On Nontraditional Trademarks
Kenneth L. Port ................................................................. 1

Almost a Century and Three Restatements After Green
It’s Time to Admit and Remedy the Nonsense of Negligence
Rory Bahadur ..................................................................... 61

Presidential Ambitions of U.S. Supreme Court Justices:
A History and an Ethical Warning
William G. Ross .................................................................. 115

Limitations Imposed on the Dual Sovereignty Doctrine by
Federal and State Governments
Thomas White .................................................................... 173

Spot Delivery: Why it is Illegal and Why Car Dealers
Have Been Getting Away With Theft For So Long
Gregory R. Ingalsbe and Raymond G. Ingalsbe ..................... 221

NOTE

Front Pay Under the FMLA
Michelle Eviston ............................................................... 259
ON NONTRADITIONAL TRADEMARKS

Kenneth L. Port*

I. INTRODUCTION

International substantive trademark law harmonization has been an abject failure. This failure is manifested in the inconsistent treatment of nontraditional trademarks. This article exposes the perceived broad protection of nontraditional trademarks for the great deception that it is. By allegedly protecting nontraditional trademarks, the United State renders true trademark harmonization impossible. The United States itself and commentators go too far in expressing this protection. In fact, nontraditional trademark protection is extremely limited, very rare and almost never enforced. Yet, the United States requires trading partners to protect nontraditional trademarks as if it is a mainstay of trademark jurisprudence in the United States, when it is not. The World Intellectual Property Organization (WIPO) and the international community are in the process of harmonizing to this fictitious standard. This will render real efforts at genuine trademark harmonization even more difficult to attain.

Part II of this article describes the setting of nontraditional trademark jurisprudence in the United States. It concludes that Americans boast of protection of odd apppellations of source, but in reality, it is very infrequently accomplished and then only at great expense.

Part III traces the etymology of the words “symbol” and “device”, the precise statutory terms upon which the protection of nontraditional trademarks rests.

Parts IV – XI address each known type of nontraditional trademark. Specific data is provided to conclude that color and sound make up virtually all nontraditional trademarks and that the data does not support claims of nontraditional trademarks playing a meaningful role in United States trademark jurisprudence.

Part XII presents a very brief explication of the Common Law of trademarks compared to the Civil Law of trademarks so that one can see the disparate views of trademark protection.

Part XIII provides a normative attack on the protection of nontraditional trademarks and concludes that, in almost all cases, they do not operate as indicators of source and therefore should not be protected.

* Professor of Law and Director, Intellectual Property Institute, William Mitchell College of Law. I am indebted to Madeline Bowie (2010) for her research assistance. I am also grateful to Jim Baker, Adjunct Professor of Law, William Mitchell College of Law and the members of my Contemporary Issues in IP class for their input and feedback of an earlier draft of this article. I am also grateful to Neal Axton, Reference Librarian, William Mitchell College of Law, for his research and input.
Part XIV concludes that nontraditional trademarks do not play the major role in American trademark jurisprudence that some want us to believe. There has been registration activity but no corresponding reported enforcement activity. As such, much is made of nontraditional trademark protection but it is a vacuous promise. The world should not harmonize to the vacuous. If we harmonize, it should be to a standard that will be effective in lowering trade barriers, increasing the free flow of goods and services, and generally making the lives of trademark holders easier, not to a fictitious standard that is almost entirely meaningless.

II. CONTEXT

“Nontraditional trademarks,” include sensory marks such as sound, color, scent, taste, and tactile marks, and also holographic marks or even the motion of a product. They play an increasingly important role in the plan of some to engage in “extreme branding.”¹ Nontraditional indicators of source purportedly become a valid trademark when they are or become distinctive of the good or service on which they are used and, in the minds of the consumers, come to identify a consistent source or origin for that good or service² and can be represented graphically. Nontraditional trademarks fall into the larger category of trade dress discussed in the Qualitex Co. v. Jacobson Products Co., 514 U.S. 159 (1995) decision and, as such, must have secondary meaning to be protected.³

The Qualitex case itself is said to be a “stunning breakthrough for nontraditional trademarks.”⁴

In fact, United States Patent and Trademark Office (PTO) application statistics indicate that there were only 93 nontraditional trademarks on record.

¹. See, e.g., EXTREME BRANDING, http://www.extremebranding.com (last visited Sept. 22, 2011). This phrase itself is a purported trademark, although unregistered. This term was recently coined by Nick Pisarsky, Note, PoTAAds-PoTAHo-Let's Call the Whole Thing Off: Trademark Protection of Product Sounds, 40 CONN. L. REV. 797, 800 (2008). A trademark application for EXTREME PERSONAL BRANDING was received by the United States Patent and Trademark Office (PTO) but was soon abandoned. See United States Patent and Trademark Office, Leg Up Marketing, LLC, Serial No. 78847574 (indicating the mark has been abandoned) available at http://tarr.uspto.gov/servlet/tarr?regser=serial&entry=78847574.


³. Fabrication Enters., Inc. v. Hygenic Corp., 64 F.3d 53 (2d Cir. 1995) (color scheme on rubber bands to denote resistance has acquired secondary meaning and is protectable); Mana Prods., Inc. v. Columbia Cosmetics Mfg., Inc., 65 F.3d 1063 (2d Cir. 1995) (no secondary meaning for black compacts in ordinary shapes used within the industry as a whole); Knitwaves Inc. v. Lollytogs Ltd., 71 F.3d 996 (2d Cir. 1995) (ecological design on sweaters is distinctive). As of yet, an unanswered question is whether product packaging that amounts to a nontraditional trademark could be inherently distinctive and protectable without possessing secondary meaning. In Walmart Stores, Inc. v. Samara Bros., Inc., 529 U.S. 205 (2000), the Supreme Court held that product packaging could be inherently distinctive but not product design. That case is silent on the nontraditional trademark discussed herein.

prior to the *Qualitex* decision.\(^5\) After the *Qualitex* decision published on March 28, 1995, there have been 688 trademark applications filed for nontraditional trademarks.\(^6\) That is, in the first 48 years of the existence of the Lanham Act, there were 93 nontraditional trademarks filed (1.93 per year).\(^7\) In the 14 years since the *Qualitex* decision, there have been 688 trademark applications for nontraditional trademarks (approximately 49 a year).\(^8\) This is an explosion of interest in nontraditional trademarks.

Contrary to the claims of some that *Qualitex* resulted in a “stunning breakthrough,”\(^9\) in fact, *Qualitex* did not have the same impact on the number of reported cases involving nontraditional trademarks.\(^10\) In the 48 years before *Qualitex*, there were 20 reported cases; in the 14 years since *Qualitex* (for which there is data), there have been 15 reported cases.\(^11\) Per year that amounts to a doubling of the number of cases (pre-*Qualitex*: .42 per year; post-*Qualitex*: 1.07 per year); however, as these numbers remain around one per year, it is hard to see this as a “stunning breakthrough” and certainly not the exponential increase that there was in applications.\(^12\)

Even so, nontraditional trademarks still make up an infinitesimally small portion of all trademark applications or registrations.\(^13\) In August of 2009,

---

5. See infra app. E.
6. See infra app. E.
7. See infra app. E.
8. See infra app. E.
11. Id.
13. See infra note 14 (comparing number of nontraditional trademarks to number of total trademarks).
approximately 0.01504% of all trademark applications on record with the PTO were for nontraditional trademarks.\textsuperscript{14}

Yet, much is made of this ability in the United States to register nontraditional trademarks\textsuperscript{15} and the United States has even placed pressure on other countries to expand the scope of other countries’ trademark laws to include the protection of nontraditional trademarks.\textsuperscript{16} The Singapore Treaty on Trademarks in 2006 was the first treaty to explicitly recognize nontraditional trademarks although it also stated that signatory states are under no obligation to adopt “new types” of trademarks.\textsuperscript{17} This interest in and support for

\begin{itemize}
\item \textsuperscript{14} A search on the Trademark Electronic Search System (TESS) maintained by the USPTO, for “(live)[LD] or (dead)[LD]” on August 28, 2009 returned 5,192,791 records. This is the total number of trademark applications on file with the PTO as of that date. The number of nontraditional trademarks gathered from TESS yielded 781 applications as of August 28, 2009. Therefore, the percentage of nontraditional trademarks on file with the PTO in any capacity, be they applications, registrations, abandoned or canceled, is only 0.01504%. For a similar search done in 2006, see Nick Pisarsky, \textit{PotAYto-PoTAHto-Let’s Call the Whole Thing Off: Trademark Protection of Product Sounds}, \textit{40 CONN. L. REV.} 797, 803 n.28 (2008) (showing that the percent of nontraditional trademarks was .00727%). Note that there is no search term that will search all nontraditional mark applications in TESS. As a result of the limitations of TESS nontraditional trademark applications must be gathered by using various search terms in addition to searching “(‘6’)[MD]”, which returns marks that cannot be visually perceived. Therefore, the 780 nontraditional trademark applications cannot be considered an absolute number as there is an unquantifiable margin of error when searching nontraditional marks on TESS.
\item \textsuperscript{16} This pressure comes in the form of free trade agreements. If a country wants a free trade agreement, the United States’ agreement contains a provision to require partner nations to protect sound and scent marks. To date, no other nontraditional marks are protected in the free trade agreements. \textit{See}, e.g., United States–Australia Free Trade Agreement, U.S.-Austl., art. 17.2(2), May 18, 2004, \textit{U.S.T. LEXIS} 162, \textit{234}, \textit{available at} \url{http://www.ustr.gov/sites/default/files/uploads/agreements/fta/Australia/asset_upload_file5141.pdf} which reads in its entirety as follows: “Neither Party may require, as a condition of registration, that marks be visually perceptible, nor may a Party deny registration of a mark solely on the ground that the sign of which it is composed is a sound or a scent.”. The free trade agreements with 13 other nations contain similar requirements. \textit{See generally}, Free Trade Agreements, Office of the United States Trade Representative \url{http://www.ustr.gov/trade-agreements/free-trade-agreements} (providing texts of other free trade agreements). The International Trademark Association (INTA) has its own model free trade agreement that requires countries to protect sound, single colors, and three-dimensional shape marks. \textit{See Model Free Trade Agreement Proposals Intellectual Property Rights: Trademark Law,} \url{http://www.inta.org/downloads/tap_modelfreetrade.pdf} (expressing intent to protect sound, single color, and three-dimensional shape marks).
\item \textsuperscript{17} Singapore Treaty on the Law of Trademarks, R. 3(5) available at \url{http://www.wipo.int/edocs/mdocs/mdocs/en/ltt_r_de/ltt_r_de_30.pdf} (“[Hologram Mark, Motion Mark, Color Mark, Position Mark] Where the application contains a statement to the effect that the mark is a hologram mark, a motion mark, a color mark or a position mark, a Contracting Party may require one or more reproductions of the mark and details concerning the mark, as prescribed by the law of that Contracting Party”). The United States position on nontraditional trademarks seems to be that the PTO will support international efforts in this regard because there is so little else where cooperation seems possible at this given moment.
\end{itemize}
nontraditional trademarks is paying off as Japan has moved to adopt nontraditional trademarks and even WIPO’s Standing Committee on the Law of Trademarks, Industrial Designs and Geographic Indications has recently taken this up in the form of a major study and report.\textsuperscript{18} At least one entire book was published by the International Trademark Association that gives the implication that protecting nontraditional trademarks is a common, routine and a frequent affair.\textsuperscript{19}

In fact, the registration of nontraditional marks in the United States is a very difficult, time consuming and expensive endeavor. Nontraditional trademark registrations constitute a minute fraction of all registered trademarks in the United States.\textsuperscript{20} Furthermore, nontraditional trademarks are extremely difficult to enforce. Of the 2,762 reported Lanham Act cases since 1947, the inception date of the Act, 35 (1.27\%) were cases regarding a nontraditional trademark and of those 13 (37\%)\textsuperscript{21} prevailed when the national average over the 60 years of the Lanham Act is that the plaintiff prevailed 50\% of the time.\textsuperscript{22}

Some of the more well-known registered nontraditional marks include the roar of the lion in the initial credit scenes of a movie (owned by MGM),\textsuperscript{23}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{nontradt_marks_litigation.pdf}
\caption{Percentage of Nontraditional Trademark Litigation where Plaintiff Prevailed 1947-2007}
\end{figure}


\textsuperscript{20} An August 28, 2009 search on TESS for ‘RN > “0” returned 2,968,267 records. This is the total number of registered trademarks on file with the PTO as of that date. The number of nontraditional trademark registrations gathered from TESS was 452 as of August 28, 2009. Therefore, the percentage of nontraditional trademark registrations on file with the PTO is a mere 0.015227\%.


\textsuperscript{22} See Mitchell Study on Trademark Litigation, Graph B: Total Infringement Claims Established/Not Established (2008), http://www.wmitchell.edu/intellectual-property/files/WMTMStudy-GraphB.pdf (comparing number of established claims to not established claims).

