.

290 See supra Part III.A.


292 Adaptive Federalism, supra note 16.

293 Note, however, that Congress has directed EPA to study hydrofracking more closely, which the agency has begun to do, releasing a draft study plan on February 8, 2011. See Hydraulic Fracturing, U.S. EPA, http://water.epa.gov/type/groundwater/uic/class2/hydraulicfracturing/index.cfm (last visited Feb. 12, 2011).

Fracking and Federalism

For instance, much of the information that the Tioga County Attorney has been using to develop a model voluntary road use agreement was obtained by the attorney from neighboring Broome County, who, faced with a local information deficit, traveled to Texas to obtain information about the kinds of issues that have arisen there due to gas production. However, information applicable to one state or region does not always translate well to another. Furthermore, information that has been publicly compiled in an ad hoc manner might not be reliable, with a good proportion subject to interest-group and media distortions.

posting positive articles about hydrofracking and alerts about efforts to limit or ban the practice). There are many more such community-operated websites and blogs, as well as more journalistic efforts, like ProPublica’s investigative reports on hydrofracking trends nationwide. See supra notes 63 & 64. In terms of the kind of general information generation that has occurred, and how unruly it can be, note that as of February 27, 2011, an Internet search for “hydraulic fracturing” and “New York” turned up 155,000 search results. Hydraulic Fracturing New York, GOOGLE, http://google.com (search “hydraulic fracturing new york”) (last visited Feb. 27, 2011). To see how EPA is likely better situated to aggregate and analyze information about hydrofracking, see Drilling Down, supra note 74 (comparing state information and data on water contamination with internal EPA documents that provide a fuller picture of hydrofracking’s harms to water bodies).

295 Telephone Interview with Judith Quigley, supra note 70.

296 Hydrofracking effects are likely to differ across states depending on hydrology and geology (i.e. aquifer depth versus depth of drilling), as well as political and economic variation. See LISA SUMI, SHALE GAS: FOCUS ON THE MARCELLUS SHALE (2008), http://www.earthworksaction.org/pubs/OGAPMarcellusShaleReport-6-12-08.pdf (discussing drilling or potential for drilling in several different shale formations across the country and suggesting how differences might lead to different impacts).

297 Accuracy or validity of either position aside, compare Hydraulic Fracturing Facts, HYDRAULIC FRACTURING.COM, http://www.hydrofracturing.com/Pages/information.aspx (last visited Mar. 14, 2011) (website that appears to be operated by Chesapeake Energy asserting that “[p]roperly conducted modern hydraulic fracturing is a safe, sophisticated, highly engineered and controlled procedure”) with John Zeiger, No Fracking Way: Ban Hydrofracking in New York, ONEARTH.ORG, http://www.onearth.org/node/1865 (last visited Mar. 14, 2011) (NRDC-published community blog post describing SGEIS as “possibly the worst document [the DEC] has ever written” and positing that “one could make the argument that it was practically written by gas
Indeed, public concern about whether anecdotal harms can be traced to hydrofracking has already led the EPA to commence a new study of the practice.\textsuperscript{298} The agency now has the opportunity to develop and respond to emerging information in a comprehensive fashion, creating a reliable record that addresses public fears, as this Note and other authors recommend.\textsuperscript{299} With more reliable information, regulators at all levels of government would be better situated to make decisions about whether or how to regulate hydrofracking.\textsuperscript{300} Regulatory flexibility to respond to company lobbyists\textsuperscript{\textsuperscript{298}}. Note that uncertainty about what harms are likely to occur that results from concern over accuracy of publicly available information has placed local officials who seek to cultivate objectivity in an awkward position, exacerbating their difficulties in preparing for incidental impacts; as Andrew Fagan explains:

When you see movies like Gasland and other things, you need some way to view those things through the lens of objectivity if you don’t have all the knowledge then you get caught up in the emotional . . . . So that’s been a real challenge, and the media doesn’t help . . . . And right now we are in a position that we can learn from mistakes that are happening in Pennsylvania, [but] in Pennsylvania the companies are showing all of these regulations they already have to follow, and DEP is happy about how they are monitoring things, and how things are going, but yet you hear about these number of concerns or issues and you keep trying to weigh that out in your mind, and saying “who do you believe?”

Interview with Andrew Fagan, supra note 213.

\textsuperscript{298} The EPA is engaged in producing a report on hydrofracking, to be released in 2012. See Hydraulic Fracturing, supra note 293. An EPA report released in 2004, which concluded that hydrofracking could not be conclusively linked to purported harms, has been widely criticized. See generally LISA SUMI, OIL AND GAS ACCOUNTABILITY PROJECT, OUR DRINKING WATER AT RISK: WHAT EPA AND THE OIL AND GAS INDUSTRY DON’T WANT US TO KNOW ABOUT HYDRAULIC FRACTURING (2005), available at http://www.earthworks action.org/pubs/DrinkingWaterAtRisk.pdf.


\textsuperscript{300} Allowing the EPA to play a pivotal role in providing and ensuring reliability of information about possible environmental harm can result in improved policy choices about new technologies, like hydrofracking, over time. Applying existing federal laws would help prevent some of hydrofracking’s
unique problems would increase as uncertainty that hinders
decisionmaking or incentivizes regulators to discount harm is
lessened.

Although matching principle proponents would likely argue
that wholly intrastate municipal drinking water supplies can be
more appropriately and efficiently regulated at the state or local
level, an adaptive approach to federalism supports the idea that a
federal environmental law like the SDWA is equally justified as
those that target clearly interstate resources like air and navigable
waterways. Analysis of the SDWA shows that applying the
matching principle and fixing a static regulatory position that
drinking water is always better left to localities or states can result
in underprotection. Water resources that states and localities can
protect most efficiently under “normal” conditions may be left
vulnerable when a new kind of industry, such as hydrofracking,
arises that has unprecedented and formerly unforeseeable
impacts. Towns can only use devices like weak wellhead
protection ordinances to prevent pollution of communal water
supplies. And as in Tioga County, where one aquifer supplies a
whole county, a wellhead protection plan will not forestall
contamination that occurs outside the borders of the municipality
that has enacted the plan. If the SDWA were to apply to
hydrofracking, federal enforcement and oversight resources would

most dramatic harms but would not remove jurisdiction from state and local
regulators. See supra Part III.A. Thus, federal regulators would still be able to
learn from innovations at lower levels of government and from industry
responses to different policies. The federal government could consciously devise
and adopt iterative strategies, finding ways to moderate what industry perceives
as draconian state postures in some instances, while inhibiting overly permissive
schemes in other states. See Carlson, supra note 139 (presenting a discussion of
iterative opportunities created when federal and state regulatory differences are
played off of one another in the context of vehicle emissions standards).

301 See Adler, supra note 128, at 157–60.
302 See Drilling Down, supra note 74.
303 Interview with Elaine Jardine, supra note 72.
304 Id. See also Clinton Street-Ballpark Valley Aquifer System Broome and
(final determination).
305 Interview with Elaine Jardine, supra note 72.
offer a higher level of assured protection. Not only would the federal government carry a bigger stick to ensure compliance, but federal intervention would cut out the formidable costs of establishing and revisiting baseline standards where the nature of novel harms is not well understood.\footnote{See e.g., Part III.C.} Furthermore, forcing energy companies to implement technologies that will prevent harm to municipal water supplies could have the beneficial impact of preventing harm to unregulated water sources, like private wells, as industry operating standards improve in order to achieve compliance with federal regulations. Hydrofracking shows that even federal laws that seem intrusive at the local or state level may be necessary to ensure proper protection of critical commonly held resources. Thus, allowing federal laws to provide a stable framework while increasing responsiveness within that framework will lead to better protection.