\textsuperscript{23} Metro-Goldwyn-Mayer Lion Corp., Reg. No. 1,395,550 (“The mark comprises a lion roaring.”).
plumeria-scented yarn that came in a sealed kit to knit a skunk (well-known, that is, among intellectual property people),\textsuperscript{24} or even the sound of a Harley Davidson motorcycle while on idle.\textsuperscript{25} Some of the less known registered nontraditional trademarks include the vertical opening motion of Lamborghini car door\textsuperscript{26} or the vertical spray of water from the back of a Yamaha personal watercraft.\textsuperscript{27} Applications for registration include some unusual things including the DNA sequence for futurist Frank Ogden\textsuperscript{28} or anthropomorphic wilted flowers yelling insults.\textsuperscript{29}

Nontraditional marks may be registered in the United States because the Lanham Act states that “any word, name symbol or device” that is used on or in connection with some good or service may be registered if it has secondary meaning.\textsuperscript{30} The United States Supreme Court has interpreted this use of “any” as actually meaning “any.”\textsuperscript{31} Therefore, the mere color alone of a green/gold press

\begin{itemize}
\item 24. *In re Clarke*, 17 U.S.P.Q.2d (BNA) 1238 (T.T.A.B. 1990); See also Celia Clarke, Reg. No. 1,639,128 (“The mark is a high impact, fresh, floral fragrance reminiscent of plumeria blossoms.”).
\item 25. Harley-Davidson, Inc., Serial No. 74,485,223 (“The mark consists of the exhaust sound of applicant’s motorcycles, produced by V-Twin, common crankpin motorcycle engines when the goods are in use.”).
\item 26. Automobile Lamborghini Holding S.p.A., Reg. No. 2,793,439 (“The mark consists of the unique motion in which the door of a vehicle is opened. The doors move parallel to the body of the vehicle but are gradually raised above the vehicle to a parallel position.”).
\item 27. Yamaha Hatsudoki Kabushiki Kaisha, U.S. Patent No. 1,946,170 (“The mark is comprised of a three dimensional spray of water issuing from the rear of a jet propelled watercraft and is generated during the operation of the watercraft.”). See generally, Lesley Matty, Note, *Rock, Paper, Scissors, Trademark? A Comparative Analysis of Motion as a Feature of Trademarks in the United States and Europe*, 14 CARDOZO J. INT’L & COMP. L. 557 (2006) (discussing motion marks). A personal watercraft is more commonly called a “jet ski.” However, JET SKI is actually the registered trademark of Kawasaki Heavy Industries, Ltd., U.S. Patent No. 1,394,547 (filed Jan. 20, 1984). Of course, this is not to be confused with Honda’s AQUATRAX (U.S. Patent No. 2,737,518 (filed August 9, 2000)) or Bombardier Recreational Products’ SEA-DOO (U.S. Patent No. 3,705,331 (filed Feb. 3, 2005)).
\item 28. Frank Ogden’s trademark application claimed its identification of goods as follows: “DNA GENE SEQUENCE TO PRODUCT [sic] FRANK OGDEN OR OTHER PROSPECTIVE HUMANS, CLONES, SURVIVAL MACHINES, OR FUTURISTIC ANIMALS THAT ARE USED TO CREATE IDEAS.” U.S. Patent Serial No. 76,016,924 (filed April 3, 2000) (abandoned Jan. 21, 2002). Frank Ogden maintains a web site at http://www.drtomorrow.com. Mr. Ogden’s registered trademark for Dr. TOMORROW was cancelled in 2004. U.S. Patent Serial No. 74,274,318 (filed May 11, 1992) (cancelled June 12, 2004). When asked, Frank Ogden responded that the reason he attempted to register his own DNA sequence was because “I THOUGHT THAT JUST THE LOOKING AT A PERSON AND BE WILLING TO PENETRATE THEIR DNA WOULD REMOVE A LOT OF CONTROL OVER ANY TAXPAYER. I WANTED TO GET THERE BEFORE SOME BUREAUCRAT GOT A BILL PASSED FOR THAT CONTROL...” E-mail to Frank Ogden (July 29, 2009) (on file with author).
\item 31. *Qualitex*, 514 U.S. at 162.
\end{itemize}
pad used in the laundry business was found to be a valid trademark. That is, any “symbol or device” that identifies source is a valid trademark in the United States if it has secondary meaning.

This article argues that nontraditional trademarks should be denied registration and be deemed unenforceable. Nontraditional trademarks are functional. Nontraditional trademarks are subjective; what one person perceives of the sensory mark is not what others perceive. As such, there can be no reliable test for when the nontraditional mark is infringed or diluted. Courts know this. Nontraditional marks succeed in trademark infringement or dilution causes of action only 37% of the time while trademark holders prevail nearly 50% of the time in all reported trademark causes of action since 1947.

More importantly, the substantive harmonization of trademark laws around the world, a contentious issue for 120 years, is proceeding. However, the substantive harmonization is to an odd standard. There now is a movement among WIPO countries to harmonize to the standard of protecting nontraditional trademarks as opposed to preventing nontraditional trademarks from becoming protectable. This is an odd movement because there have been so few significant nontraditional trademark registrations in the United States and so few nontraditional trademarks that are successfully enforced. Even if harmonized, these numbers are so small that they could not possibly make any difference whatsoever. Furthermore, this doctrinal harmonization will not be effective because the underlying jurisprudence of trademark protection between the United States and Civil Law countries will remain distinct.

Therefore, the goal of harmonization, which is to some degree to create approximations between two or more systems to reduce transaction costs, will not be recognized because the theory on which protection is based will remain distinct and the numbers of possible nontraditional trademarks remain extremely

32. Id. at 166.
33. Id. at 162-63.
34. For purposes of this article, as product design or product configuration marks have become mainstream, trade dress marks will not be considered nontraditional.
small. As such, this harmonization is doomed to fail as it is numerically infinitesimal and a misguided and mistrusted jurisprudence.

The United States is the only country in the world that protects “any” word, name or even device as trademarks. Although Japan is considering legislation that would extend trademark protection to nontraditional marks,38 no other country has gone so far as to pass express legislation in this manner.39 The fact that the United States’ system of trademark jurisprudence conceptually supports and validates nontraditional trademarks is a manifestation of how far the United States is, conceptually, from the trademark laws of the Civil Law countries. That is, by protecting “any” symbol or device, the United States trademark jurisprudence is out of sync with the rest of the world. Although some countries are attempting to doctrinally harmonize their trademark law to the peculiar standard of protecting nontraditional trademarks, they will not succeed in harmonizing trademark law in this area. There are so few nontraditional trademarks that providing them protection will not affect anything in the larger universe of trademark jurisprudence.

This is significant because most of the people on the planet do not live in Common Law jurisdictions. In fact, as the following graphs40 make explicit, the vast majority of the world population lives in either strict Civil Law jurisdictions or in some combination. Only a third of the world’s population lives in strictly Common Law countries, and even then registration plays a significant role in trademark right generation. Less than a third of all countries are Common Law nations.

In the United States, the trademark right is a right of exclusion.41 One is allowed to exclude others only to the extent one uses a mark and for long as one uses the mark.42 In the Civil Law world, a trademark is much more. In countries like France,43 Germany44 or Japan,45 the registered trademark is property.46 In these countries, a trademark is, by definition, a registered mark.47 Registration in Civil Law countries is harder to obtain than in the United States and registered marks are more difficult to lose. As the Civil Law of trademarks treats the trademark as property and in the United States it is a mere right of exclusion, in


45. KENNETH L. PORT, TRADEMARK AND UNFAIR COMPETITION LAW AND POLICY IN JAPAN 5-6 (2007).

46. Section 4 of Council Regulation that established the CTM system is titled “Community Trademarks as Objects of Property” and Article 19 of the Regulation gives in rem jurisdiction over trademarks. Council Regulation 40/94, 1994 O.J. (L 11) 1.

47. In Japan, for example, Article 18 of the Trademark Act says that the trademark right subsists only upon registration. Trademark Act, Law No. 127 of 1959, art. 18 (Japan), translated and available at http://www.japoneselawtranslation.go.jp/law/detail?id=45&vm=04&re=02. See also Max Vern, A Trademark Visa -- Aspects of International Trademark Use and Protection, 88 J. PAT. & TRADEMARK OFF. SOC’Y 847, 848 (2006) (“[t]he tenet of the Civil Law system [is] that there are no trademark rights without a trademark registration”). See generally KENNETH L. PORT, TRADEMARK AND UNFAIR COMPETITION LAW AND POLICY IN JAPAN 77-81 (2007)(discussing McDonald’s late trademark registration and protection in Japan).
the 125 years since the Paris Convention, the world has made nearly no progress in harmonizing substantive trademark law. This lack of progress has not been because countries were not trying.

In fact, much effort has been given to the job of harmonizing trademark laws. Success has been elusive, however, because the Civil Law of trademarks is so conceptually disparate from the Common Law of trademarks. To be meaningful, in harmonizing trademark law, both the theory and the doctrine of trademark law must be harmonized. To merely adopt, for example, a trademark system where rights are created by registration, as in the Civil Law world, not by use, would be unconstitutional in the United States. That is, it would take a constitutional amendment for the United States to grant rights based on registration, not on use.

In fact, there are major and much more meaningful issues in substantive trademark harmonization other than, for example, whether Japan will allow the registration of the motion of a Lamborghini door. The world cannot agree on when trademark rights are created. Is it upon use or upon registration? The world cannot agree on what constitutes a famous mark and to what degree such a famous mark

---

The world is morbidly fixed to the notion of territoriality of trademark laws. All these issues are far more meaningful than whether a country protects the motion of the spray of water behind a personal watercraft.

As subjective aesthetic indicators, nontraditional marks do not fulfill any justifications for trademark protection. When a mark satisfies the justifications for trademark protection, it becomes appropriate to grant the trademark holder a right to exclude others from using the mark. It makes competitive sense to do this because we want and need companies to invest in trademarks to lower the transaction costs of purchasing goods and services. For example, the stronger the mark, the lower the search costs consumers incur when attempting to buy a product they like or want.

However, nontraditional marks do not lower transaction costs. In fact, as they are subjective and indefinite, they actually increase transaction costs for the consumer. As such, they are inappropriate trademark subject matter as they do not further the purpose of trademark protection. They also waste scarce resources of the PTO. As nontraditional trademarks are a very rare breed, harmonizing to this standard is a curious development for a world where trademark laws and policies are so distinct. As the harmonization efforts that exist regarding nontraditional marks are only on a doctrinal level and not a theoretical level, the idea of trademark protection will remain disparate and distinct, even if some countries harmonize doctrine to allow for the protection of nontraditional trademarks.

III. HISTORY OF THE WORDS “SYMBOL” AND “DEVICE”

Today, nontraditional trademarks are deemed protectable as they constitute either a “symbol” or a “device” as used in the Lanham Act. To determine the


53. Personal Watercrafts (PWC) are very popular in the United States with approximately 20 million Americans riding them each year and annoyingly zipping around lakes and rivers, which are otherwise quiet and peaceful bodies of water. In 2005 there were 1.55 million PWCs in use throughout the country. PWC sales generate a considerable amount of revenue each year in the United States. In 2005, the industry grossed $7.6 million in sales and sold more than 80,000 units. Personal Watercraft Industry Association, Personal Watercraft Sales, http://www.pwia.org/faq/sales.aspx (last visited October 1, 2009).


56. Landes & Posner, supra note 54 at 270-71; but see Beebe, supra note 55 (criticizing the economic analysis of trademark law).

appropriateness of this determination, it is instructive to trace the history of the use of the words “symbol” and “device” as they have been used in the definition of trademarks over time. When viewed from a historical context, these words, as originally construed, did not support the notion of nontraditional trademarks.

Professor Lars Smith is correct when he claims that the use of the word “device” in the definition of trademarks was in the original Lanham Act, however, it did not start there. In fact, it appears that use of the words “symbol” and “device” as part of a definition of a trademark originally appeared in an English case from 1836. It is made confusing because “device” was not used in the definition of a trademark in any of the trademark legislation that preceded the Lanham Act and “symbol” was only used in the Act of 1920. Both the terms “symbol” and “device” are derived from a concept much older than the modern trademark laws.

The earliest known use of the words “symbol” and “device” to define a trademark occurred in the English case of Knott v. Morgan in 1836. In this case, the London Conveyance Company operated “omnibuses,” horse drawn bus-like vehicles, pictured below, in London. Morgan, the defendant, used the terms CONVEYANCE COMPANY and LONDON CONVEYANCE COMPANY on competing omnibuses. Morgan also used a green and gold star and garter type symbol, pictured below. The plaintiff had been using a similar symbol with similar coloring for some time prior to the defendant.
The English court held that the plaintiff did not have exclusive rights in or to the words “London Conveyance Company” or “Conveyance Company,” but it did deserve an injunction because of the defendant’s use of these words in conjunction with the same or similar symbols.  

The defendant was enjoined using the same symbols and:

[From running, or in any manner using or causing to be used, for the conveyance of passengers [sic], his omnibus in the bill mentioned, or any other omnibus, having painted, stamped, printed, or written thereon the words or names “London Conveyance,” or “Original Conveyance for Company,” or any other names, words, or devices painted, stamped, printed, or written thereon, in such manner as to form or be a colourable imitation of the names, words and devices painted, stamped, printed, or written on the omnibuses of the Plaintiffs . . . .]

In this case, the word “device” and/or “symbol” does not relate to any nontraditional trademark. In fact, the word “device” is used synonymously with the words “design mark.” That is, it merely identifies the colors and shapes of the star and garter type symbol as depicted below.

The first known case where the word “device” appears in the United States in a trademark case was in 1840. In *Bell v. Locke*, the court held that the DEMOCRATIC REPUBLICAN NEW ERA newspaper trademark was not infringed by NEW ERA newspaper. The court found that a distinguishing feature was the fact that the defendant used a large portrait of “a Bird of Jove” and the plaintiff used a rather modest picture of an eagle in conjunction with

---

69. *Id.* at 613.
70. *Id.* at 613-14.
71. *See Knott v. Morgan, 48 Eng. Rep. 610 (1836); Upton, supra note 61 (using the words “device” and “symbol”).
72. *Id.*
73. *See Bell v. Locke, 8 Paige Ch. 75 (N.Y. Ch. 1840) (using the word “device”).
74. *Bell v. Locke, 8 Paige Ch. 75 (N.Y. Ch. 1840).*
75. An eagle.
their word marks.76 The court stated as follows: “But surely no one can mistake
the wide-spread wings and the warlike attitude of ‘the bird of Jove,’ which
occupies so large a space in the heading of the defendant’s paper, for the very
modest device which sustains the Democratic Republicanism of the
complainant’s New Era.”77

Clearly, in this case the word “device” was meant to indicate the plaintiff’s
design mark78 that consisted of an eagle that provided a distinguishing feature to
the defendant’s use of the word marks. The “device” spoken of, seems to be the
relative eagles alone or the relative eagles in conjunction with the relative word
marks.79 There was no mention of the word “symbol.”80 Regardless, the word
“device” is synonymous here with “design mark,” the words we would use today
to describe such a trademark.