\textit{C. Avoiding Regulatory Commons Problems}

In addition to demonstrating that an adaptive approach is preferable to a matching approach, New York’s experience with hydrofracking sheds light on some aspects of operative incentives within the regulatory commons. As in New York, states may have conflicting regulatory mandates, which result in incentive confusion and distortion. For example, the DEC’s goal is to promote the maximum extent of exploitation of gas resources while protecting correlative rights and the environment—goals that are in conflict.\footnote{See e.g., Part III.C.} Giving the states primacy over environmental problems while tasking them with striking a balance between economic growth and environmental protection may make it less likely they will have incentives to include localities in permitting and land-use decisions.\footnote{N.Y. Comp. Codes R. & Regs. tit. 6 § 550.1 (2010); About DEC, N.Y. St. Dep’t of Envtl. Conservation, http://www.dec.ny.gov/24.html (last visited Apr. 7, 2011). For Commissioner Grannis’ perspective on these conflicting directives, see Baca, supra note 272.} In an atmosphere where there is pressure

\footnote{Especially where there is evidence that state regulators are already more growth-oriented than federal regulators to begin with. See Contextual}
Fracking and Federalism

to compromise, states have incentives to stifle legitimate opposition; they can easily characterize such opposition as “NIMBYism” and dismiss it. For example, in New York, meaningful opportunity to express local opposition and participate in critical decisions over how hydrofracking will occur is foreclosed by the State’s usurpation of traditional zoning power and local SPDES and SWPPP permitting processes. Contracted authority leads to jurisdictional insecurity at the local level; meanwhile, the State has not provided adequate mechanisms for protection. However, if federal law were to apply, the push to begin hydrofracking would be slowed dramatically as energy companies would be forced to comply with federal permitting processes. State and local regulators would have more time and more resources to focus on and refine strategies best suited to their historic expertise, such as cooperative compliance with the federal environmental laws, land use and natural resources planning, and municipal zoning. Shared regulatory responsibility and a slowed pace could help diminish the high stakes atmosphere that currently drives New York to stifle meaningful local input, and would lead to enhanced deliberation and increased jurisdictional confidence overall.

Finally, the pervasive role of compromise in preference setting at the local level suggests that avoiding regulatory commons problems involves awareness of differences in risk tolerances across the three tiers of government. As is occurring in New York, states and localities may be bound to accept higher levels of risk than many citizens prefer because the combination of profits and harm-prevention strategies that rely on industry skew incentive structures and align risk tolerances with industry preferences.
addition, local actors have little choice but to accept an activity, even if they would prefer to keep it out of their environs. By contrast, regulators at the federal level are more insulated from immediate risks of harm, direct benefits, and complicated mixtures of the two. Thus, federal regulators may be better poised to set and enforce baseline standards that can prevent any group or community from accepting unacceptably high levels of risk. In addition, varying the level of federal oversight according to type of activity rather than the type of harms an activity causes—as is occurring with hydrofracking—results in inconsistent application of laws that adds to jurisdictional confusion.

While adding another layer of regulatory authority will introduce a new group of actors, with their own incentive structures that could arguably add to jurisdictional confusion, a federal role seems nonetheless necessary to ensure adequate protection. Where state and local oversight and enforcement resources are not adequate to meet an environmental law challenge, a federal role would assure that energy companies exercise a minimum level of environmental precaution. Local governments would then not be compelled to enter unenforceable agreements that result in the same, or likely far less, protection. Local governments could spend their scarce resources on initiatives that contribute to the long-term stability of their communities rather than on scrambling to prepare or on negotiations with energy companies with short-term and myopic negative and anti-fracking doesn’t help because then we won’t get the help of the energy companies.” Telephone Interview with Dick Le Count, supra note 71. For analysis of dynamics of risk allocation and economies of scale, see Contextual Environmental Federalism, supra note 16, at 121–22.

313 See Contextual Environmental Federalism, supra note 16, at 111–12 (advancing an “economy of scale” argument for federal jurisdiction on the basis it is better poised to balance risk tolerances).

314 See id.

315 E.g., the exemption of hydrofracking, a construction activity that creates a high level of disturbance, from the state SWPPP program seems arbitrary, given that local regulators generally have input into permitting of local construction activities.

316 See Contextual Environmental Federalism, supra note 16, at 126.

317 See supra Part III.C.
interests. This outcome seems entirely consistent with the purposes of federalism—that responsibilities at each level of government are correlated to resources and expectations about accountability.

Application of federal laws along with state and local regulations can also play an important signaling role by setting normative priorities. These priorities serve to reinforce clarity over, and confidence in, regulatory authority to meet novel challenges. As in the hydrofracking example, failure to apply environmental laws in a consistent fashion decreases regulator assurance in jurisdictional authority. Because it seems clear that the federal government is steadfastly interested in developing natural gas resources, an increased federal role in the regulation of hydrofracking is not likely to result in insurmountable barriers to production and may provide legal, conceptual, and normative stability that is key to avoiding regulatory commons problems.

CONCLUSION

New York’s attempt to regulate hydrofracking has provided an opportunity to scrutinize the appropriateness of various approaches to environmental federalism. At first glance, hydrofracking seems like an activity that could be best regulated at the state and local

---

318 Elaine Jardine explained “[Hydrofracking] takes up a good chunk of my time.” Interview with Elaine Jardine, supra note 72.
319 See Contextual Environmental Federalism, supra note 16, at 128.
320 See Stewart, supra note 120, at 1210–11.
321 See supra note 279.
322 See supra Part III.A (describing extent of Congressional support for exemptions for hydraulic fracturing).
323 See supra Part III.A. The most aggressive Congressional attempt to address hydrofracking to date has sought to remove the exemption for hydraulic fracturing under the SDWA. Fracturing Responsibility and Awareness of Chemicals Act of 2009, H.R. 2766, 111th Cong. (2009); Fracturing Responsibility and Awareness of Chemicals Act of 2009, S. 1215, 111th Cong. (2009). However, neither the House nor the Senate bill made it out of committee. See Bill Summary and Status, THOMAS (LIBRARY OF CONGRESS), http://thomas.loc.gov/cgi-bin/bdquery/z?d111:h.r.02766: (last visited Apr. 12, 2011).
level, given that its most immediate effects are localized and intrastate. Therefore, hydrofracking and New York State ostensibly make a good regulatory match, as Congress has mandated.

However, as analysis of New York’s experience shows, matching approaches rely on what are perhaps overly simplistic assumptions about the roles that risk and uncertainty play in setting and satisfying preferences. In New York, regulators have been required to balance interests that are easy to price and relatively certain against those that are difficult to quantify and uncertain; further, they have been under pressure to make sensitive decisions about important trade-offs quickly. With little regulatory control, localities have been forced to align their interests with industry interests in attempts to assure safety and maximize benefits, and then hope for the best. Existing state and local laws are likely to prove inadequate to prevent or redress harms from hydrofracking, should they occur, because oversight and enforcement resources are limited. In turn, concern over state and local capabilities to address harms itself threatens to inhibit gas production. Thus, hydrofracking in New York suggests that state primacy—and thus the matching principle—may not result in adequate public and environmental protection or optimal economic efficiency.

New York’s experience with hydrofracking underscores the need for an adaptive framework even with regulatory commons risks. A federal role will lead to some costs, which include initial delays in fracking activity and increased spending to achieve compliance with federal laws. However, underregulation can

324 See supra Part IV.A.  
325 See supra Part III.A.  
326 See supra Part IV.  
327 See supra Parts II–III.  
328 See supra Part III.C.  
329 See supra Part III.C.  
330 See supra Part IV.A.  
331 See supra Part IV.  
332 See supra Parts IV.B–C.  
333 See supra Part IV.A.
Fracking and Federalism

lead to unnecessary, tragic, and irreversible costs borne by those populations least equipped to bear them.\textsuperscript{334} A vigorous and responsive framework of laws with roles for each level of government would increase jurisdictional confidence and eliminate or mitigate uncertainty about risk.\textsuperscript{335} Federal regulation would also prevent hindrance of beneficial economic activity by ensuring that regulators and the public have access to accurate and adequate information and by providing necessary oversight and enforcement resources.\textsuperscript{336} Thus, allowing an active, yet adaptive, federal role in the regulation of activities that threaten to cause substantial harms will better protect public health and the environment as well as the virtues and values of federalism.\textsuperscript{337}

\begin{flushright}
\begin{flushright}
\footnotesize
334 \textit{See supra} Parts II.C & III.B.
335 \textit{See supra} Parts IV.B–C.
336 \textit{See supra} Parts IV.B–C.
337 \textit{See supra} Parts IV.B–C.
\end{flushright}
\end{flushright}
DEFENDING PAID SICK LEAVE IN NEW YORK CITY

Rebeccah Golubock Watson*

INTRODUCTION

Maximino Santos, a chef at a New York City deli, developed colds easily. But he often went to work sick for fear of losing his job. In order to protect the food he was preparing from his coughs, he bought face masks with his own money. Many employees like Mr. Santos also do not take time off because they cannot afford it. For Mr. Santos, and many other employees, serious consequences follow: Mr. Santos developed pneumonia, and others are fired for missing work.