In 1844, the Circuit Court of Massachusetts characterized the packaging
used to identify the plaintiff’s thread as “devices.”81 The court agreed with the
initial bill and found the following:

[T]he defendant has extensively manufactured and sold spurious threads
of inferior quality, put up on spools similar to those used by
complainants, and colored, stamped and labeled, and enclosed in
envelopes resembling exactly the spools, labels, devices, trade marks
and envelopes used by the complainants; that this conduct of the
defendant has greatly injured the complainants, not only by depriving
them of the profits on the sale of large quantities of their own threads,
but by the prejudice to the reputation of the article manufactured by
them, caused by the inferior quality of the threads sold by the
defendant.82

In this case, the use of the word “device” seems to mean the entire
conglomeration of things that the plaintiff did to differentiate its goods from
others. To be sure, the colors red and black were part of this device, but, just as
assuredly, not as used separately from this overall source identifying device.

The next occurrence of the word “device” used in relationship to a trademark
in the United States was in 1846.83 This time, the Chancery Court of New York
refused to enjoin the defendant’s use of a bee-hive and foliage mark on a match
box because the plaintiff’s mark was made slovenly and the defendant’s “device”

76. Bell, 8 Paige Ch. at 77.
77. Id.
78. A design mark consists of graphics in place of or in addition to letters to identify source. See Boston Duck Tours, LP v. Super Duck Tours, LLC, 531 F.3d 1 (1st Cir. 2008) (involving duck graphics with the names of tour companies).
79. See Bell v. Locke, 8 Paige Ch. 75 (N.Y. Ch. 1840) (comparing the newspapers).
80. Id.
81. Taylor v. Carpenter, 3 Story 458, 23 F.Cas. 742 (1844).
82. Taylor v. Carpenter, 3 Story 458, 23 F.Cas. 742 (1844).
83. See Patridge v. Menck, 2 Sand. Ch. 622 (N.Y. Ch. 1846) (using the word “device”); UPTON, supra note 61, at 34-36.
was made elegantly. 84 In this case, as in all of the above cases regarding device, the court calls a “device” what a court or commentator today would call a “design mark.” 85

The word “symbol” does not appear in the cases from 1836 until 1849. In the 1849 case, *Amoskeag Mfg. Co. v. Spear*, 86 the court clearly claims that the definition of trademarks should extend to “symbols” as follows:

Every manufacturer and every merchant for whom goods are manufactured, has an unquestionable right to distinguish the goods that he manufactures or sells, by a peculiar mark or device, in order that they may be known as his in the market for which he intends them, and that he may thus secure the profits that their superior repute as his, may be the means of gaining.

... The owner of an original trade mark, has an undoubted right to be protected in the exclusive use of all the marks, forms or symbols that were appropriated as designating the true origin or ownership of the article or fabric to which they were affixed; but he has no right to an exclusive use of any words, letters, figures or symbols, which have no relation to the origin or ownership of the goods, but are only meant to indicate their name or quality. He has no right to appropriate a sign or symbol which, from the nature of the fact which it is used to signify, others may employ with equal truth, and therefore have an equal right to employ, for the same purpose. 87

In 1853 the Supreme Court of New York that the definition of trademark would include both the word “symbol” and “device:” 88

The principle is well settled that a manufacturer may, by priority of appropriation of names, letters, marks, or symbols of any kind, to distinguish his manufactures, acquire a property therein, as a trade mark, for the invasion of which an action for damages will lie; and in the exclusive use of which he may have protection, when necessary, by injunction. ... In all cases where names, signs, marks, brands, labels, words or devices of any kind can be advantageously used to designate the goods or property, or particular place of business, of a person engaged in trade or manufacture, or any similar business, he may adopt and use such as he pleases, which are adapted to that end and have not been before appropriated; and no other person can lawfully imitate

---

84. Patridge v. Menck, 2 Sand. Ch. 622 (N.Y. Ch. 1846); UPTON, supra note 61, at 34-36.
85. See Patridge v. Menck, 2 Sand. Ch. 622 (N.Y. Ch. 1846) (using the word “device” to describe the image on a matchbox).
87. Id. at 605-06; UPTON, supra note 61, at 86.
88. Stokes v. Landgraff, 17 Barb. 608 (N.Y. Sup. Ct. 1853); UPTON, supra note 61, at 87.
them, and by that means sell his own goods or property, or carry on his
business, as the goods, property or business of the former.89

By 1860, Francis H. Upton, the primary trademark commentator of the day,
reported that the definition of trademark included both the words “device” and
“symbol.”90 “A trade mark is the name, symbol, figure, letter, form or device,
adopted and used, by a manufacturer, or merchant, in order to designate the
goods that he manufactures, or sells . . . .”91

In 1877, the U.S. Supreme Court gave trademarks the expansive definition
that is very similar to the definition used by courts today.92 In McLean v.
Fleming, the Court defined a trademark as “a name, symbol, figure, letter, form,
or device . . . .”93

Subsequently, in 1878, another trademark commentator of the time, Charles
E. Coddington, defined the trademark in the following terms: “A trademark is a
name, symbol, figure, letter, form or device, adopted and used by a manufacturer
or merchant to designate the goods he manufactures or sells and to distinguish
them from the goods of another.”94

One of the earliest reported alleged nontraditional trademark case where
“device” was used to describe the trademark was A. Leschen & Sons Rope Co. v.
Broderick & Bascom Rope Co.95 In this case, a red stripe was placed on wire
rope.96 It was this colored strand that was claimed to be the plaintiff’s
trademark.97 The United States Supreme Court held that color alone was too
indefinite to constitute a trademark.98 The court called the colored strand
imprecisely defined and “manifestly too broad.”99

In Coca-Cola Co. v. Koke Co. of America, in referring to trademarks as a
“device,” the court held that “[t]he product including the coloring matter is free
to all who can make it if no extrinsic deceiving element is present.”100 Although

---

89. Id.
90. UPTON, supra note 61, at 9.
91. UPTON, supra note 61, at 9.
92. See McLean v. Fleming, 96 U.S. 245, 254 (1877) (describing a trademark as “a name,
symbol, figure, letter, form, or device”).
93. McLeav v. Fleming, 96 U.S. 245, 254 (1877) (citing UPTON, supra note 61 as controlling
authority).
94. CHARLES E. CODDINGTON, A DIGEST OF THE LAW OF TRADEMARKS § 22 (1878). As support
for this proposition, the author relies on the following cases: (1869) Phila. Ct. of Com. Pleas.,
Ferguson v. Davol Mills, 7 Phila. 253; S. C., 2 Brewster 314. However, these cases could not be
located and appear lost.
95. See A. Leschen & Sons Rope Co. v. Broderick & Bascom Rope Co., 201 U.S. 166 (1906)
(using the word “device”).
96. A. Leschen & Sons Rope Co. v. Broderick & Bascom Rope Co., 201 U.S. 166, 170
(1906).
97. Id. at 170.
98. Id. at 170. “Whether mere color can constitute a valid trade-mark may admit of doubt.” Id.
at 171.
99. Id. at 171.
100. Coca-Cola Co. v. Koke Co. of Am., 254 U.S. 143, 147 (1920).
the mark COCA-COLA has secondary meaning, it did not give the plaintiff a right to prevent the use of “dope” as used on a brown, bubbly beverage.

All of this significant use of the words “device” and “symbol” to define a trademark did not motivate the drafters of the Trademark Acts of 1870, 1876, 1881, 1882, 1905, or 1920 to include the word “device” in defining a trademark and only the Trademark Act of 1920 used the word “symbol.” Therefore, when the Lanham Act takes up the yoke of “symbol” and “device” it had some common law history on which to rely; however, that history actually teaches away from the current broad understanding of these words. Most importantly, in 1988 Congress expressly “expanded” the definition of a trademark and retained “symbol or device” so as not to preclude the registration of colors, shapes, sounds or configurations where they function as trademarks.

The problem, of course, was that “symbol” and “device” had almost exclusively meant design marks, not nontraditional trademarks. In the only case where it seems to have been used to define color strips as used on rope, the Court rejected that claim as too broad. Then, in 1995, when the Supreme Court relied on this very language to claim that color alone is appropriate subject

101. Id. at 147.
108. Id.
110.  See supra Part III.
111. In a Senate Report accompanying the Trademark Law Revision Act of 1988 Congress expressed its intent to broaden the definition of “trademark” and retain “symbol and device” to encompass nontraditional trademarks such as color, sound, etc. S. REP. NO. 100-515, at 44 (1988).
112. Id. See also The United States Trademark Association Trademark Review Commission Report and Recommendations to USTA President and Board of Directors, 77 TRADEMARK REP. 375, 421 (Sept.-Oct. 1987) (expressing intent to allow registration of things “such things as a color, shape, smell, sound, or configuration which functions as a mark”); Moon-Ki Chai, Note, Protection of Fragrances Under the Post-Sale Confusion Doctrine, 80 TRADEMARK REP. 368, 371-72 (July-Aug. 1990) (“The Committee ‘determined that the terms ‘symbol, or device’ should not be deleted or narrowed to preclude registration of such things as color, shape, smell, sound, or configuration which functions as a mark.’”)
113. See supra Part III.
matter for a trademark, it is not difficult to see how people might conclude that this was a “stunning breakthrough” for nontraditional trademarks.\(^{115}\)

Little in the etymological history of the words “symbol” and “device” hints at the expansive meaning now ascribed to those terms.\(^{116}\) In historical terms, these words have changed from meaning a design mark to now meaning the vertical motion of a car door opening or the vertical spray of water from the back of a personal water craft.\(^{117}\) As Congress expressly admitted in 1988, this is an “expansion” of what heretofore had been the trademark right.\(^{118}\) As Gilson says, this is “stunning.”\(^{119}\) In context, this expansion is not explained or warranted. As persuasive plaintiff attorneys gradually made their cases, the expansion occurred. This expansion can be ascribed to the persuasive powers of the common law lawyer, not to any particular legislative enactment. In this context, we can see that the \textit{Qualitex}\(^{120}\) case was a radical departure from the 150 years of common law case history that preceded it.\(^{121}\)

IV. COLOR

The issue of registering nontraditional marks received a major incentive when the Supreme Court held that there was no \textit{per se} rule against color alone being granted trademark status.\(^ {122}\) In \textit{Qualitex v. Jacobson}, the Supreme Court flippantly dismissed objections to color alone being a trademark that had stood for decades.\(^ {123}\)

In \textit{Qualitex}, the Supreme Court determined that the previous belief that there was a “\textit{per se}” rule against granting trademark protection for color alone was in error.\(^ {124}\) In \textit{Qualitex}, Justice Breyer’s opinion squarely addresses whether or not a color alone may be suitable trademark subject matter.\(^ {125}\) Justice Breyer concluded that because the Lanham Act says “any word, name, symbol, or device” may be a valid trademark, as long as it acts to identify source, color alone may also play the role of a valid trademark.\(^ {126}\)

Although taken for granted now in the United States and sometimes misstated,\(^ {127}\) the United States remains the only jurisdiction in the world to so

\(^{115}\) See Gilson & LaLonde, supra note 4.
\(^{116}\) See supra Part III.
\(^{117}\) See supra notes 26-27.
\(^{118}\) See supra note 111.
\(^{119}\) See Gilson & LaLonde, supra note 4.
\(^{121}\) See supra Part III.
\(^{122}\) See \textit{Qualitex}, 514 U.S. 159.
\(^{123}\) Id. See infra Part XIII.
\(^{125}\) Id.
broadly recognize rights in color alone. The European Union, the World Intellectual Property Organization and the Madrid Protocol all arguably allow for and make specific provisions for the registration of color alone as a trademark, however, judicial support in that context is sparse.

In the United States, registration of color alone is justified by categorizing it as a “device” under the Lanham Act. This was contemplated when the Lanham Act was amended in 1988. That is, because the statutory definition of a trademark now includes the terms “symbol or device,” and this no longer means “design mark,” the possibility of registering things, such as color, that are not traditional word marks is supported by the statute itself.

This result has been harshly criticized by commentators. Perhaps the most persuasive and direct attack came recently from Ann Bartow. Professor Bartow argued that color alone should not be recognized as a trademark. To do so, she argues, allows companies to monopolize the aesthetic and communicative attributes of color. Given that color is also inherently functional and the consequent uncertainties created by protecting color, Professor Bartow would never allow color alone to act as a trademark.

Professor Bartow seems correct. Protecting color per se seems to extend the trademark beyond where it is needed to encourage use by corporations to reduce transaction costs. More importantly, though, color alone is not a frequently
litigated subject matter, even though the Court in Qualitex seemed to believe it was worthy of Supreme Court attention.138

Virtually all of the reported cases regarding nontraditional trademarks are litigation over color.139 Of the total reported cases since 1947, less than 1% of the total claim color alone as a trademark;140 however, of the 35 nontraditional trademark cases reported 30 (86%) of them are about color.141 Although there are 345 trademark applications for color alone (44% of the total), 152 of these marks (nearly 45%) were either abandoned before they reached registration or were cancelled for nonuse once registered.142 It takes on average 37 months to obtain a registration for a color mark.143 Therefore, it appears that commentators and courts alike exaggerate the significance of this nontraditional trademark.

Arguably, color may be the least subjective of the nontraditional trademarks. Using the Pantone Matching System, a color-coding system, subjectivity may be lessened.144 This coding system gives every shade of every color a number.

141. Id.
142. See infra app. C.
143. See infra app. D.
144. See generally Pantone, http://www.pantone.com (associating identifying numbers with a range of colors).
145. Id.
The Pantone Matching System is an internationally recognized color-coding system. Pantone Corporation lists over 1000 colors and it has assigned each color a number code. This allows Pantone to match colors uniformly worldwide. Therefore, one might claim Pantone Matching System 300, which is a blue color. However, for Thrifty, Inc., claiming this number was not enough to overcome an Examiner’s rejection based on a claim that the Applicant failed to provide an adequate description of the mark.

However, even if the subjectivity of color marks is lessened with the Pantone Matching System, its use in American trademark law would be limited. Context is an important factor considering whether a mark is infringed. That is, courts will not tolerate taking a mark out of context and, for example, doing a side-by-side comparison to determine if the two marks are confusingly similar. Therefore, although it is, of course possible to run two different colors through a spectrograph to determine similarity, no court would admit that into evidence. To determine similarity, the mark must be viewed as consumers confront the mark in the marketplace. Consumers do not look at the color of a product through a spectrograph before making a purchase; therefore, it would not be admissible.

V. SOUND

There have been 336 applications for sound before the PTO. This amounts to over 43% of all nontraditional trademarks. Sound and color (44%) make up 87% of all nontraditional trademarks. Four of the 35 reported cases (11%) are sound marks and the plaintiff has prevailed only once.

The PTO apparently thinks rather lightly of sound marks as the only list of these marks appears in the “Kids” section of the PTO website which also includes things like games, puzzles other links, and the most peculiar looking swashbuckling Pac-Man-like character yielding a wooden sword labeled “stop

146. Id.
147. Id.
148. Id.
149. In re Thrifty, 274 F.3d 1349 (Fed. Cir. 2001) (sustaining rejection where applicant’s claimed color blue on a wall was materially altered when applied to several other objects).
152. See infra app. A.
153. See infra app. A.
154. See infra app. A.
piracy.” The peculiar part is that the character appears to be attractive to a child interested in pirates, but the content of that portion of the site is clearly intended for someone well educated in the art of knock-off goods.

Figure 1: PTO’s Graphic User Interface for Children of “Pirated” Marks

It certainly appears that the registrants or applicants do not take these applications as lightly as the PTO does. Reviewing the application records using the TARR (Trademark Applications and Registrations Retrieval) database of the PTO discloses the fact that applicants for sound marks do not have an easy time obtaining registrations. It takes an average of just under 24 months to obtain a trademark registration after application. The average pendency for all marks is 13.9 months. These applications are often refused for being functional or for failing to function as a trademark that can be used to identify and distinguish applicant’s goods from others and to identify source. That is, it takes significantly more time and therefore costs significantly more to obtain a registration for a sound mark rather than a traditional mark.

Perhaps the most famous case involving a claimed sound as a trademark occurred in 1994 when Harley Davidson filed a trademark application for the sound of its motorcycle idling. In the application, Harley Davidson described this sound as saying the word “potato”, slowly, over and over.

158. Whoever in the PTO dreamed up this graphic certainly doesn’t have kids.
159. See infra app. B.
160. See infra app. D.
Davidson applied for this mark in 1994, it was allowed by the PTO in 1995 and five years later in 2000 it was expressly abandoned by Harley Davidson after several parties, including Suzuki, Honda, Kawasaki and Polaris, filed oppositions and three months before the TTAB sustained those oppositions.

One of the first sound trademarks to be registered in the United States was filed in 1970 by NBC Universal, Inc. for the now famous “NBC chimes.” The identification of goods for this registration reads as follows:


Of course, some sound marks have attained an iconic reputation. Everyone knows the roar of the MGM lion, first used in 1924, and the Harlem Globetrotters whistling the song Sweet Georgia Brown, first used in 1948. However, few people would recognize “Hisamitsu” sung over the sound of four musical tones, E, A, E, and F sharp. The first three notes being eighth notes and the final note being a tied eighth and half note,” for medicated transdermal patches or the loon call of the Minnesota State Lottery. That is, few have become iconic, like the Menards song about how we can all “Save Big Money at Menards.” Most of the trademark applications or registrations for sound marks went the way of Anheuser-Busch’s howling wolf. In fact, of the 336

---

165 See supra note 163.  
168 Id.  
170 Id.  
172 Id.  
sound mark applications in the 62-year history of the Lanham Act, 28 were cancelled for nonuse of the registered mark and 112 were abandoned before they reached the registration stage.\footnote{177}

Sound marks constitute 43\% of all nontraditional mark applications, yet 140 of the initial 336 applications didn’t survive to registration or were cancelled thereafter.\footnote{178} Very few would be deemed “famous” under the Dilution Act.\footnote{179}

VI. SCENT

The most popular case where the scent of a product was recognized as a trademark was \textit{In re Clarke}.\footnote{180} In this case, Celia Clarke, doing business as Clarke’s OSEWEZ, produced a yarn that smelled like plumeria flowers.\footnote{181} The Trademark Trial and Appeal Board found the scent to be distinctive of the yarn, possessed secondary meaning and, therefore, upheld Mrs. Clarke’s application to register the mark as used on kits where the purchaser would knit a skunk that smelled like plumeria.\footnote{182}

Perhaps no American has played a larger role in getting scents recognized as trademarks than James E. Hawes. Hawes first advocated for the registration of scent trademarks in 1989.\footnote{183} Hawes represented Cecilia Clarke in her petition before the TTAB to register the scent of her yarn in 1990.\footnote{184} Unfortunately for Hawes, on September 29, 1997, Clarke’s registered trademark for the plumeria scented thread was canceled by the PTO for a failure to file a Continuing Use Affidavit under Section 8 of the Lanham Act.\footnote{185} That is, the mark was canceled for nonuse.

It appears that the registration for the scent of Clarke’s yarn was more significant for Hawes than it was for Clarke. Although this case is depicted by some as a major breakthrough that others ought to follow,\footnote{186} very few applicants have, in fact, obtained trademark registrations for scents and in other places this

\begin{footnotesize}
\footnote{177}{See infra app. C.}
\footnote{178}{See infra app. C.}
\footnote{179}{\textit{See David S. Welkowitz, Trademark Dilution: Federal State, and International Law} 177 (2002); 15 U.S.C. §1125(c) (fame is ephemeral in nature and difficult to identify, which is why Congress laid out multiple factors including “(i) The duration, extent, and geographic reach of advertising and publicity of the mark, whether advertised or publicized by the owner or third parties. (ii) The amount, volume, and geographic extent of sales of goods or services offered under the mark. (iii) The extent of actual recognition of the mark. (iv) Whether the mark was registered under the Act of March 3, 1881, or the Act of February 20, 1905, or on the principal register.”)}
\footnote{180}{\textit{In re Clarke}, 17 U.S.P.Q.2d (BNA) 1238 (T.T.A.B. 1990).}
\footnote{181}{Id.}
\footnote{182}{Id.}
\footnote{183}{James E. Hawes, \textit{Fragrances as Trademark}, 79 TRADEMARK REP. 134 (1989).}
\footnote{184}{\textit{In re Clarke}, 17 U.S.P.Q.2d (BNA) 1238 (T.T.A.B. 1990).}
\footnote{186}{Faye M. Hammersley, Comment, \textit{The Smell of Success: Trade Dress Protection for Scent Marks}, 2 MARQ. INTELL. PROP. L. REV. 105 (1998).}
\end{footnotesize}
idea has been harshly criticized. In fact, there are currently only 12 registered trademarks for scents in the United States and only 28 applications. Nine of the 28 applications have either been abandoned or subsequently cancelled for nonuse. It takes an average of 26 months to obtain a registration. There are no reported cases where a scent trademark was at issue.

In fact, lessons from osphresiology, the science of smells, teach us that scents are inappropriate for trademark protection. The actual scent that one recognizes is affected by temperature and humidity. Therefore, not only is the scent that is projected from an article different according to these and other variables, the receiver also does not sense the same scent from time to time. Human beings have an enormous capacity to remember sounds and sights. The human brain can only remember sixteen smells for a long period of time. Also, science has not determined the role that, for example, simple cold remedies such as Zicam can play in impairing one’s sense of smell.

Few scent marks are registered in European Union. The “smell of fresh cut grass,” for example, has been issued for use on tennis balls.

Therefore, because scents are perceived differently over time and because each person’s subjective olfactory sense differs based, for example, on whether one has a cold or not, scent trademarks do not possess the incident of consistency that is required to make a scent worthy of trademark protection.

VII. TASTE

To date, there have been nine trademark applications in the United States for taste. Thus far, no registrations have been issued and all nine applications

188. See infra app. A.
189. See infra app. C.
190. See infra app. D.
191. Churovich, supra note 187, at 301.
198. See infra app. A.
199. See infra app. A.
have been abandoned. Although a flavor may be a “device” under the Lanham Act and therefore not statutorily barred, no one has convinced a court or the PTO that a flavor deserves trademark protection. The best case for flavor trademarks apparently is in the pharmaceutical industry.

In a very closely watched case by flavor lovers, N.V. Organon, a Dutch company, failed to convince the TTAB that its flavor orange deserved trademark protection on antidepressant medication.

Organon had argued that its orange flavor was a different orange flavor than other medicines on the market with an orange flavor. It argued that the orange flavor has nothing to do with the product itself. It did not make the product cheaper or stronger or weaker in dosage. It also argued that doctors when prescribing the drug do not concern themselves with the flavor of the drug. Organon also argued that the orange flavor did not affect the function of the medication—regardless of how it tastes, the medicine still works the same.

The TTAB disagreed. The TTAB held that the orange flavor could not function as a trademark because pharmaceutical companies add flavor to their medicines on a frequent basis. Organon did not attempt to register a distinct flavor of orange; it attempted to register the flavor orange. Even if a unique flavor of orange was at issue, not only did the applicant not supply any evidence as to this flavor, but flavor itself is too subjective to be perceived in a constant way by consumers and therefore consumers could not use the flavor to distinguish Organon’s products other products.

Some predict that the recognition of flavors as trademarks may be inevitable. This may be possible when, as argued by Hobbs, something like an orange flavored National Law Journal becomes something people might buy. In the advertisements for Coke Zero, the Coke Cola Company satirizes attorneys and attempts to get fictitious legal advice that would support their clam for trademark protection of Coke Zero. Although humorous, the point of the

200.  See infra app. C.
201.  See infra app. A.
204.  Id. at 1641.
205.  Id. at 1641.
206.  Id. at 1641.
207.  Id. at 1641.
208.  Id. at 1641.
210.  Id. at 1649-50.
211.  Id. at 1650.
212.  Id. at 1650.
214.  Id.
215.  For a humorous take on Coke Cola attempting to obtain trademark rights to its beverage
advertisement is obvious. Coke considers the flavor of its beverage to be an indicator of the source of its product. Coke, to date, has not yet filed a trademark application nor been a party to a reported case regarding the flavor of Coke Zero.

VIII. TACTILE

In the United States, there is one mark registered for the feel or texture of the mark.\(^{216}\) There have been five applications\(^ {217}\) and only one reported case (where the plaintiff did not prevail).\(^ {218}\) The only registered mark is owned by American Wholesale Wine and Spirits, Inc., which registered the velvet texture covering on a bottle of wine.\(^ {219}\) This registration issued in October of 2006.\(^ {220}\) There were two other tactile marks that were applied for registration for crystals applied to ties\(^ {221}\) and crystals applied to art reproductions.\(^ {222}\) Both marks were refused registration because the marks failed to operate as a trademark and appeared to be merely decorations.\(^ {223}\) The applicant, Karen’s Krystals, Inc., did not respond to the Office Actions and so the marks were abandoned by the PTO.\(^ {224}\) Worldwide, there appears to be only one other trademark registered claiming the feel of an appellation.\(^ {225}\) In Germany, the word UNDERBERG in Braille for use on beer and other beverages is registered.\(^ {226}\)

known as Coke Zero and enforcing those rights against itself, see Coke sues Coke Zero for Infringement, http://www.youtube.com/watch?v=pv8YgrqUCVU (Uploaded July 26, 2006).

216. See infra app. A.

217. See infra app. A.


224. See supra notes 221-22.


226. Id.
Tactile marks remain a very odd commodity for trademark offices worldwide. To date, only three applications have been filed worldwide for the feel of a mark.\textsuperscript{227} Even so, one influential commentator has noted apparently approving of tactile marks that “[i]n terms of a brand, the feeling and textures of a product can play on a consumer’s emotions and can also relate directly to his or her perception of quality.”\textsuperscript{228}

IX. HOLOGRAM

As the Lanham Act and the \textit{Qualitex} decision refer to “any” word, name, symbol or device, some creative entities have even attempted to apply for holograms as their trademark.\textsuperscript{229} In \textit{In re The Upper Deck Co.}, the Applicant attempted to register a mark it described as follows:

The mark consists of a hologram device which is applied to the goods, trading cards. The mark is discrete from and does not constitute a part of the subject matter of the trading card. Neither the size nor the shape of the hologram device, nor any content which may be represented within the hologram device, nor the positioning of the hologram device on the trading card are claimed as features of the mark.\textsuperscript{230}

This application was denied by the TTAB on the grounds that the holographic device failed to function as a trademark.\textsuperscript{231} Additionally, by definition a holographic mark would be perceivable from multiple angles and therefore more than one view of the mark here was applied for registration, the TTAB found that this necessitated the finding that more than one mark was applied for registration, something that the Lanham Act prohibits.\textsuperscript{232} That is, the Applicant was attempting to register the idea of a hologram on playing cards, not a specific, single mark.\textsuperscript{233} Based on this understanding of the interaction between the Lanham Act and holographic marks, it is difficult to see how a holographic mark could ever be registered.

However the PTO has registered 11 of 19 applications for trademarks that consist of a hologram.\textsuperscript{234} Thirteen of the 19 applications have been abandoned or cancelled for nonuse.\textsuperscript{235} It takes an average of almost 30 months to convince the PTO that a hologram mark should be registered.\textsuperscript{236} Once again, the data does

\begin{itemize}
\item \textsuperscript{227} See supra Part VIII.
\item \textsuperscript{230} Id. at 1689.
\item \textsuperscript{231} Id. at 1691-92.
\item \textsuperscript{232} Id. at 1689-91.
\item \textsuperscript{233} Id. at 1689-91.
\item \textsuperscript{234} See infra app. A.
\item \textsuperscript{235} See infra app. C.
\item \textsuperscript{236} See infra app. D.
\end{itemize}
not support the conclusion that hologram marks are a significant part of the American trademark universe. There were no reported cases where a hologram mark was the subject of the trademark litigation.