No law required Mr. Santos’ employer to allow him to take time off from work when he or a family member was sick. Like

---

* J.D., Brooklyn Law School 2012, B.A., Harvard University, 2004. I would like to thank Paul Sonn of the National Employment Law Project and Sherry Liewant of A Better Balance for their comments on drafts and their commitment to this work. I owe a debt of gratitude to Priti Trivedi, Bill Spirer, and the entire staff of the Journal of Law and Policy for their cheerful and tireless editorial assistance. Finally, deepest thanks to my family, especially to my mother, Carol Golubock, a fearless advocate for workers’ rights, and my father, Arthur Watson, her partner in crime.


2 Id.

3 Id.


42% of all U.S. private sector employees, Mr. Santos has no right to take sick leave. However, San Francisco, Washington, D.C. and Milwaukee have passed local laws that require all businesses within their jurisdiction to provide some paid sick days for their workers. Inspired by these measures, and concerned about the public health implications of an estimated 1.3 million workers in New York City lacking even a single day of paid sick time, Council member Gale Brewer introduced the Paid Sick Time Act in August, 2009 with thirty-six co-sponsors in a fifty-one-member City Council. Despite the bill’s veto-proof majority, Speaker Christine Quinn decided not to bring it to a vote, citing concerns about the bill’s effect on small businesses. Supporters of the bill have vowed to continue to push for its

---


10 Paid Sick Leave Ordinance, MILWAUKEE, WIS. CODE OF ORDNANCES § 112.

11 REISS ET AL., supra note 7, at 4.

12 The bill was re-introduced in March 2010 when the next session of the Council convened. Paid Sick Time Act, N.Y.C. Council Int. No. 0097-2010 (introduced Mar. 25, 2010), available at http://legistar.council.nyc.gov/Legislation.aspx (enter “Int 0097-2010” into “Search” box; select “All Years”; then select “Search Legislation.”).

The Paid Sick Time Act, along with its D.C., San Francisco, and Milwaukee counterparts, is illustrative of a broader phenomenon: localities crafting innovative legislation pursuant to their home rule power. Many cities and municipalities have relied on their home rule power—authority delegated to cities and municipalities to pass substantive laws affecting those within their borders—when enacting a wide range of legislation, from alternative voting systems, campaign finance reform, minimum wage floors, and gay rights legislation. This trend has prompted scholars to liken localities to “laboratories of democracy,” a phrase Justice Brandeis used to refer to the states. Often these local enactments seek to regulate businesses by imposing additional costs or new guidelines. The result is that businesses are likely to challenge local laws on federal and state law preemption.
grounds. As one scholar has commented, “[t]he examples are legion, but the story is familiar: when a city adopts a new policy that differs from state law and may harm some segment of the business community, a preemption challenge is almost certain to follow.”

In anticipation of such a challenge, this Note addresses whether the New York City Council has the authority to pass the Paid Sick Time Act, pursuant to the New York Municipal Home Rule Law, and whether state and federal laws governing employee benefits would preempt this local proposal. Part I of this Note establishes the need for paid sick leave by exploring the gaps left by current federal and state law addressing employment leave. Part II discusses the parameters of New York home rule doctrine, and lays out the principles of state law preemption of local law, arguing specifically that the Paid Sick Time Act does not conflict with the New York Minimum Wage Act. Part III analyzes federal preemption doctrine and asserts that neither the Family and Medical Leave Act (FMLA) nor the Employee Retirement Income Security Act (ERISA) preempts the Paid Sick Time Act.

I. NO PROTECTION FOR SICK WORKERS

A. The Need for Paid Sick Leave

No federal law guarantees even a minimum number of paid sick days for U.S. workers to use for ordinary illness for themselves or their children. Advocates for paid sick leave argue that workers should have a fundamental right to take time off to take their children to the doctor or care for short-term, commonplace illnesses. It is essential that this leave be protected by law, they say, to ensure that employees will not be fired or penalized for taking time off to see the doctor or care for their sick

22 Diller, supra note 15, at 1114.
23 Id. at 1115.
Paid Sick Leave in New York

children. Furthermore, advocates argue that paid sick leave is necessary from a public health standpoint because employees like Mr. Santos need time off to recuperate from illnesses that, if not treated preventatively or early on, would only recur and prevent the employee from working or affect others in the workplace. Moreover, sick leave is not only necessary for the health of the individual employee, but for her “co-workers, customers, classmates, and in turn, their families.”

Statistics reflect the dire situation many workers face when it comes to paid sick leave. Only about half of private employees are provided sick leave by their employer, and such leave is usually in the form of accrued time off that employees use for their own non-work-related illness. Paid sick leave is out of reach for over one-third of women in workplaces of fifteen employees or more. Access to paid leave divides along class lines as well. Not surprisingly, paid sick days are often not available to low-wage workers. Of the top quartile of American workers in terms of wages, 81% have paid leave, when only a third of the lowest quartile of workers do. This deep discrepancy may be one of the reasons why there is now a growing urgency for paid leave. Many Americans may take their paid leave for granted, and therefore fail to appreciate the toll an unpaid day of work could take on a


28 REIS ET AL., supra note 7, at 16.


30 Id. at 6.


32 Id.
moderate or low-wage worker. In 2000, three out of four workers who reported that they needed leave did not take it because they could not afford to lose wages. New studies support advocates’ contention that paid sick leave promotes public health. A new report by the Restaurant Opportunities Center found that nearly 90 percent of all restaurant workers in New York City report having no paid sick days, and that over 63 percent report having cooked or served food while sick.

Little progress has been made, both federally and at the state level, to address the need for paid sick days. Six states—Hawaii, Maine, Minnesota, Oregon, Washington, and Wisconsin—allow workers who already have paid sick days for personal illness to use them to care for certain family members as well. At the federal level, Representative Rosa DeLauro and Senator Ted Kennedy introduced the Healthy Families Act in 2009, which would require employers to provide seven paid sick days to employees who work more than thirty hours a week. The bill died in committee in the 111th Session of Congress, and has not yet been introduced in the 112th Session.

37 MINN. STAT. § 181.9413 (2010).
40 WIS. STAT. § 103.10 (2010).
42 S.1152 § 5(a)(1). The bill allows employees to take fifty-six hours of leave, which amounts to seven eight-hour workdays. Id.
43 Search Bill Summary & Status, THOMAS, http://thomas.loc.gov/home/LegislativeData.php (select “111th Congress,” type “Healthy Families Act” into the search form, click “search” and then choose “Healthy Families Act” from
Paid Sick Leave in New York

B. Family Leave: A Different Animal

The only law that comes close to addressing workers’ need for sick leave is the FMLA.\(^{44}\) But the FMLA only covers serious illnesses or childbirth, is available only to full-time workers, and, perhaps most importantly, is unpaid.\(^{45}\) Even those Americans who are eligible for the FMLA are not protected by law for taking time off for short-term illnesses. In addition, they are not paid for time taken.

The FMLA requires employers with fifty or more employees to grant up to twelve weeks of job-protected unpaid leave to employees in the event of the birth or adoption of a child or a serious health condition affecting the employee or the employee’s child, spouse, or parent.\(^{46}\) To be eligible for the leave, employees must have been employed for at least twelve months by the employer, with at least 1,250 hours of service for that employer over the past twelve months.\(^{47}\)

The FMLA has problems of its own. While beneficial to some, it fails to adequately cover broad categories of workers; the benefits extend to only half of the U.S. workforce.\(^{48}\) The FMLA disproportionately benefits middle- and upper-income workers who can afford to take unpaid leave.\(^{49}\) More than half of all low-wage workers are excluded from the FMLA, as they are employed by small businesses. In addition, low-wage workers are less likely to satisfy the FMLA’s yearlong employment and hour requirements, as more than half of them are part-time and spend less than a year on the job.\(^{50}\) Finally, part-time workers are often denied benefits under the FMLA because they have yet to work the requisite 1,250 hours.\(^{51}\) This in turn disproportionately hurts low-

---


\(^{45}\) §§ 2611-2612.

\(^{46}\) Id.

\(^{47}\) § 2611(2)(A).

\(^{48}\) Lester, supra note 29, at 2.

\(^{49}\) O’Leary, supra note 33, at 39.

\(^{50}\) Id. at 44.

\(^{51}\) Id.
wage women, since women comprise two thirds of the part-time workforce. Many family leave advocates anticipated this result, predicting that the FMLA would be “a shadow benefit” for many workers as long as the leave was unpaid.

But even if the FMLA were strengthened to include part-time employees, or those who have been with the employer for less than a year, the problem remains that it is limited to long-term leave for childbirth or serious illness, excluding the many common, short-term sicknesses that plague American workers and their families and it does not require paid leave.

C. New York City’s Paid Sick Time Act

The Paid Sick Time Act would allow employees to take paid sick leave for themselves and to care for family members. It is designed to provide paid sick leave for workers without overly burdening employers. The Paid Sick Time Act would require employers to allow employees to accrue sick time based on hours worked. For every thirty hours worked, the employee would gain one hour of paid sick leave. Although the employee would begin accruing sick leave once her employment begins, she could not use her accrued time until she has worked for ninety days. The bill would impose a cap on the amount of paid sick leave allowed: for

52 Id.
53 Emily A. Hayes, Bridging The Gap Between Work and Family: Accomplishing the Goals of the Family and Medical Leave Act Of 1993, 42 WM. & MARY L. REV. 1507, 1523 (2001). It is also worthwhile to note that the FMLA’s unpaid leave makes the United States the least generous industrialized nation when it comes to family leave. Lester, supra note 29, at 3. In addition, the U.S. is the only country among twenty-two countries ranked highly in terms of economic and human development that does not guarantee that workers receive paid sick days or paid sick leave. JODY HEYMAN ET. AL., CTR. FOR ECON. & POLICY RESEARCH, CONTAGION NATION: A COMPARISON OF PAID SICK DAY POLICIES IN 22 COUNTRIES 1 (May 2009), available at http://www.cepr.net/documents/publications/paid-sick-days-2009-05.pdf.
55 Id. § 2(e).
56 Id. § 2(e)(2).
57 Id. § 2(e)(7).
Paid Sick Leave in New York

businesses with fewer than twenty employees, workers would be allowed forty hours of paid sick leave; for larger businesses, workers could take off up to seventy-two hours. An employee could use the leave for her own mental or physical illness, diagnosis, or preventative care, and she could take leave to care for the illness of a spouse, child, parent, grandparent, or domestic partner. The Paid Sick Time Act would also allow employees to take paid leave in the event of a business or school emergency. An employer found in violation of the law would be liable for a minimum civil penalty of one thousand dollars for each violation.

The Paid Sick Time Act seeks to protect workers from losing their jobs when forced to take off work to care for their own short-term illnesses, or to care for a sick family member. It also seeks to promote public health by enabling service workers like Maximino Santos to take time off to stave off a serious illness, and to prevent spreading his illness to others.

II. HOME RULE AND STATE LAW PREEMPTION

A. The Power of Home Rule: Protecting Well-Being

Pursuant to the Constitution of the State of New York (“New York Constitution”) and the New York Municipal Home Rule Law, local governments possess broad powers to enact laws relating to the welfare of citizens within their municipality. The New York Constitution provides that “every local government shall have power to adopt and amend local laws not inconsistent with the provisions of this constitution or any general law.” The scope of this legislative power includes laws relating to government, protection, order, conduct, and the safety, health and

---

58 Id. § 2(c)(8).
59 Id. § 2(d)(1)(i-ii). The bill defines “family member” broadly to include parents-in-law, domestic partners, and the parents of a domestic partner. Id. § 2(b)(7).
60 Id. § 2(d)(iii).
61 Id. § 2(l)(1).
62 See N.Y. Const. art. IX, § 2(c); N.Y. Mun. Home Rule § 10(1)(a)(12).
63 N.Y. Const. art. IX, § 2(c)(i-ii).
well-being of New Yorkers.\textsuperscript{64} New York City’s police powers are further established in the New York City Charter, which provides, in pertinent part, that the City Council “shall have power to adopt local laws which it deems appropriate, which are not inconsistent with the provisions of this charter or with the constitution or laws of the United States or this state.”\textsuperscript{65} These laws must be, \textit{inter alia}, “for the preservation of the public health, comfort, peace and prosperity of the city and its inhabitants.”\textsuperscript{66} However, the New York Constitution mandates that such laws must not conflict with the Constitution’s own terms, or any state law related to “property, affairs, or government,”\textsuperscript{67} and the Municipal Home Rule Law proscribes cities from adopting local laws inconsistent with the state constitution or any general state law.\textsuperscript{68}

New York City has enacted a broad range of legislation aimed at safeguarding the health, safety, and welfare of its residents. For example, New York City has recently required food establishments to post the calorie counts for their food\textsuperscript{69} and provided that building owners, when first taking over a new building, must offer temporary employment to the predecessor employer’s service employees.\textsuperscript{70} New York courts have interpreted home rule power broadly to authorize many of these local enactments.\textsuperscript{71} However, the City has also lost its share of home rule battles. In 1962, a state preemption challenge invalidated the City Council’s minimum wage ordinance,\textsuperscript{72} and in 2005, the Appellate Division found that

\textsuperscript{64} Id. § 2(c)(10); N.Y. MUN. HOME RULE § 10(1)(a)(12).
\textsuperscript{65} N.Y.C. CHARTER § 28(a).
\textsuperscript{66} Id.
\textsuperscript{67} N.Y. CONST. art. IX, § 2(c).
\textsuperscript{68} N.Y. MUN. HOME RULE § 10(1)(ii).
\textsuperscript{69} 24 R.C.N.Y. § 81.50 (2009).
\textsuperscript{70} N.Y.C. ADMIN. CODE § 22-505 (2009).
state law preempted the City Council’s legislation requiring that City contracts only be awarded to businesses providing spousal benefits for same-sex partners.73

B. Occupying the Field: New York State Preemption Doctrine

A measure enacted by a New York city or municipality is generally preempted by state law when it expressly conflicts with the state law, or where the State has established an intent to “preempt an entire field”74 or to create “uniformity” in that field.75 Field preemption is evidenced in legislative history, or from a statute’s “comprehensive and detailed regulatory scheme in a particular area.”76 When a state court determines that a state provision occupies the field, it will generally invalidate a local measure that proscribes a practice which the state law “considers acceptable,” even if the state law does not “expressly speak[]” to it, or it limits rights granted in state law.77

In one of the seminal New York home rule cases concerning employment benefits, Wholesale Laundry Bd. of Trade v. City of New York, the New York Appellate Division invalidated a city law setting wage thresholds because the measure regulated a field fully occupied by a state scheme, and in doing so prohibited what the state law allowed.78 The New York City Minimum Wage Act, enacted by the City Council in 1962, required all employers in New York City

---

74 ILC Data Device Corp. v. Cnty. of Suffolk, 588 N.Y.S.2d 845, 849 (N.Y. App. Div. 1992), appeal dismissed, 613 N.E.2d 965 (N.Y. 1993); see also N.Y. State Club Ass’n v. City of New York, 505 N.E.2d 915, 917 (N.Y. 1987) ("[T]he City may not exercise its police power when the Legislature has restricted such an exercise by preempting the area of regulation.").
76 Id. at 566.
77 ILC Data Device Corp., 588 N.Y.S.2d at 850.
to pay a minimum wage of $1.25, and not less than $1.50 beginning one year from the date the law became effective.\textsuperscript{79} At the time, the State’s Minimum Wage Act set a minimum wage of $1 per hour during a two-year time period, and then then provided for specific increases in subsequent two-year periods.\textsuperscript{80} During each such period, the State Minimum Wage Act allowed for an increase in the wage only if another wage was established under the State Minimum Wage Act.\textsuperscript{81} Specifically, the Commissioner of the Department of Labor could investigate the need for higher wages in various occupations, and could then appoint a wage board to make recommendations as to whether a higher wage should be fixed by the Commissioner.\textsuperscript{82}

The local law’s opponents had two main arguments: (1) New York State’s Home Rule Law prevented the City from enacting such a measure, and (2) the local law was invalid because it was “inconsistent with State-wide legislation on the same subject.”\textsuperscript{83} The \textit{Wholesale} court found that it did not need to address the home rule issue because the local law was inconsistent with the State’s Labor Law.\textsuperscript{84} Although the City argued that its law “extends but does not run counter to the State statute,” the court found that the local law made impermissible what would be allowed under the state law.\textsuperscript{85} Moreover, the court held that the state’s minimum wage statute meant to occupy the entire field because it had an “elaborate machinery” in place for shaping wage thresholds.\textsuperscript{86}

New York courts have also vitiated a local law on preemption grounds when it imposed “prerequisite ‘additional restrictions’ on rights under State law, so as to inhibit the operation of the State’s general laws.”\textsuperscript{87} In \textit{Council of the City of New York v. Bloomberg},