X. MOTION

In 1957, the PTO granted a registration for the motion of a coin spinning on top of a hard surface to connote a particular Bank in Duluth, Minnesota. This is one of the earliest registered nontraditional trademarks in America. Since then, the PTO has granted registrations for motion marks such as preprogrammed search lights projected into the sky for use on search lights, to promote public awareness of the negative effects of smoking, a "stylized capital letter ‘Y’ within a loosely defined circle orbited clockwise by a second, smaller, loosely defined circle containing a stylized question mark," and, to beat all, for use on automobiles, the motion of a Lamborghini car door opening vertically rather than horizontally.

There are apparently three types of motion trademarks. The first is computer generated moving image marks such as Nokia’s Community Trademark registration described as “The mark comprises an animation which consists of four images depicting hands coming together, shown in succession from left to right and from top to bottom.” The second type of motion marks consists of the motion of the product itself as in the vertical motion of the Lamborghini doors or the fan-shaped spray that leave the back of a Yamaha personal watercraft. The third type of motion mark is a hand gesture.

The complexity of claiming motion as a trademark came to a head in the Diamond Cutter dispute with Jay-Z. A former TV variety professional


238. Id.


245. For further documentation of these and other motion marks and some extremely helpful analysis, see Lesley Matty, Note, Rock, Paper, Scissors, Trademark? A Comparative Analysis of Motion as a Feature of Trademarks in the United States and Europe, 14 CARDOZO J. INT’L & COMP. L. 557 (2006).

246. Plaintiff Diamond Dallas Page’s First Amended Complaint for (1) Trademark Infringement; (2) False Designation of Origin; (3) Federal Dilution; (4) Unfair Competition; (5) Injury to Business Reputation; (6) Copyright Infringement (7) Misappropriation of Trade and
wrestler by the name of Page Falkinburg, whose wrestling name was “Diamond Dallas” claimed to have coined and adopted as a trademark a hand gesture that can best be described as attempting to make a diamond shape out of one’s two hands. To make the gesture, with open hands, face your palms out. Then, bring your forefingers and thumbs together just so they touch. If you do that, Diamond Dallas claims you have infringed his trademark. This is the gesture that Jay-Z made in a picture on an album cover and so Diamond Dallas filed suit.

This is a perfect example of why the TTAB’s objection to holographic marks ought to apply to motion marks as well. That is, Diamond Dallas is claiming a trademark in the idea of a hand gesture rather than a specific mark that can be identified. Therefore, motion marks should not be registered and should be unenforceable.

There are actually 18 total trademark applications for motion marks and 13 subsisting registrations. It takes a mere 20 months to obtain trademark registration status for a motion. There were no reported cases where a motion mark was the subject of the litigation.

XI. THE SKY IS THE LIMIT

With the venerable explosion of interest in nontraditional trademarks, applicants and claimants alike in the United States are interpreting “any” to really mean “any.” Therefore, if distinctive and not functional, one well-respected commentator claims that even “sales techniques” would be appropriate subject matter for trademark protection. Gilson claims as follows:

adoption procedures for Cabbage Patch Dolls were found to be protectable trade dress as were the layout and overall appearance of mail-order catalogs and a sales brochure, car service reminder letters and the “combination and arrangement” of features in a sales report. Advertisements for sunglasses with actors posing as vampires—and other vampire-related marketing—were found to be protectable trade dress, as was the overall image of Marlboro cigarette advertising.
If you cannot use a cowboy or scary people wearing sunglasses without violating someone else’s trademark rights, I submit that we have taken what used to be a firmly grounded right of prior appropriation to support and maintain fair competition and moved it right past the stratosphere and right to the mesosphere. As one reporter has put it, “What’s next? A fee for looking?”

XII. HARMONIZATION

There are three competing understandings of the word “harmonization.” On one hand, it could mean an effort to make all trademark law in the world uniform so that trademark owners know the bounds of protection anywhere in the world by understanding the laws of one country, thereby allowing trademark owners a higher degree of confidence in the worldwide market. It might also be understood to provide “an approximation of different rules in order to minimize any conflict that their differences might generate; . . . not necessarily imply[ing] the replication of rules.” Or harmonization may mean an effort to create a relationship between things that implies the creation of accord or consonance.

One way or the other, the comments herein are equally applicable to each and every definition of or motivation for harmonization.

A. Civil Law

The Civil Law of trademarks is oriented on the registration system and more oriented to the benefit of the trademark owner. Trademark rights subsist upon registration. With few exceptions, without registration, the trademark right does not exist. In Japan, for example, the most common word for “trademark”

254. A further example is the registered mark of a British female face for use on educational services. See Sally Ramage, Reg. No. 3,440,915 available at http://tarr.uspto.gov/tarr?regser=serial&entry=79%2F035414&action=Request+Status (“The mark consists of the color brown appears in the woman’s eyes and portions of her hair, the color red appears in portions of her hair, the color pink appears in her lips, and the color tan appears in the woman’s face and neck.”).


261. These exceptions come in the form of limited rights granted to the first user of a mark that becomes famous without registration. In such a setting, subsequent registrants must accept and allow such prior use. See, e.g., KENNETH L. PORT, TRADEMARK AND UNFAIR COMPETITION LAW AND POLICY IN JAPAN 85-90 (2007) (discussing the effect of use in Japan).

262. Max Vern, A Trademark Visa -- Aspects of International Trademark Use and Protection,
(shohyo) is not even used to describe an appellation of source unless it is registered.263

The criteria for registration are very plainly spelled out. The various Civil Law trademark statutes state the specific types of trademarks that may be registered and those that may not be registered.264 Although a WIPO survey indicates that many Civil Law countries allow registration of nontraditional trademarks, very few actually have registered significant numbers of nontraditional trademarks.265

Since the inception of the CTM system in 1996,266 there has been some application activity for nontraditional marks in, for example, the Community Trademark system for the European Union. The data on CTM applications for nontraditional trademarks post 1996 is as follows.267


263. KENNETH L. PORT, TRADEMARK AND UNFAIR COMPETITION LAW AND POLICY IN JAPAN 95 (2007).

264. See Trademark Act, Law No. 127 of 1959, arts. 3-4 (Japan), translated and available at http://www.japaneselawtranslation.go.jp/law/detail/?id=45&vm=04&re=02 (prescribing six types of things that can be registered and 19 types of marks that may not be registered in Japan); Article L711-1 (France) available at http://195.83.177.9/upl/pdf/code_35.pdf (providing examples of marks worthy of registration). The EU denied the registration of a mark that consisted of “balsamic fruity scent with a slight hint of cinnamon” for educational services and food and drink because the chemical formula did not satisfy the technical requirement that the mark be able to be depicted graphically. Case C-273/00, Sieckmann v. Deutsches Patent- und Markenamt, 2002 E.C.R. I-11737, available at http://curia.europa.eu/jcms/jcms/j_6/ (search for Case C-273/00).


266. The Community trademark (CTM) system was introduced in 1996 by The Trade Marks and Designs Registration Office of the European Union (OHIM). CTM covers the entire European Union and is valid in all 27 Member States. To date there are approximately 600,000 CTM registered trademarks. The Trade Marks and Designs Registration Office of the European Union, Trade Marks, http://oami.europa.eu/ows/rtw/pages/CTM/index.en.do;.

267 Color marks are excluded because the database does not differentiate between color alone marks and color as an attribute of the mark. This study regards color alone as a mark. See The Trade Marks and Designs Registration Office of the European Union: CTM-ONLINE – Trade Mark Consultation Service – Basic, http://oami.europa.eu/CTMOnline/RequestManager/en_SearchBasic (last visited Sept. 22, 2011) (searchable database does not allow searching for color alone as a mark).
The two things in this graph that are striking are that the number of sound marks issued has dropped significantly and that the number of sound marks applied for has risen sharply. This inverse relationship between the number of applications for sound marks and the number that are actually registered is indicative of the lack of experience Civil Law countries have with examining nontraditional trademarks.

To allow the registration of nontraditional trademarks (or any trademark whatsoever), the Civil Law requires a specific statute stating the specific type of mark to be registered. This is because one can say that Civil Law systems, in general, are far more positivist in orientation than the Common Law system of the United States.\footnote{268. Errol E. Meidinger, Learning Sustainability, 10 BUFF. ENVTL. L.J. 211, 253 (2003); Alan Uzelac, Comparative Theme: UNCITRAL Notes on Arbitral Proceedings: A Regional View, 4 CROAT. ARB. Y.B. 135, 153 (1997); Karen Engle, Comparative Law as Exposing the Foreign System's Internal Critique: An Introduction, 1997 UTAH L. REV. 359, 362 (1997) (quoting Marie-Claire Belleau, The “Juristes Inquiets”: Legal Classicism and Criticism in Early Twentieth-
interpret and apply it. Although that is an overly broad generalization that is used as a litmus test in confirmation hearings of Supreme Court nominees before the United States Senate, it is more accurate in Civil Law jurisdictions than in the United States.

Trademark statutes in Civil Law jurisdictions are written with precision so that interpretation is not necessary. If a statute needs to be interpreted, it is synonymous with saying that that statute needs to be revised or amended. Civil Law systems do not recognize the notion of stare decisis. As judges are only supposed to applying the statute in a more mechanical manner, there is no need for judges to look for guidance outside of the statute itself.

Most Civil Law systems would not allow a trial by jury for a trademark violation. Rather, the initial trial would be a before a panel of judges. In many Civil Law systems, judges are rotated frequently. It is common during a trial to have different judges rotate in and out of the panel. As judging is a more mechanical endeavor where judges are simply applying the statute as written to any given set of facts, it should not matter who, personally, might be on trial panel.

Discovery in Civil Law systems is quite limited. As pleading is fact-based, plaintiffs are not allowed to go on fishing expeditions during discovery. Litigants must produce any document specifically named, but, for

---


270. See, e.g., Confirmation of Ruth Bader Ginsburg as Supreme Court Justice Before S. Judiciary Comm., 103d Cong. (1993) [hereinafter Confirmation Hearing] (Ruth Bader Ginsburg stated that judges do not make law, they interpret and apply the law).

271. Merryman & Pérez-Perdomo, supra note 268, at 61 (law is a precise science and should be able of clear communication. “The law is thus enriched, and enriched by a purely scientific method”) (quoting Rudolph Sohm, The Institutes: A Textbook of the History and System of Roman Private Law 30 (James Crawford Ledlie, trans., Erwin Grueber ed., 3d ed. 1907)).


273. Merryman & Pérez-Perdomo, supra note 268, at 45-47.


275. Id. (stating that judges are rotated “frequently.”)


277. See Mirjan Damaska, Presentation of Evidence and Factfinding Precision, 123 U. Pa. L. Rev. 1083, 1088-89 (1975) (discussing role of the judge in gathering facts); Benjamin Kaplan et al., Phases of German Civil Procedure I, 71 Harv. L. Rev. 1193, 1234-35 (1958) (“proof-taking is conducted mainly by the court” in Germany).

278. See Stephen N. Subrin, Discovery in Global Perspective: Are We Nuts?, 52 DePaul L.
example, defendants are never asked open-ended interrogatories or requests for documents for which one would need a truck to deliver the responses.

The appeal process in Civil Law countries is a continuation of the original trial. Once the trial court comes to a conclusion, the appellate court retries the case. One can produce new evidence, give different responses, change the theory of the case, etc. Litigants are not bound to the trial record of things said. Therefore, there is no concept of standard of review when a Civil Law appellate court considers a trial court outcome. In fact, because judges rotate, in a lengthy trial, it is not uncommon for a trial court judge to be reassigned to the appellate bench while the appeal is pending. Such a Civil Law judge would not recuse him or herself. Rather, this judge would join the appellate bench for an appeal of a trial that judge tried.

In Civil Law countries, trademark related surveys are either non-existent or not relied on anywhere near as much as in the United States. As such, a Civil Law judge may only hear about how the public views the mark, whether they would be confused by a competing use, or how strong the mark is from the litigants themselves. Surveys are considered far too unreliable to be used in a


279. See Kuo-Chang Huang, INTRODUCING DISCOVERY INTO CIVIL LAW 207-08 (2003) (restriction on interrogatories intended to prevent overly broad “fishing expeditions”).

280. See also Brendan Sweeney, The Role of Damages in Regulating Horizontal Price-Fixing: Comparing the Situation in the United States, Europe and Australia, 30 MELBOURNE U. L.R. 837 (2006) (the job of collecting documents is a judicial function).

281. See, e.g., Class Action (InterScope Communications 1991) (A film starring Gene Hackman where the defense attempts to bury the plaintiff in a mountain of discovery to try to win the case).


283. Id. at 460-61.

284. Id. at 460-61.

285. Id. at 461.

286. See John H. Langbein, The German Advantage in Civil Procedure, 52 U. CHI. L. REV. 823, 855-58 (1985) (stating that there is “[n]o presumption of correctness attaches to the initial judgment” in German appellate courts.)

287. See Ramseyer, supra note 274, at 116 n. 16 (stating that Japanese judges rotate “frequently”).


On the other hand, individual judges make up their own mind regarding these questions. Therefore, it is not unusual in Civil Law jurisdictions for judges to reach decisions that are clearly not responsive to general public perception. Considering there are no juries, this means that the only way for a judge to answer whether the mark is strong, for example, is to look at the evidence him or herself and reach a conclusion.

The reason it is more appropriate to treat trademarks in such a positivist manner is because Civil Law jurisdictions consider the trademark itself as property. When one owns a registration one owns a trademark. There is no distinction conceptually made between the trademark and the trademark right. All rights in trademarks are determined by the scope of the classification of goods or services claimed in the registration itself. It would be appropriate to say that a trademark owner owns the right to use the mark on all goods in the classification for which it is registered.

That is, trademarks in Civil Law jurisdictions are corporeal things. Civil Law judges do not make tortured analogies to “quasi property rights” as they do in the United States. Trademarks are property, period.

The theory that trademarks are property is doctrinally evidenced by the fact that trademark rights in Civil Law jurisdictions are severable. A trademark

---

292. See MERRYMAN & PÉREZ-PERDOMO, supra note 268, at 117-18 (explaining that civil judges both gather and weigh evidence).
294. See MERRYMAN & PÉREZ-PERDOMO, supra note 268, at 117-18 (explaining that civil judges both gather and weigh evidence).
295. For Japan, see Makoto Amino, Shohyo (Trademarks) 47 (1992); for France, see Article L713-1 available at http://195.83.177.9/upl/pdf/code_35.pdf (“Registration of a mark shall confer on its owner a right of property. . .”).
296. For France, see L. 713-1 available at http://195.83.177.9/upl/pdf/code_35.pdf (“Registration of a mark shall confer on its owner a right of property in that mark for the goods and services he has designated.”).
owner in a Civil Law country can sell the same mark to one entity to make bicycles and another entity to make cameras. The quality control of a licensee’s trademark use need not be supervised by the licensor in Civil Law countries. There is no tracking of the goodwill. Trademarks can be assigned in gross and used under naked licenses. Doctrinally, trademarks act like property in Civil Law countries.

Therefore, as property, trademarks in Civil Law countries are far harder to obtain and far more difficult of which to be divested. Although the trademark registration process itself in Civil Law jurisdictions is no more arduous than in the United States and usually much easier, the very fact is that one must obtain a trademark registration before one “owns a trademark.”

As such, there are very high transaction costs involved in Civil Law countries that are not recognized in the United States. There is a very high cost of knowledge. A user of a trademark has to, somehow, know or learn that a registration must be filed. The cost of this knowledge is not measurable. Additionally, an entity that wishes to own a trademark has to hire a lawyer (or other specialist) to file the trademark application. It would be a very unlikely occurrence for a private individual to file a trademark application pro se in Civil Law countries.

Once registered, no Civil Law trademark office has the authority to cancel trademarks on their own motion for something like nonuse. It takes a specific motion by a party who now wants to use and register the mark to obtain a cancellation for nonuse. That is, in order to obtain a registration and in order to maintain the trademark, use need never occur. Usually, three years of nonuse opens the door for aggrieved parties to file a motion to cancel a trademark for nonuse. The evidentiary burden is similarly high for other common reasons to

---

298. For Japan, see Japanese Trademark Act, art. 24(1) available at http://www.japaneselawtranslation.go.jp/law/detail/?id=45&vm=02&re=02&new=1 (“trademark right may be divided for each of the designated goods”); for France, see Law No. 92-597 of July 1, 1992, on the Intellectual Property Code (Legislative Part) as last amended by Law No. 94-102 of February 5, 1994, book 7, title 1, chapter 4, L. 714-1 available at http://195.83.177.9/upl/pdf/code_35.pdf (“The rights under a mark may be transferred in whole or in part, independently of the company that exploits them or has them exploited.”).

299. For Japan, see Kazuko Matsuo, Trademarks, in 4 Doing Business in Japan Sec. 6.01 – Sec. 6.12 (Zentaro Kitagawa ed., 2007) (entire chapter of treatise on licensing intellectual property in Japan and never mentions quality control).

300. See Kazuko Matsuo, Intellectual Property Protection, Doing Business in Japan § 6.01[d] (Zentaro Kitagawa ed., 2009) (“The prior rule that assignment of a trademark independent of the business good will concerned was not permitted was abolished, permitting freer transfer of registered trademarks.”); Makoto Amino, Shohyo [Trademarks] 647-54 (1995).


303. Id.

petition a trademark to cancel a trademark registration such as nondistinctiveness or functionality.

Because the trademark in Civil Law countries is property, it is harder to lose, harder to obtain and more easily confused. That is, in Civil Law countries, trademarks are construed much more broadly than in the United States.

The Civil Law of trademarks find their origin in providing businesses an ability to protect their “signs” as corporeal rights. The origin of the Common Law trademark is to prevent a consumer from being confused.

The Civil Law of trademark jurisprudence focuses on the functions of trademarks. When a function of the trademark is harmed, the Civil Law world calls that trademark infringement.

The first function of trademarks is the source indicating function. This is the only role that trademarks play in the United States. This function allows the trademark claimant to use the mark to indicate that it is the sole creator of the good or service to which the mark is attached.

The second function is the quality guarantee function. By using the same trademark on the same goods or services, customers can rely on the fact that the quality of the goods or services will be consistent. Not that the quality will be

---


306. See Timothy W. Blakely, Beyond the International Harmonization of Trademark Law: The Community Trade Mark as a Model of Unitary Transnational Trademark Protection, 149 U.PA. L.REV. 309, 309 (2000) (stating that traditionally, trademarks were “understood to indicate the origin of a product.”).


308. Shoen Ono, Overview of Japanese Trademark Law, Chapter 5, p. 7-8 (1992) translated and available at http://www.iip.or.jp/e/e_publication/ono/index.html (chapters are not consecutively numbered); Makoto Amino, Shohyo (Trademarks) 71-2; Ethan Horwitz, FRA-3 World Trademark Law and Practice, France § 3.01 (Mathew Bender, Rev. ed.) (In French law, “the sole condition [of trademark validity] is that it must be distinctive with respect to the articles it identifies,” citing Mathely, Letter from France, (1969)).

309. See Radiance A. Walters, Partial Forfeiture: The Best Compromise in Trademark Licensing Protocol, 91 J. PAT. & TRADEMARK OFF. SOC’Y 126, 128 (2009) (“According to the statute, the purpose of a trademark is to identify a good or service to the consuming public and ensure consistency in quality.”).

310. Shoen Ono, Overview of Japanese Trademark Law, Chapter (1999) translated and available at http://www.iip.or.jp/e/e_publication/ono/index.html (chapters are not consecutively numbered); Makoto Amino, Shohyo (Trademarks) 73-74; 2-3 Ethan Horwitz, World Trademark Law and Practice § 3.01 (Mathew Bender, Rev. Ed.) (Under German law, trademarks “must be likely to serve the consumer as a means of recognizing the goods of the applicant and to distinguish those goods from the goods of competitors. A mark cannot be registered if it is not capable of distinguishing the goods upon which it appears, if it consists solely of figures or letters, or if it consists of words which indicate the kind, time and place of production, the nature, purpose, price, quantity or weight of the goods.”) (citations omitted).
good or bad, just consistent. Therefore, a handbag purchased at JC Penny will have different quality expectation than one purchased at Coach.

The third function is the consistent good function. When a specific trademark is used, the consumer can immediately recognize that a specific good or range of goods is associated with the use of this mark. Therefore, when a consumer sees the trademark Kodak, the consumer knows that the range of goods associated with that mark will somehow include something to do with cameras. In the United States, when the producer creates a close association between the good and the trademark, the mark will be rendered generic. In Civil Law countries, this ability to broadly protect the good’s function renders the dilution doctrine unattractive.

Finally, the advertising function of trademarks is of growing import. The use of trademarks plays a very important role in advertising a good or service. Take Coach purses, for example. Just saying the word “Coach” connotes expensive and luxurious purses and handbags for women. Coach need not say

---

311. Shoen Ono, Shohyoho (Trademark Law) 49 (1994); Makoto Amino, Shohyo (Trademarks) 774-75. A Danish decision from 1983 held that confusingly similar names could be both eligible for registration as trademarks because the two companies were in totally different fields, thus there would be no confusion of one product line with another. 2-3 Ethan Horwitz, World Trademark Law and Practice, Denmark § 3.07, n. 16 (Mathew Bender, Rev. Ed.) (“A trade mark can only stop the registration of a company name if the company actually is doing business with identical or similar goods or services for which the trade mark is registered but a trade mark registration can only stop a person using the trade mark as a name if the trade mark has been put on the official list of names. The rights to names and trade marks are obviously not reciprocal, but the rights derived from names are on the other hand more narrow in scope, as identity or near-identity until now has been paramount to the ability to stop infringing use or abuse of a name or a trade mark. In this connection, it seems that a trade mark registration combined with the provisions under the law protecting names would form the best possible protection in Denmark for trade marks derived from eg a company name, a founder’s name etc. It should be considered desirable to register secondary names for registered companies or to put the registered trade mark on the official list of names as identical trade marks are then prevented from being used even for different goods or services.”) (quoting Memorex A/S v. Memotak, Aps, decision of the Supreme Court dated February 15, 1983, ULR 1983, 341, [1983] EIPR D-247.).

312. See, e.g., King-Seeley Thermos Co. v. Aladdin Industries, Inc., 321 F. 2d 577 (2d Cir. 1963) (the product named “thermos” became the word used to describe a vacuum-insulated container that would keep beverages hot or cold and thus was generic).

313. See Shoen Ono, Shohyoho (Trademark Law) 49 (1994); Makoto Amino, Shohyo (Trademarks) 74-75. Argentina has struggled repeatedly with the importance of advertising in trademark law. Aracama Zorraquin describes the evolution of the law this way: “Those who are familiar with the evolution of Argentine case law concerning the protection of advertising phrases (slogans) will remember that for a time such phrases were accepted as being registrable under trademark law. A subsequent development brought about the prohibition of their registration as such, whereupon case law, after reverting to the original position, progressed towards a situation where phrases of advertising character had to be protected by Law No. 11, 723 on Copyright, whereas those that possessed distinctive character were to be protected by the Law on Trademarks. This distinction, which is highly subjective and very difficult to make in practice, has been removed by the new Law, which allows registration as trademarks of phrases for advertising purposes that meet the requirements of registrability specified by the Law on Trademarks.” 1-3 Ethan Horwitz, World Trademark Law and Practice, Argentina § 3.01, n. 11 (Mathew Bender, Rev. Ed.) (citing Ernesto Aracama Zorraquin, New Legislation in Argentina, 1982 Ind. Prop. 88, 93).
anything about their product. Just saying the word is enough. Even lesser known, less expensive brands of purses utilize the advertising function of trademarks. As producers attempt to raise brand awareness and make sales, they focus on the use of their trademark to operate as a short cut to their reputation, whether they are a high-end producer or not.

Civil Law countries protect these functions of trademarks. Harming any one of these functions results in a claim of trademark infringement. That is, the trademark right in Civil Law countries is far broader than in the United States.

The protection of nontraditional trademarks in a Civil Law system is problematic because of the need to be able to depict marks graphically. In the United States, a verbal description of a scent mark, for example, sufficed; however, in the CTM and before German judges, a chemical formula of a scent was deemed insufficient. Under TRIPS, “members may require, as a condition of registration, that signs be visually perceptible.” Civil Law systems and the CTM take this rather literally.

However, the graphic representation requirement is satisfied by a “stave divided into measures and showing, in particular, a clef, musical notes and rests whose form indicates the relative value and, where necessary accidentals.” Color may satisfy the graphic representation requirement if an “internationally recognized identification code” is used.

B. Common Law (United States)

Essentially, each element of the mini-primer on the Civil Law of trademarks from above is answered in the diametric opposite when in the United States. All


319. As the United States is the last true believer of the common law jurisdictions and the largest economy in the world, it makes a perfect contrast to Civil Law countries. Most other common law countries such a Great Britain or Canada have already moved their trademark systems closer to Civil Law countries. Australia may be an exception to this. In Australia, although nontraditional trademarks like color are registered and enforced, they remain a “relatively rare beast.” Campbell Thompson & Bill Ladas, How Green Is My Trade Mark? Woolworths v. BP, 29 EUR. INTELL. PROP. REV. 29, 29 (2007).
It is this use that makes the federal regulation of trademarks constitutional in the United States. Trademarks are not writings. Therefore, the Patent and Copyright Clause of the United States Constitution does not apply. Trademarks may be regulated by Congress under the Commerce Clause. For there to be commerce which Congress can regulate, there must be use in interstate commerce. Although there are advantages to obtaining a federal registration in the United States, the trademark right itself subsists upon use of a mark, not on registration.