\textsuperscript{79} \textit{Id.} at 863.
\textsuperscript{80} \textit{Id.} at 864.
\textsuperscript{81} \textit{Id.}
\textsuperscript{82} \textit{Id.}
\textsuperscript{83} \textit{Id.} at 863–64.
\textsuperscript{84} \textit{Id.} at 864.
\textsuperscript{85} \textit{See id.}
\textsuperscript{86} \textit{Id.} at 865.
Paid Sick Leave in New York

the New York Appellate Division held that New York City’s Equal Benefits Law, which prevented the City from contracting with vendors that did not provide spousal benefits to domestic partners, conflicted with General Municipal Law 103(1) and ERISA.88 General Municipal Law 103(1) requires that the State obtain the best services at a competitive price.89 Because the Equal Benefits Law “expressly excludes a class of potential bidders for a reason unrelated to the quality or price of the goods or services they offer,” the court held that the state statute invalidated the local law.90

Despite the Wholesale ruling, New York courts have been cautious to find field preemption when the state and local measures regulate a similar subject but with different approaches, and the state law does not evince the intent to occupy the field. In People v. Cook, the court considered a challenge to a New York City law that taxed cigarettes in proportion to the amount of tar and nicotine they contained.91 The defendant in Cook argued that the law was invalid because state law permitted businesses to sell cigarettes without a tax differential. The Cook court disagreed, finding that despite the state’s existing general regulation of cigarettes, it had not “spoken through its own laws in relation to price differentials based on tar and nicotine content. The fact that the State also taxes cigarettes has no significant relation to the price-differential aspect of the city’s enactment, and therefore cannot be said to create an inconsistency.”92

The court in Cook also did not accept defendant’s contention that the City was preempted from regulating cigarettes on the basis of nicotine content because the State’s regulation tacitly permitted such conduct. The court found that this interpretation of preemption doctrine was too loose, stating “[i]f this were the rule, the power of local governments to regulate would be illusory. Any


88 Id. at 109–10. See Part III, infra, for a longer discussion of ERISA preemption.
89 Bloomberg, 791 N.Y.S.2d at 109.
90 Id.
91 People v. Cook, 312 N.E.2d 452, 454 (N.Y. 1974).
92 Id. at 457.
time that the State law is silent on a subject, the likelihood is that a local law regulating that subject will prohibit something permitted elsewhere in the State. The court drew a distinction between the circumstances before it, in which the state law was silent on a subject, and other circumstances in which the state has “acted upon a subject, and in so acting has evidenced a desire that its regulations should pre-empt the possibility of varying local regulations,” as in Wholesale Laundry. Under the latter circumstances, the State law preempts the local law. Thus, Cook illustrated a fairly cautious approach towards field preemption under circumstances where there is significant overlap between the two areas of regulation, yet the state has not acted to preempt the entire field.

The Cook court also established important precedent regarding New York City’s home rule authority to protect health, finding that the City is granted the authority to promote health by the New York Constitution, the Municipal Home Rule Law, and the New York City Charter. If the City’s law is properly related to the power to promote health, the court held, there are only two exceptions to its legitimacy: (1) if the local law is inconsistent with constitutional law, or (2) if the Legislature prohibits it. The Cook court relies on a test set forth in Good Humor Corp. v. City of New York, where the New York Court of Appeals vitiates an anti-peddler ordinance “because it was not passed for the purpose of advancing health or general welfare but rather, in the words of the city council, ‘to prevent unfair competition by itinerant peddlers with storekeepers who pay rent and various taxes.’” The Good Humor test requires that the local law have a “substantial relation to a legitimate, authorized purpose” granted to the City. The Cook court found that the City’s new tax structure for cigarettes had a “substantial relation” to health.

93 Id.
94 Id.
95 Id. at 455.
96 Id.
97 Id. at 456–57 (quoting Good Humor Corp. v. City of New York, 49 N.E.2d 153, 155 (N.Y. 1943)).
98 Id. at 457.
Other Court of Appeals decisions adopt Cook’s narrow approach to field preemption. In *Myerson v. Lentini Bros. Moving and Storage Co.*, the court considered a city law protecting consumers from deceptive practices by moving companies despite a state transportation law that regulated moving company rates.\(^9^9\) The court found that the state law, which required companies to obtain a permit and file and publish their rates, fulfilled a different purpose than the local law, which proscribed unethical business practices.\(^1^0^0\) The state law was designed to “foster sound economic conditions in the industry” and to promote “reasonable charges…without unjust discriminations,” which was distinct from the local measure’s effort to curtail specific behavior.\(^1^0^1\)

Furthermore, the court held, the state statute did not establish the intention to occupy the entire field of moving company regulation.\(^1^0^2\)

The Court of Appeals applied a similar approach to preemption in reviewing a more recent challenge to a city law regulating rental car agencies in New York City.\(^1^0^3\) The local provision prohibited rental agencies from determining whether to rent to a customer, or to adjust rental rates, based on the customer’s residence. In *Hertz Corp. v. City of New York*, the court considered several state statutes, one of which established a comprehensive scheme regulating all motor vehicles in the state, including those owned by rental companies. Another provision specifically regulated the rental car industry, prohibiting discrimination based on age, race, gender, and credit card ownership, among other statuses.\(^1^0^4\) Yet none of the provisions, the court held, addressed the specific issue of whether rental companies could deny service or alter rates based on an individual’s residence, and thus the state law did not preempt.\(^1^0^5\)

---

\(^1^0^0\) *Id.* at 807.
\(^1^0^1\) *Id.*
\(^1^0^2\) *Id.*
\(^1^0^3\) *Hertz Corp. v. City of New York*, 607 N.E.2d 784 (N.Y. 1992).
\(^1^0^4\) *Id.* at 786.
\(^1^0^5\) *Id.*
scope or so detailed” to preempt all local rental company regulation.106

The Wholesale, Bloomberg, Cook, Myerson, and Hertz cases offer some (at times contradictory) principles that New York courts follow when determining whether state or federal laws preempt local laws. First, Wholesale stands for the proposition that a local law cannot prohibit what a state or federal statute explicitly permits when the state evinces the intention to preempt the entire field. Under Bloomberg, a local law cannot impose additional “prerequisites” on a state law that limit its scope. However, under Cook, a local law can proscribe something that a state law permits, provided that the State is silent, and therefore tacitly permissive, on that subject. The conflict arises when the State has “acted” on an issue and, in so doing, indicated that it seeks to occupy the field. In addition, Cook grants broad authority to cities to promote health pursuant to their home rule power. A court will most likely find field preemption if a state law establishes a comprehensive regulatory scheme over a specific area, as with the state minimum wage regime in Wholesale. But, under Cook, Myerson, and Hertz, a court will uphold the local law—even if it regulates in an area that is quite similar to that treated by the state statute—as long as the aspects of the regulation are different, and the state law does not establish an intent to preempt the field.

C. The Paid Sick Time Act and State Law Preemption

The Paid Sick Time Act represents an effort by the City Council to assert its home rule authority for the health and welfare of its citizens. At the threshold, it must be determined whether the Council has this authority. If it does, the second issue is whether the State Minimum Wage Law would preempt the bill if it were to become law.

A New York court would first ask whether New York’s Home Rule Law empowers the City to enact such legislation. Under Cook, a court would likely find that the City was acting pursuant to its police powers to promote health, and that the legislation has a

106 Id.
“substantial relation to a legitimate, authorized purpose.” The legislative intent section of the Paid Sick Time Act asserts that the primary thrust of the Act is to “ensure a healthier and more productive work force in New York City” and to promote public health. Thus, it seems clear that the City seeks to enact the Paid Sick Time Act “for the preservation of the public health, comfort, peace and prosperity of the city and its inhabitants.”

After establishing that the Paid Sick Time Act falls within the bounds of proper local legislation under New York’s home rule doctrine, the court would then ask whether the bill is preempted by state law. The court would first look for any evidence of a direct conflict between the Minimum Wage Act, or other state law, and the proposed legislation. An argument could be made that, under the Minimum Wage Act, the word “wages” includes other employee benefits, such as sick leave. The Minimum Wage Act does not explicitly define wages, but it does establish that they “include allowances . . . for gratuities and, when furnished by the employer to employees, for meals, lodging, apparel, and other such items, services and facilities.” The question, then, is whether a court could find that sick leave benefits could be included under “other items, services and facilities.” That is unlikely. There is no case law, nor is there employer practice, to suggest that courts in New York permit employers to view wages as anything other than compensation. The New York Minimum Wage Act does not prescribe floors for the amount of sick leave, or other benefits, as it does for wages, and employers are not permitted to compensate benefits with wages under the Minimum Wage Act.

There is one exception to this last rule. Pursuant to the New York State Labor Law, public contractors are required to pay “prevailing wages,” as set by the New York Department of Labor, and can credit benefits, or “supplements,” towards these wages.