320. In re Trade-Mark Cases, 100 U.S. 82, 94 (1879); United Drug Co. v. Theodore Rectanus Co., 248 U.S. 90, 97-100 (1918).
321. Jerome Gilson, Karin Green, & Anne Gilson LaLonde, 1-3 GILSON ON TRADEMARKS § 3.02 (Matthew Bender); 15 U.S.C. § 1052(d). Proctor and Gamble Company v. Johnson and Johnson, Inc. 485 F. Supp. 1185 (S.D.N.Y. 1980) aff’d at 636 F.2d 1203 (2d Cir. 1980); In re Compagnie Generale Maritime, 993 F.2d 841, 854 (CAFC 1993) (“The ‘use in commerce’ requirement is the Constitutional basis for enactment of the Lanham Act, . . .) (Judge Nies, dissenting); Buti v. Perosa, S.R.L., 139 F.3d 98, 102 (2d Cir. 1998) cert. denied, 525 U.S. 826, 119 S. Ct. 73, 142 L. Ed. 2d 57 (1998)(“use in commerce” requirement of the Lanham Act is simply a jurisdictional predicate to any law passed by Congress under the Commerce Clause. The “history and text of the Lanham Act show that ‘use in commerce’ reflects Congress’s intent to legislate to the limits of its authority under the Commerce Clause”)(quoting United We Stand America, Inc. v. United We Stand, America New York, Inc., 128 F.3d 86, 92 (2d Cir. 1997)).
322. See In re Trade-Mark Cases, 100 U.S. 82, 94 (1879) (discussing difference in a writing and a trademark).
323. U.S. CONST. art. I, § 8, cl. 3.
324. 1) Notice: Perhaps one of the most significant reasons to register a trademark regards notice of rights. Section 22 (15 U.S.C. §1072) provides for constructive notice of claimed rights under the Lanham Act. That is, once a trademark is registered, all would-be infringers are charged with knowledge of the registration and the benefits derived thereunder. In litigation, this could have a significant impact on the outcome of any given case. See Melissa E. Roth, Something Old, Something New, Something Borrowed, Something Blue: A New Tradition In Nontraditional Trademark Registrations, 27 CARDOZO L. REV. 457, 469-471 (2005) (registration provides constructive notice). 2) Deterrence: Closely related to the notice function, trademark registration owners can claim use of the ® symbol on goods or services. Although not precisely documented as such, it is thought that use of the ® symbol provides some deterrence to infringers. See Philip J. Greene, Trademark Counsel in the Federal Government -- A Practitioner’s Perspective, 86 J. PAT. & TRADEMARK OFF. SOC’Y 361, 375 (2004) (“the (R) symbol adjacent to the mark will help dispel the notion that government marks are in the public domain”). 3) National protection: Once a trademark is registered, the owner of that registration can claim national priority. If this registration is based on an intent to use a mark, that priority date reverts back to the actual application date of the registration. See Laurinda L. Hicks & James R. Holbein, Convergence of National Intellectual Property Norms in International Trading Agreements, 12 AM. INT’L L. & POL’Y 769, 777 (“Although federal registration grants substantive and procedural rights to the trademark owner, the absence of such federal registration does not place the mark in the public domain.”). 4) Incontestability: Once a registered trademark is used for five consecutive years, the owner of that registration can claim the registration has become “incontestable” under Section 15 of the Lanham Act. As will be seen, this, too, can provide many positive consequences for the plaintiff/trademark registration owner. The stakes over incontestability went up remarkably in 2009 when the Eighth Circuit held that when a mark became incontestable, it so changed the circumstances that collateral estoppel did not prevent the plaintiff from suing the defendant a second time on the same operative facts. B&B Hardware, Inc. v. Hargis Indus., 569 F.3d 383, 388 (8th Cir. 2009). 5) Cybersquatting: It is far easier to establish a bona fide commercial interest in a registered trademark for purposes of
As we adequately demonstrated in the *Qualitex* decision, it does not need an act of Congress to expand the notion of trademark protection to nontraditional trademarks.\(^\text{326}\) In that case, the Supreme Court held that “any” appellation of source could be claimed as a trademark as long as there was secondary meaning and that there ought to be no per se exclusion of any trademark subject matter.\(^\text{327}\) Therefore, as was stated earlier, after the *Qualitex* case, trademark applications for nontraditional trademarks in the United States exploded.\(^\text{328}\) Congress did not amend the Lanham Act. There was one judicial holding by the Supreme Court and the entire scenario for nontraditional trademarks changed.\(^\text{329}\)

In the United States there may be a litmus test used to attempt to exclude activist judges where judges must swear to the “application of law” god and forgo the “judges making law” devil,\(^\text{330}\) but there is plenty of evidence that United States judges have no choice but to make law.\(^\text{331}\) As in the *Qualitex* case,
the Supreme Court may not have made a law allowing for nontraditional trademarks; however, if only 54 nontraditional trademarks were filed in nearly 50 years before the Qualitex decision and 352 were filed in the 15 years after that case, it is hard to argue that the result of the Qualitex case was anything other than making law. Applying the old statute in a new way made law. That is all that happened in the Qualitex case.

Of course, in the United States, there is general pleading for lawsuits, and discovery is liberal and there are jury trials. Today, roughly one half of all trademark trials are before juries. In addition, surveys are greatly used and relied upon. In some district courts, if a litigant does not proffer a survey it is presumed that the survey would not support the proposition that is being argued. Surveys in the United States have become a more precise science. These are not general opinion surveys, which are not worth much. Good trademark surveys establish foundation, establish that the respondent is knowledgeable about the questions asked and can be submitted into evidence as reliable.

The trademark in the United States is never owned. In the United States, only the right to exclude others to the extent the mark is used and as long as the mark is used would be considered a property right. A property right of exclusion

---

336. World Wrestling Fed’n. Entm’t, Inc. v. Big Dog Holdings, Inc., 280 F. Supp. 2d 413 (WD Penn. 2003) (“The Court is aware that WWE is not legally required to conduct a confusion survey. Such a failure, particularly when the trademark owner is financially able, justifies an inference ‘that the plaintiff believes the results of the survey will be unfavorable.’”) (quoting Charles Jacquin et Cie, Inc. v. Destileria Serralles, Inc., 921 F.2d 467, 475 (3d Cir. 1990)).
and a property right to a corporeal thing are very different things. In the United States the trademark right is a right of exclusion; in Civil Law nations a trademark is a corporeal thing that is owned.

The United States system of trademark protection is based on fair competition.\textsuperscript{339} Trademark law is one subset of unfair competition law. It has its roots in the English tort of deceit where the deceived consumer had standing to sue, not an aggrieved trademark holder.\textsuperscript{340} The trademark laws in the United States are founded on the notion that the consumer has a right to be free of deception.\textsuperscript{341}

In the United States today, there are three competing interests that need to be simultaneously balanced. The first is this consumer’s right to be free of deception.\textsuperscript{342} However, if the system over-protects the consumer, trademark holders may use trademarks less and thereby society would not receive that benefit of trademark use. Consumers’ search costs would go up if trademark holders used trademarks less often and consumers would pay more for their goods and services.\textsuperscript{343}

Therefore, the rights of trademark holders also must be balanced in this equation. Trademark holders must receive adequate incentive to continue to use trademarks.\textsuperscript{344} Trademark holders need to feel that their trademarks are appropriately protected and enforced. Trademark holders need to continue to invest in trademarks so that society’s search costs are lessened and society can take advantage of the short-cut function that trademarks play.\textsuperscript{345}

As a subset of unfair competition law, trademark law and policy also must include the rights of third parties and their ability to compete.\textsuperscript{346} If trademark rights are too strong in the United States, third parties will not be able to compete. Third parties’ costs of market access would go up if they have to defend too many law suits or if they have to work too hard to clear trademarks before using the marks.\textsuperscript{347} Additionally, third parties would suffer from too

\begin{itemize}
\item \textsuperscript{340} J. Thomas McCarthy, 1 McCarthy on Trademarks and Unfair Competition § 5:2 (1996 4th ed.).
\item \textsuperscript{341} Id.
\item \textsuperscript{342} BellSouth Advertising & Publishing Corp. v. Real Color Pages, Inc., 792 F. Supp. 775, 780-81 (M.D. Fla. 1991).
\item \textsuperscript{344} Id.
\item \textsuperscript{345} Id.
\item \textsuperscript{346} See Kenneth L. Port, The Congressional Expansion of American Trademark Law: A Civil Law System in the Making, 35 WAKE FOREST L. REV. 827, 831 (2000) (“The economic rationale for trademark protection is undermined [by expansion of rights] because the expansion actually increases total transaction costs rather than being, at least, cost neutral . . . [T]he expansion undercuts the legal rationale for trademark protection because it tips the scales in favor of protecting the goodwill of the trademark owner rather than balancing the interests of the consumer to be free from confusion and third parties’ right to compete.”).
\item \textsuperscript{347} Kenneth L. Port, Trademark Extortion: The End of Trademark Law, 65 WASH. & LEE L.
strong protection of trademarks because the range of available marks that third parties might use would shrink.

All three of these interests need to be balanced. Historically, the United States has been focused on the equality of each of these three pillars of trademark protection. The data on reported trademark cases seems to bear out this fact. In the last 60 years, trademark claimants have prevailed in their claims of infringement 50% of the time. This rate is shrinking. Today, fewer trademark claimants prevail in their claims of infringement than they did just a decade ago.

Therefore, there is a critical gap in the theoretical justification of protecting trademarks in Civil Law countries compared to the United States. This theoretical difference makes harmonization of any sort difficult. Even if the doctrine is harmonized, the underlying theory will remain diverse. This distinction is why trademark harmonization has not been successful to date. In order to succeed, the doctrine and the theory must be harmonized.

XIII. THE NORMATIVE CASE AGAINST NONTRADITIONAL MARKS

Normatively, with few exceptions, nontraditional trademarks do not satisfy the expectations we have for trademarks. Even doctrinally, there is little justification for the protection of nontraditional trademarks. Nontraditional trademarks are functional. Nontraditional trademarks do not act as indicators of source. Nontraditional trademarks are byzantine elements of a good or service. They do not indicate that the good or service comes from a consistent source. Even though nontraditional trademarks have recognized a huge spike post-
Qualitex, nontraditional trademarks are very rarely found valid and infringed by Article III style court.

Jacobson, the losing party in Qualitex v. Jacobson, actually provided the most the perfect objections to all of nontraditional trademarks, not just color. These are addressed below.

A. Sense Depletion

Jacobson argued that to protect a color would ultimately limit the number of colors available for a competitor to use. When applying this notion to nontraditional marks, we might call it sense depletion.

351. Id. at 168.
When the human brain can only remember sixteen scents for a long period of time, allowing anyone to register a smell seems to work great hardship, a great depletion, of that sense. Although there are many sounds available in the world, allowing anyone to exclude others from the use of sound would inappropriately drive up transaction costs without the related benefit to society of the trademark use. That is, making a product sound attractive is decoration. Decoration is considered to be aesthetically functional and therefore not protected.

It seems that the sense depletion theory is very important in field of nontraditional trademarks. Allowing them to be protected depletes the field and does not result in the expected reduction in transaction costs.

B. Sense Confusion

With sense confusion, we are worried about articulating a standard by which a use could be judged as infringing or not infringing. All Circuits have come up with tests to be applied to traditional marks. This is known as the likelihood of confusion analysis.

With nontraditional marks, what will be the test? Simple variations of the nontraditional trademark might result in big differences in perception because nontraditional trademarks rely on senses that are arbitrary and not well understood. These senses are not consistent between people. They, in fact, are inherently unreliable.

With a traditional mark, it can be projected to a jury. The jury can look at the alleged infringing mark, such as KOKE and compare it to the word COKE. The jury can think about whether KOKE infringes COKE. It does not have to agree with what COKE is, how it should be perceived, or how it is to be known. Those are givens in traditional trademarks.

With nontraditional marks, the jury has to first decide what the mark is before it can go on to decide if the mark has been infringed. When nontraditional trademarks are invoked, reasonable jurors would think differently of the nontraditional trademark. No two jurors would ever agree on the scent of a mark or the motion of a mark. These things are subjective and variable. When

---

352. See Churovich, supra n. 187, at 302.
354. W.T. Rogers Co. v. Keene, 778 F.2d 334, 339-40 (7th Cir. 1985) (commercial need precludes trademark protection); see also Jerre B. Swann, The Design of Restaurant Interiors - A New Approach to Aesthetic Functionality, 76 TRADEMARK REP. 408, 408 (1986) (“Primarily attractive features serve ‘other than a trade-mark purpose.’”); 1-2 GILSON ON TRADEMARKS § 2.11 (“…consumers may not perceive certain nontraditional marks as trademarks at all. They may see them as merely decorative, as an inherent part of the product or as an attempt to amuse, rather than as indicating the source of the goods”).
355. 5-5 GILSON ON TRADEMARKS § 5.02.
356. 1-1 GILSON ON TRADEMARKS § 1.04.
357. See, e.g., supra Parts VI-VII.
a scent mark is not perceived the same between two people or even by one person given weather patterns, etc. these marks seem too indeterminant to warrant trademark protection. There would be far too much sense confusion to make sense of these trademarks.

C. Ability to Act as a Trademark

The biggest worry with nontraditional trademarks is whether they are actually acting as a trademark at all. Consider, for example, the Yamaha personal watercraft spray. Is this so-called trademark acting to identify Yamaha personal watercrafts or is it a nifty product feature that people buy because of the product feature, not because they are relying on it to identify source. Yet again, we bump into the notion of aesthetic functionality.

With many nontraditional marks, one does not have to enter the murky lands of aesthetic functionality to oppose a trademark. Most nontraditional trademarks do not pass the functionality test.359 Most nontraditional marks are initially refused registration because of this doctrine. Nontraditional trademarks play a role in dictating the function of a product. Many of these relate to human senses that we do not clearly understand. One court denied trademark protection for the color of boat engines.360 That court held that black motors worked to make an object look smaller and therefore was playing a function of the product.361

If the nebulous notion of a color contributing to our size perception of the product is worthy of analysis, I suggest that we do not understand the role nontraditional trademarks is playing in our purchasing decisions. Do we buy the black engine because we think engine should be black or because it identifies source? Most likely, we have an inconsistent image in our minds about what color things should be or how they should smell, etc. When Mrs. Clark makes her skunks smell like plumeria,362 we are not only relieved, we like it. We do not buy the skunks because the smell indicated source. Plumeria is a very fine scent. We buy the skunks because we like the smell of plumeria, not because it indicates source.

As such, nontraditional trademarks do not play the role of indicating source and they cannot. Nontraditional trademarks merely make a product more attractive. The product is purchased by buyers not because of the source denoting capacity of the nontraditional mark but rather because of an extremely inconsistent, subjective notion of what is “good” or “attractive.”

361. Id. at 1529.
362. See supra notes 180-86 and accompanying text.
D. Is it Worth it?

All this raises a very important question: is it worth it? I submit that it is not. Nontraditional trademarks are much harder to get than traditional marks. On average, nontraditional marks take more than 30 months to be registered.\textsuperscript{363} Traditional marks take far less time.\textsuperscript{364} Nearly half of all nontraditional trademarks applied for registration no longer exist.\textsuperscript{365} The applications have gone abandoned or the registrations have been cancelled for nonuse. Courts do not enforce nontraditional trademarks.

Therefore, although obtaining a registration for a nontraditional mark is possible in the United States, it is not worth it. It is not worth it from the trademark claimants’ viewpoint as it takes so long and most are lost or abandoned. It is not worth it from society’s point of view as allowing this protection does not reduce transaction costs. Society gets nothing in return for allowing the protection of nontraditional trademarks. Nontraditional marks do not play the same role of reducing transaction costs so that our products cost less.\textsuperscript{366} Nontraditional trademarks merely adorn and beautify our products. Although I’m all in favor of pretty goods, pretty is not a trademark. As such, trademark protection for nontraditional marks is not worth the investment.

XIV. CONCLUSION

There must be a reason why trademark harmonization has been so difficult to realize. The world knows that a lack of harmonization contributes, in significant ways, to the cost of doing business internationally. Harmonizing trademark laws would be good for all. Yet, it has remained elusive.