109 N.Y.C. CHARTER § 28(a).
110 N.Y. LAB. Law § 651(7) (McKinney 2010).
111 LAB. § 220(3)(a) (requiring that workers in the construction industry receive a wage determined by the prevailing wage rate of that locality); Id. 220(3)(b) (requiring that supplements paid to the employee are determined by
“Supplements” are defined as “all remuneration for employment” which does not constitute wages.\textsuperscript{112} This remuneration includes “health, welfare, [and] non-occupational disability,” among others.\textsuperscript{113} These prevailing wage provisions offer the only instance in New York where employers may supplant wages with benefits. If the New York State legislature had wanted to provide such an option to other employers, it could have done so explicitly in the State Minimum Wage Law, as it did under the prevailing wage provisions of the New York State Labor Law. Because the Legislature has expressly permitted employers to substitute benefits for wages in some instances, it is even less likely that a court would find that the Legislature meant to imply this exchange existed in the State Minimum Wage Law. This makes an express finding of preemption even less plausible.

If a New York court were to consider how a court of another state has dealt with an express preemption challenge to local paid sick legislation under state minimum wage law, it might look to a recent Wisconsin state court decision on the Milwaukee Paid Sick Leave Ordinance.\textsuperscript{114} Although three cities have passed paid sick leave legislation,\textsuperscript{115} only Milwaukee’s bill has been challenged in court. The Milwaukee Ordinance was passed by ballot initiative and then struck down by the trial court because of the wording of the ballot question.\textsuperscript{116} The Wisconsin Supreme Court remanded after a split decision, and the case is currently pending in the Court of Appeals. The Ordinance is similar to the Paid Sick Time Act. It requires all Milwaukee employers to grant employees paid sick leave that accrues by the number of hours worked.\textsuperscript{117} Employees may take the leave for their own health, or in order to care for a

\begin{itemize}
\item \textsuperscript{112} Id. § 220(5)(b).
\item \textsuperscript{113} Id.
\item \textsuperscript{114} Paid Sick Leave Ordinance, MILWAUKEE, WIS. CODE OF ORDINANCES § 112 (2010).
\item \textsuperscript{115} See supra notes 8-10.
\item \textsuperscript{116} Metro. Milwaukee Ass’n of Commerce v. City of Milwaukee, No. 08CV018220, 2009 WL 2578536 (Wis. Cir. June 12, 2009), vacated, 787 N.W.2d 847 (Wis. 2010), divided decision, 789 N.W.2d 734 (Wis. 2010).
\item \textsuperscript{117} MILWAUKEE, WIS. CODE OF ORDINANCES § 112-3(2).
\end{itemize}
family member. The court considered state statutes governing wages and employment benefits, including the Wisconsin FMLA and the Wisconsin Living Wage Act, and found that none preempted the Paid Sick Leave Ordinance.

The Wisconsin Living Wage Act establishes minimum wage standards, defining “wage” as “any compensation for labor measured by time, piece or otherwise.” As in New York, a state agency sets the wage floors pursuant to state law, but “it does not prescribe a minimum for benefits, such as sick leave.” The court found that sick leave and wages were “fundamentally different.” Unlike for a wage,

an employer need not compensate its employee with paid sick leave if the employee does not qualify for its use. Compensation under the Living Wage Act is limited to a minimum wage amount and tips . . . . Therefore, other employment benefits, including the payment of sick leave, could not qualify as “compensation” under the state law.

Thus, the court found that the Living Wage Act did not preempt the Ordinance because it involved compensation for labor, and the Ordinance instead governed “a benefit that the employee must be eligible to receive.” Although the Wisconsin court applied its own state labor law in its decision, the guiding principal behind the analysis would probably be the same in a New York state court.

Opponents of paid sick time legislation might counter that, under Wholesale, the local law cannot bar what state law permits—in this case, the right of employers to grant sick leave as they see

---

118 Id. § 112–5(1)(b).
120 WIS. STAT. § 104.01(8) (West 2010).
121 See N.Y. LAB. LAW § 652-659 (McKinney 2010) (setting forth the authority of the Labor Commissioner to establish a wage board and to raise the minimum wage in specific industries according to the wage board’s recommendations).
122 Id. Milwaukee Ass’n of Commerce, 2009 WL 2578536 at *16.
123 Id.
124 Id.
125 Id.
fit. In his testimony against the proposed paid sick time legislation, Jack Friedman, the Executive Director of the Queens Chamber of Commerce and a representative of the Five Boro Chamber Alliance, stated that the Alliance opposed “government deprivation of our ability to determine the appropriate benefit package for our employees.” Although this appears to be framed as a policy argument, it could also be tied in to a legal argument supporting preemption. Perhaps opponents of the Paid Sick Time Act would argue that state law permits employers to regulate employee benefits within the confines of state and federal law, and therefore cities cannot interfere with the permissive stance established by the Legislature. This argument could also rely on Bloomberg, contending that the local law imposes “prerequisite “additional restrictions” on those state law rights, to the effect of “inhibiting the operation of the State’s general laws.”

However, it is not likely that this line of reasoning would succeed under Cook. The court in Cook found that the success of a preemption claim does not hinge on the state law’s silence on a subject. Instead, under Cook, the court must look to whether the state has sought to occupy the field on the subject.

This might lead to the court to ask whether the state had indicated that it wished to establish a uniform system of benefits provision, and thus occupy the field. The Wholesale court found that the State Minimum Wage Act evinced an intention to occupy the field of setting wages thresholds, and therefore the City Council’s attempt to raise wage levels for New York City was invalid. In Wholesale, this field was limited to minimum wage thresholds. For the state law to preempt the Paid Sick Time Act, however, an argument would have to be made that the Act intended to preempt the entire field of employee benefits.

The New York Legislature’s purpose in enacting the State

---

Minimum Wage Act was not nearly so broad: the Legislature intended to preserve the health of New Yorkers through the establishment and maintenance of certain wage floors. In its Statement of Public Policy, the Legislature expressed a concern for individuals employed in the State at wages insufficient to provide adequate maintenance for themselves and their families.129 This employment, the Legislature found, “impairs the health, efficiency, and well-being of the persons so employed.”130 As the Legislature concluded: “[I]t is the declared policy of the state of New York that such conditions be eliminated as rapidly as practicable . . . . To this end minimum wage standards shall be established and maintained.”131

The Wholesale court found that the state meant to occupy the field of minimum wage standards because it had established an “elaborate machinery” that would dictate the rise of the minimum wage.132 Identifying this elaborate machinery, the court pointed to provisions that created a wage board and empowered the Labor Commissioner to instigate a wage board process should she see a need for wages to increase in a particular industry. This machinery does not apply to the Paid Sick Time Act, as there is no indication that the wage board process was meant to address issues of employee benefits. Furthermore, no evidence of a “comprehensive and detailed regulatory scheme” exists in state law that is particular to the area of employee benefits.133 Thus, a state preemption challenge to the Paid Sick Time Act is more akin to Cook, Myerson, or Hertz than Wholesale or Cook. In the former, the Court of Appeals generally held that the fact that the local and state measures at issue share some common ground is not dispositive. Unless the State has expressed a desire to occupy the field, the two laws can co-exist as long as they do not regulate the same aspect of the same subject.

Even if the Legislature did not intend to preempt the field of

129 LAB. § 650.
130 Id.
131 Id.
132 Wholesale Laundry, 234 N.Y.S.2d at 865.
133 N.Y. State Club Ass’n v. City of New York, 505 N.E.2d 915, 917 (N.Y. 1987).
employee benefits, paid sick leave opponents might argue that the paid sick leave legislation would jeopardize the Legislature’s desire for “uniformity in a given field.”\textsuperscript{134} Opponents may also argue that implementing a generous sick leave policy in New York City could hurt local businesses and confer an unfair advantage upon businesses outside the City, thereby incentivizing businesses to leave the City altogether. Although these are strong policy arguments for the opponents of the bill, these arguments are not grounded in the legal question of preemption. The State Legislature has not created a uniform statute to structure and regulate all employee benefit programs, and therefore it has not shown the requisite intent to create uniformity in this area.

III. FEDERAL PREEMPTION

Federal preemption cases test the boundaries of state sovereignty and therefore frequently come before the circuits and the Supreme Court. The Supremacy Clause of Article VI of the United States Constitution governs the federal preemption doctrine.\textsuperscript{135} Under Article VI, federal law “shall be the supreme Law of the Land . . . any Thing in the Constitution or Laws of any State to the Contrary notwithstanding.”\textsuperscript{136} State law conflicting with federal statute is “without effect.”\textsuperscript{137}

State law is preempted when, \textit{inter alia}, it seeks to regulate a field solely occupied by the federal government.\textsuperscript{138} Congress may evince an intent for federal law to occupy the field by constructing a “scheme of federal regulation . . . so pervasive as to make reasonable the inference that Congress left no room for the States to supplement it,” or where Congressional law “touches a field in which federal interest is so dominant that the federal system will

\begin{flushleft}
\textsuperscript{134} \textit{In re} City of New York, 863 N.Y.S.2d 564, 567 (N.Y. Sup. Ct. 2008).
\textsuperscript{136} U.S. CONST. art. VI, cl. 2.
\end{flushleft}
be assumed to preclude enforcement of state laws on the same subject.” Federal law also preempts when the state provision is in direct conflict with federal law. This occurs when a party cannot comply with both state and federal law, or where the state law precludes the fulfillment of Congress’s goals.

A. Federal Laws and NYC Paid Sick Leave

A possibility remains that various federal laws preempt the Paid Sick Time Act, including the FMLA and ERISA. The FMLA contains an express provision allowing employees to take advantage of benefits more generous than the FMLA:

Nothing in this Act or any amendment made by this Act shall be construed to supersede any provision of any State or local law that provides greater family or medical leave rights than the rights established under this Act or any amendment made by this Act.

In dicta, the Supreme Court interpreted this clause to mean that Congress required a minimum of twelve weeks of unpaid leave and allows states the option to develop more generous leave plans. The Southern District of New York has found that this provision establishes that the Act will not curtail rights established by any state or local law. This clause is further proof that Congress did not wish for federal law—and therefore federal courts—to control the field in this area of litigation. Rather, Congress intended that the FMLA serve as a complement to state

140 English, 496 U.S. at 79.
141 Id.
142 Id. (quoting Hines v. Davidowitz, 312 U.S. 52, 67 (1941)).
143 Family and Medical Leave Act, 29 U.S.C.A. § 2651(b) (West 2010).

The Court’s language, “thus leaving States free to provide their employees with more family-leave time,” suggests that the Court may have only been referring to state employees. But even if this were true, the case still stands for the general proposition that states may supplement the FMLA without being preempted.
Along these lines, a Maryland District court found that “no part
of the [FMLA] evinces an attempt by Congress to ‘occupy the

The Wisconsin circuit court heard state law preemption
arguments involving the Milwaukee Paid Sick Leave Ordinance
and the Wisconsin FMLA (WFMLA).\footnote{Findlay v. Phe, Inc., No. 1:99CV00054, 1999 U.S. Dist. LEXIS 9761 at
*8 (M.D.N.C. Apr. 16, 1999) (quoting Family and Medical Leave Act, 29 U.S.C. § 2651).} The WFMLA, like the
Federal FMLA, allows an employee to take unpaid leave for
serious health conditions and in order to care of family
members.\footnote{Metro. Milwaukee Ass’n of Commerce v. City of Milwaukee, No. 08CV018220, 2009 WL 2578536, at *16 (Wis. Cir. June 12, 2009). Because
New York has no FMLA provisions in state law, and the WFMLA is
substantially similar to the federal FMLA for the purposes of this Note, the
WFMLA analysis is included in the federal preemption section.} Opponents of the Milwaukee Paid Sick Leave
Ordinance argued that the WFMLA preempted the Ordinance for
two reasons: (1) the leave allowed by the state law overlapped with
the uses of paid sick leave, and (2) the WFMLA affirmatively does
not require an employer to pay the employee for this leave.\footnote{See Wis. Stat. § 103.10 (West 2002).} The
court however, found that the WFMLA, like the Federal FMLA,
has an express provision providing that employees may substitute
paid or unpaid leave for the family or medical leave governed by
the Act, and that this provision “provides a carve-out for a paid
sick leave system provided voluntarily or otherwise from an
employer.”\footnote{See Id. § 103.10(5)(a), which states, “[t]his section does not entitle
an employee to receive wages or salary while taking family leave or medical
leave.”}

Even if a New York court were to hold that the FMLA’s
explicit provision allowing more generous leave does not bar a
preemption argument, it would most likely still find that the Paid
Sick Time Act does not conflict with the FMLA. The court would

first look to whether the FMLA contains an explicit preemption clause, which it does not. Then, the court would consider whether the statute contains an implied preemption, or an intention to occupy the field.\(^{151}\) The court would likely find that Congress’s purposes for creating the FMLA, as expressed in the Findings and the Purposes sections of the Act, do not encompass the kind of short-term, paid sick leave that the NYC Council was seeking to require, and that Congress had not intended for the FMLA to occupy the field of employee benefits.

Congress’s general purpose for the FMLA was to allow employees to care for newborns and for family members with serious health problems.\(^{152}\) Under the Purposes section of the Act, employees are entitled to take “reasonable leave for medical reasons, for the birth or adoption of a child, and for the care of a child, spouse, or parent who has a serious health condition.”\(^{153}\) “Reasonable leave for medical reasons” raises the possibility that the leave could be for a short duration, and therefore more akin to sick leave than child rearing or care-giving for a family member with a long-term illness. The regulations define “reasonable leave” somewhat narrowly.\(^{154}\) Employees cannot take the FMLA leave for parental or family care except in cases of birth, foster care and adoption, or a family member’s serious illness.

In order for an employee’s own sick leave to qualify under the FMLA, it must be for a “serious health condition,” which is defined as an illness or condition that involves inpatient care in a medical care facility or “continuing treatment” by a health care

\(^{152}\) Family and Medical Leave Act, 29 U.S.C.A. § 2601(b) (West 2010).
\(^{153}\) Id.
\(^{154}\) Id. Qualifying leave is defined in the regulations as follows:
(1) For birth of a son or daughter, and to care for the newborn child, (2) For placement with the employee of a son or daughter for adoption or foster care, (3) To care for the employee’s spouse, son, daughter, or parent with a serious health condition, (4) Because of a serious health condition that makes the employee unable to perform the functions of the employee’s job.
\(^{155}\) For more on what constitutes “continuing treatment,” see 29 C.F.R. §
provider. The FMLA also allows employees to take intermittent or reduced leave because of the serious health condition of the employee or of a child, spouse, or parent. These specifications are strong evidence that the short-term sick leave available under the Paid Sick Time Act is not covered by the FMLA. They also show that the FMLA did not intend, implicitly or expressly, to preempt local or state law concerning short-term sick leave, nor did it intend to “occupy the field” of sick leave generally.

Although the FMLA and the Paid Sick Time Act are similar in that they both allow employees to take leave to care for themselves or for family members, the FMLA leave is unpaid and applies only to serious health conditions. The Paid Sick Time Act, on the other hand, provides for paid leave and has a much broader definition of what kinds of illnesses qualify for such leave. Under the Paid Sick Time Act, an employer must permit an employee to use paid leave for: (1) the care, diagnosis, or preventive care for the employee’s physical or mental “illness, injury, or health condition;—or (2) to care for a family member with a mental or physical illness or condition; or—and (3) if the employee’s place of business, or his or her child’s school, closes because of a public health emergency. Thus, under the Paid Sick Time Act, an employee can claim the need to take time off for the prevention or diagnosis of, or treatment for, a wide array of conditions. After three consecutive days of leave, an employer can require reasonable documentation that the leave is for a qualifying illness, in the form of a signed


156 Family and Medical Leave Act, 29 U.S.C.A § 2611(11)(A)–(B) (West 2010). Under the FMLA regulations, “the common cold, the flu, ear aches, upset stomach, minor ulcers, headaches other than migraine, routine dental or orthodontia problems, periodontal disease, etc., are examples of conditions that do not meet the definition of a serious health condition and do not qualify for FMLA leave” unless “complications arise.” 29 C.F.R. § 825.113(d) (2011). “Mental illness or allergies may be serious health conditions, but only if all the conditions of this section are met.” Id.


159 Id. § 2(d)(1)(ii)–(iii).
Paid Sick Leave in New York

letter from a licensed health care provider.\textsuperscript{160}

It would be reasonable for a court to find that there is some overlap between the kinds of leave covered in the FMLA and the Paid Sick Time Act. For example, Mr. Santos would be able to receive paid sick leave and the FMLA leave when he finally received his pneumonia diagnosis from the doctor, because pneumonia rises to the level of a “serious illness” that qualifies for the FMLA.

It could be argued that this overlap invalidates the local law because it confounds or interferes with the purposes of the federal law. That was the argument raised by the plaintiff in the Wisconsin case, Metropolitan Milwaukee Association of Commerce. There, the Metropolitan Milwaukee Association of Commerce (MMAC) argued that the Milwaukee Paid Sick Leave Ordinance covered illnesses that qualify for leave under the WFMLA, and thus the local law conflicted with the WFMLA.\textsuperscript{161} The court agreed that there were certain similarities between the WFMLA and the Paid Sick Leave Ordinance, and that the two laws did in fact overlap.\textsuperscript{162} Specifically, the Milwaukee Paid Sick Leave Ordinance provides paid leave for various illnesses that are also covered by the WFMLA. But the court ultimately found that these overlapping provisions do not vitiate the Paid Sick Leave Ordinance,\textsuperscript{163} and instead stated that “the addition of \textit{paid} leave provides a complement [to] rather than a conflict” with the state law.\textsuperscript{164} In fact, the court suggests, an employee could elect to substitute unpaid time off for paid time.\textsuperscript{165} Therefore, the court found that the Wisconsin legislature “has not expressly withdrawn the power of the City to act, does not logically conflict with the [WFMLA], does not defeat the purpose of the [WFMLA], and does not violate the spirit of the [WFMLA].”\textsuperscript{166}

\textsuperscript{160} Id. § 2(d)(3).
\textsuperscript{161} Metropolitan Milwaukee Ass’n of Commerce v. City of Milwaukee, No. 08CV018220, 2009 WL 2578536, *17 (Wis. Cir. June 12, 2009).
\textsuperscript{162} Id.
\textsuperscript{163} Id. at *18.
\textsuperscript{164} Id.
\textsuperscript{165} Id.
\textsuperscript{166} Id.
Although the Wisconsin court applied Wisconsin state preemption doctrine in its analysis of the WFMLA preemption, a New York court applying federal preemption doctrine would most likely reach a similar finding. Under the case law on federal preemption discussed supra, this kind of overlap between federal and local law does not suffice to constitute federal preemption. In order for the FMLA to preempt the Paid Sick Time Act, it would have to be “impossible for a private party to comply with both state and federal requirements,”167 or the local law must present an “obstacle to the accomplishment and execution of the full purposes and objectives of Congress.”168 Requiring employers to provide paid leave that buttresses the FMLA simply would not present an obstacle so great such as to frustrate Congress’s objectives, nor would it make it “impossible” for a private employer to manage both kinds of leave.

B. ERISA Preemption

ERISA regulates employee benefit plans maintained by private employers. The statute deals exclusively with a “plan, fund, or program . . . established or maintained by an employer or by an employee organization, or by both,” to provide any of the types of welfare or pension benefits described in the statute for employee participants or their beneficiaries.169 A court’s main concern regarding ERISA preemption would likely be whether the Paid Sick Time Act constitutes is a “welfare plan.”170

In its ERISA regulations, the U.S. Department of Labor (USDOL) sheds light on the parameters of a welfare plan by listing

168 Id. (quoting Hines v. Davidowitz, 312 U.S. 52, 67 (1941)).
170 29 U.S.C.A. §1002(1). A “welfare plan” is defined as plan, fund, or program “established or maintained by an employer or by an employee organization . . . for the purpose of providing for its participants or their beneficiaries through the purchase of insurance or otherwise, (A) medical, surgical, or hospital care or benefits, or benefits in the event of sickness, accident, disability, death, or unemployment.” Id.
programs that do not fall into the specifications. Included in those plans is the

[p]ayment of an employee’s normal compensation, out of the employer’s general assets, on account of periods of time during which the employee is physically or mentally unable to perform his or her duties, or is otherwise absent for medical reasons (such as pregnancy, a physical examination or psychiatric treatment).\textsuperscript{171}

Courts might consider whether an “absence for medical reasons” would constitute the kind of paid leave permitted under the Paid Sick Time Act.

The USDOL addressed just this question in a 1994 Advisory Letter.\textsuperscript{172} An employer, Parisian Inc., had a paid sick leave plan for employees who worked thirty-five or more hours per week and had worked for six months for the employer. The benefits, which were paid by the employer’s general assets, would accrue based on the employee’s years of service. USDOL ruled that, although the employer’s sick leave policy provides “benefits in the event of sickness,” as identified in section 3(1)(A) of ERISA, it is a “payroll practice,” as described in § 2510.3-1(b), and therefore does not constitute an employee welfare benefit plan within the meaning of section 3(1). It is the Department’s position that an employer’s practice of paying no more than an employee’s normal compensation during periods of absence due to illness, out of the general assets of the employer, falls within the exception to coverage carved out in the regulation.\textsuperscript{173}

A Wisconsin district court has held that employees can substitute benefits provided by the employer for the FMLA benefits. In \textit{Aurora Medical Group v. Department of Workforce Development}, the court held that an employee was permitted to substitute paid sick time provided by her employer for unpaid sick

\textsuperscript{171} 29 C.F.R. § 2510.3-1(b)(2) (2011).
\textsuperscript{173} \textit{Id.}
time allowed under the FMLA. The court found that the FMLA preempted ERISA. “[T]o the extent to which ERISA is amended by the FMLA, ERISA must yield to any provisions of the WFMLA providing greater family leave rights than those provided by the FMLA.”

To support this finding, the court relied on the U.S. Senate’s 1993 floor debate over the FMLA. During the debate, Senator Dodd, one of the legislation’s main sponsors, was asked whether ERISA would prohibit employees from using accrued paid leave, regardless of its source, in lieu of any leave provided under ERISA. Senator Dodd’s response was that the FMLA was “intended to supersede ERISA and any Federal Law.” Congress’s intent, he maintained, was to ensure that ERISA and other federal law did not “undercut[]” family and medical leave laws at the state level. States that allowed employees to replace unpaid family leave with accrued paid leave would continue to be permitted to do so, “so long as those State law provisions are at least as generous as those of this Federal legislation.”

A New York court ruling on ERISA preemption and the Paid Sick Leave Act would also most likely consider Bloomberg, which found that the City’s equal employment benefits law was preempted by ERISA. The court held that the Equal Benefits Law at issue “mandate[s] employee benefit structures or their administration,” and therefore conflicted with ERISA. Despite the fact that this requirement is conditionally based on whether the vendor chooses to contract with the City, the court nonetheless held that the Law “connect[s] with a core concern of ERISA, impermissibly interferes with its goal of uniform plan

---

174 Aurora Med. Grp. v. Dep’t of Workforce Dev., 602 N.W.2d 111 (Wis. Ct. App. 1999), aff’d, 612 N.W.2d 646 (Wis. 2000).
175 Id. at 114.
176 Id. at 114–15.
177 Id. at 115.
178 Id.
180 Id. at 110 (quoting N.Y.S. Conf. of Blue Cross & Blue Shield Plans v. Travelers Ins. Co., 514 U.S. 645 (1995)).
administration, and is thus preempted.” 181 The local law invalidated by the Bloomberg court is distinguishable from the Paid Sick Time Act, however. With respect to the Equal Benefits Law, New York City governed all “employee benefits,” which ranged from health insurance and pension and retirement benefits, to sick leave, life insurance, bereavement policies, and tuition reimbursement. 182 These benefits align much more closely with an ERISA welfare plan, and do not fall into any exceptions laid out in the statute.

CONCLUSION

Low-earners, part-time workers, and women are hit hardest by their own or their family’s sickness. Paradoxically, they are also the groups that are least likely to have paid sick days. Advocates for paid sick leave legislation have made compelling arguments about the need for all workers to have a minimal number of days they can take off when they or a family member is ill. In addition, advocates point out the public health hazards triggered when restaurant workers prepare food while sick, or when children go to school sick because their parents cannot afford to take time off to care for them. For these and other reasons, some states and localities have created new floors on minimum worker benefits. These new laws rely on the authority of local and state governments to act for the welfare of the state or municipality. But this authority is called into question, particularly by businesses and chambers of commerce. Defending the authority of legislative bodies to pass health and welfare legislation on the state and local level involves a multi-pronged approach. States, municipalities, and advocates must work to defend local laws against claims of preemption by federal law and interference with state and federal regulatory schemes, and local laws must be justified under home rule provisions.

The Paid Sick Time Act is a fitting exercise of local home rule authority and does not conflict with state and federal laws. The

181 Bloomberg, 791 N.Y.S.2d at 110.
182 N.Y.C. ADMIN. CODE § 6-126(b)(7) (West 2009).
State Minimum Wage Act establishes a comprehensive scheme for wage regulation, but does not regulate employee benefits. The FMLA has an explicit preemption provision, allowing states and localities to craft family leave provisions that go beyond the scope of the statute. Finally, ERISA regulates employment welfare plans, but a USDOL Advisory Letter suggests that paid sick leave falls into one of the statute’s enumerated exceptions.