The reason trademark doctrine has not been harmonized is because the underlying theory of trademark protection between the Civil Law world on one hand and the Common Law world, primarily the United States, on the other hand are so disparate.

The protection of nontraditional trademarks in the United States is an extreme manifestation of this divergence. With one holding from one Supreme Court decision, nontraditional trademark applications for registration increased exponentially. That decision merely said that there were no per se restrictions on the subject matter of trademarks. “Any” meant any.

Civil Law jurisdictions do not have that flexibility. As positivists, they must wait for their legislatures to enumerate the specific scope of trademark protection. Their Examiners have very little experience in examining nontraditional trademarks, to the extent applications are received.

\begin{itemize}
\item \textsuperscript{363} See infra app. D.
\item \textsuperscript{364} See supra notes 159-162 and accompanying text.
\item \textsuperscript{365} See infra app. C.
\item \textsuperscript{366} See supra notes 54-56 and accompanying text.
\end{itemize}
Therefore, until the theoretical orientation of trademark policy between the Civil Law and the United States is harmonized, trademark harmonization efforts will remain elusive. The United States exacerbates this issue by protecting “any” indicator source. Civil Law countries make this exacerbation worse by attempting to harmonize to the standard of protecting nontraditional trademarks, meanwhile adhering to the Civil Law notion of trademark protection. That is, this doctrinal harmonization will not be meaningful because it affects so few trademarks and the trademark jurisprudence remains disparate. As such, transaction costs incurred because of disparate trademark systems around the world will not be reduced. This is truly harmonization that simply does not matter.

The protection of nontraditional trademarks is further rendered suspect in the United States by the fact that there has been such a spike in trademark applications for nontraditional trademarks post-\textit{Qualitex}, but the rate of litigation has remained nearly static. That is, trademark owners are interested in applying to register nontraditional marks, but they are not interested in enforcing them. This raises the question of why? If trademark holders are interested in registering nontraditional marks, why are they not enforcing them? When they do, they only succeed 1/3 of the time.

If, for example, Japan harmonizes to the American standard of protecting nontraditional trademarks, Japan will be making several mistakes. The first mistake is to think that nontraditional trademarks add up to something in the United States. They do not. Even though \textit{Qualitex} held that color alone could be protected and an explosion of interest developed regarding nontraditional trademarks, the numbers still do not add up to much. The second mistake is to believe that other countries are engaged in the protection of nontraditional trademarks based on the survey from WIPO. They are not. Even though the Singapore Treaty explicitly addresses nontraditional marks, they remain a very rare entity in Civil Law systems. Finally, if Japan merely adopts a nontraditional trademark system without thought to the underlying theory of trademark protection, the harmonization will be empty.

These mistakes work to prove the point: real trademark harmonization has a long way to go. We should not get distracted by meaningless claims like protecting (or not) Frank Ogden’s DNA sequence.
### APPENDIX A

#### Percentage Nontraditional Marks Actually Registered

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Sound Mark Applications</th>
<th>Total Registered Sound Marks</th>
<th>Percentage Total Sound Marks Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound</td>
<td>336</td>
<td>171</td>
<td>50.89%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Scent Mark Applications</th>
<th>Total Registered Scent Marks</th>
<th>Percentage Total Scent Marks Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scent</td>
<td>28</td>
<td>12</td>
<td>42.86%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Color Mark Applications</th>
<th>Total Registered Color Marks</th>
<th>Percentage Total Color Marks Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>345</td>
<td>242</td>
<td>70.14%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Taste Mark Applications</th>
<th>Total Registered Taste Marks</th>
<th>Percentage Total Taste Marks Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taste</td>
<td>9</td>
<td>0</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Tactile Mark Applications</th>
<th>Total Registered Tactile Marks</th>
<th>Percentage Total Tactile Marks Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tactile</td>
<td>5</td>
<td>1</td>
<td>20.00%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total No Description Mark Applications</th>
<th>Total Registered No Description Marks</th>
<th>Percentage Total No Description Marks Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Description</td>
<td>16</td>
<td>1</td>
<td>6.25%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Other Mark Applications</th>
<th>Total Registered Other Marks</th>
<th>Percentage Total Other Marks Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other</td>
<td>5</td>
<td>2</td>
<td>40.00%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Motion Applications</th>
<th>Total Registered Motion Marks</th>
<th>Percentage Total Motion Marks Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motion</td>
<td>18</td>
<td>13</td>
<td>72.22%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Hologram Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hologram</td>
<td>19</td>
</tr>
<tr>
<td>Hologram</td>
<td>Total Registered Hologram Marks</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td></td>
<td>Percentage Total Hologram Marks Registered</td>
</tr>
<tr>
<td>Total Nontraditional Mark Applications (incl. pending)</td>
<td>781</td>
</tr>
<tr>
<td>TOTAL</td>
<td>Total Registered Nontraditional Marks</td>
</tr>
<tr>
<td></td>
<td>Percentage Total Nontraditional Marks Registered</td>
</tr>
</tbody>
</table>
## APPENDIX B

### Percentage Pending Nontraditional Mark Applications

<table>
<thead>
<tr>
<th>mark</th>
<th>total applications</th>
<th>total pending</th>
<th>percentage pending</th>
</tr>
</thead>
<tbody>
<tr>
<td>sound</td>
<td>336</td>
<td>52</td>
<td>15.48%</td>
</tr>
<tr>
<td>scent</td>
<td>28</td>
<td>6</td>
<td>21.43%</td>
</tr>
<tr>
<td>color</td>
<td>345</td>
<td>1</td>
<td>0.29%</td>
</tr>
<tr>
<td>taste</td>
<td>9</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>tactile</td>
<td>5</td>
<td>1</td>
<td>20.00%</td>
</tr>
<tr>
<td>no description</td>
<td>16</td>
<td>2</td>
<td>12.50%</td>
</tr>
<tr>
<td>other</td>
<td>5</td>
<td>2</td>
<td>40.00%</td>
</tr>
<tr>
<td>motion</td>
<td>18</td>
<td>2</td>
<td>11.11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td><strong>Hologram</strong></td>
<td><strong>Total Hologram Applications</strong></td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Total Pending Hologram Marks</strong></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Percentage Total Hologram Marks Pending</strong></td>
<td>15.79%</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>Total Nontraditional Mark Applications</strong></td>
<td>781</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Total Pending Nontraditional Marks</strong></td>
<td>69</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Percentage Total Nontraditional Marks Pending</strong></td>
<td>8.83%</td>
<td></td>
</tr>
</tbody>
</table>
### APPENDIX C

<table>
<thead>
<tr>
<th></th>
<th>Percentage Abandoned Nontraditional Marks</th>
<th>Percentage Cancelled Nontraditional Marks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sound Mark Applications</td>
<td>336</td>
<td>Total Sound Mark Applications 336</td>
</tr>
<tr>
<td>Total Abandoned Sound Marks</td>
<td>112</td>
<td>Total Cancelled Sound Marks 28</td>
</tr>
<tr>
<td>Percentage Total Sound Marks Abandoned</td>
<td>33.33%</td>
<td>Percentage Total Sound Marks Cancelled 8.33%</td>
</tr>
<tr>
<td>Total Scent Mark Applications</td>
<td>28</td>
<td>Total Scent Mark Applications 28</td>
</tr>
<tr>
<td>Total Abandoned Scent Marks</td>
<td>7</td>
<td>Total Cancelled Scent Marks 2</td>
</tr>
<tr>
<td>Percentage Total Scent Marks Abandoned</td>
<td>25.00%</td>
<td>Percentage Total Scent Marks Cancelled 7.14%</td>
</tr>
<tr>
<td>Total Color Mark Applications</td>
<td>345</td>
<td>Total Color Mark Applications 345</td>
</tr>
<tr>
<td>Total Abandoned Color Marks</td>
<td>102</td>
<td>Total Cancelled Color Marks 50</td>
</tr>
<tr>
<td>Percentage Total Color Marks Abandoned</td>
<td>29.57%</td>
<td>Percentage Total Color Marks Cancelled 14.49%</td>
</tr>
<tr>
<td>Total Taste Mark Applications</td>
<td>9</td>
<td>Total Taste Mark Applications 9</td>
</tr>
<tr>
<td>Total Taste</td>
<td>9</td>
<td>Total Cancelled Taste Marks 0</td>
</tr>
<tr>
<td>Abandoned Taste Marks</td>
<td>Percentage Total Taste Marks Abandoned</td>
<td>100.00%</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Total Tactile Mark Applications</td>
<td>5</td>
<td>Total Tactile Mark Applications</td>
</tr>
<tr>
<td>Total Abandoned Tactile Marks</td>
<td>3</td>
<td>Tactile</td>
</tr>
<tr>
<td>Percentage Total Tactile Marks Abandoned</td>
<td>60.00%</td>
<td>Percentage Total Tactile Marks Cancelled</td>
</tr>
<tr>
<td>Total No Description Mark Applications</td>
<td>16</td>
<td>Total No Description Mark Applications</td>
</tr>
<tr>
<td>Total Abandoned No Description Marks</td>
<td>11</td>
<td>No Description</td>
</tr>
<tr>
<td>Percentage Total No Description Marks Abandoned</td>
<td>68.75%</td>
<td>Percentage Total No Description Marks Cancelled</td>
</tr>
<tr>
<td>Total Other Mark Applications</td>
<td>5</td>
<td>Total Other Mark Applications</td>
</tr>
<tr>
<td>Total Abandoned Other Marks</td>
<td>1</td>
<td>Other</td>
</tr>
<tr>
<td>Percentage Total Other Marks Abandoned</td>
<td>20.00%</td>
<td>Percentage Total Other Marks Cancelled</td>
</tr>
<tr>
<td></td>
<td>Total Motion Applications</td>
<td>18</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Total Abandoned Motion Marks</td>
<td>2</td>
<td>Motion</td>
</tr>
<tr>
<td>Percentage Total Motion Marks Abandoned</td>
<td>11.11%</td>
<td>Total Cancelled Motion Marks</td>
</tr>
<tr>
<td>Total Hologram Applications</td>
<td>19</td>
<td>Total Hologram Applications</td>
</tr>
<tr>
<td>Total Abandoned Hologram Marks</td>
<td>5</td>
<td>Hologram</td>
</tr>
<tr>
<td>Percentage Total Hologram Marks Abandoned</td>
<td>26.32%</td>
<td>Percentage Total Hologram Marks Cancelled</td>
</tr>
<tr>
<td>Total Nontraditional Mark Applications</td>
<td>781</td>
<td>Total Nontraditional Mark Applications</td>
</tr>
<tr>
<td>Total Abandoned Nontraditional Marks</td>
<td>252</td>
<td>TOTAL</td>
</tr>
<tr>
<td>Percentage Total Nontraditional Marks Abandoned</td>
<td>32.27%</td>
<td>Percentage Total Nontraditional Marks Cancelled</td>
</tr>
</tbody>
</table>
### APPENDIX D

<table>
<thead>
<tr>
<th></th>
<th>Average Time Between Application Filing and Registration (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sound:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>Scent:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>Color:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>Taste:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>Tactile:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>No Description:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>Other:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>Motion:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>Hologram:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
<tr>
<td><strong>Total Nontraditional Marks:</strong></td>
<td>Average Time Between Application Filing and Registration</td>
</tr>
</tbody>
</table>
### APPENDIX E

<table>
<thead>
<tr>
<th>Mark Type</th>
<th>Total Post-Qualitex Mark Applications</th>
<th>Percentage Post-Qualitex Mark Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sound</strong></td>
<td>295</td>
<td>87.80%</td>
</tr>
<tr>
<td>Total Sound Mark Applications</td>
<td>336</td>
<td></td>
</tr>
<tr>
<td><strong>Scent</strong></td>
<td>27</td>
<td>96.43%</td>
</tr>
<tr>
<td>Total Scent Mark Applications</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td><strong>Color</strong></td>
<td>302</td>
<td>87.54%</td>
</tr>
<tr>
<td>Total Color Mark Applications</td>
<td>345</td>
<td></td>
</tr>
<tr>
<td><strong>Taste</strong></td>
<td>9</td>
<td>100.00%</td>
</tr>
<tr>
<td>Total Taste Mark Applications</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td><strong>Tactile</strong></td>
<td>5</td>
<td>100.00%</td>
</tr>
<tr>
<td>Total Tactile Mark Applications</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td><strong>No Description</strong></td>
<td>9</td>
<td>56.25%</td>
</tr>
<tr>
<td>Total No Description Mark Applications</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td>5</td>
<td>100.00%</td>
</tr>
<tr>
<td>Total Other Mark Applications</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td><strong>Motion</strong></td>
<td>17</td>
<td>94.44%</td>
</tr>
<tr>
<td>Total Motion Mark Applications</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Total Hologram Mark Applications</strong></td>
<td>19</td>
<td></td>
</tr>
<tr>
<td><strong>Total Post-Qualitex Hologram Mark Applications</strong></td>
<td>19</td>
<td></td>
</tr>
<tr>
<td><strong>Percentage Post-Qualitex Hologram Mark Applications</strong></td>
<td>100.00%</td>
<td></td>
</tr>
<tr>
<td><strong>Total Nontraditional Mark Applications</strong></td>
<td>781</td>
<td></td>
</tr>
<tr>
<td><strong>Total Post-Qualitex Nontraditional Mark Applications</strong></td>
<td>688</td>
<td></td>
</tr>
<tr>
<td><strong>Percentage Total Post-Qualitex Nontraditional Mark Applications</strong></td>
<td>88.09%</td>
<td></td>
</tr>
</tbody>
</table>
I. INTRODUCTION

In a negligence action, once the plaintiff proves the first two elements of the cause of action, duty and breach, causation presents the next hurdle. If the harm would not have occurred “but for” the unreasonable conduct, then cause-in-fact exists. Finally, if actual harm occurs as a result of the unreasonable conduct, proximate cause determines whether liability exists.

In 1923, Leon Green published a law review article suggesting that unreasonable conduct and proximate cause were indistinguishable. Green’s
position was based on the fact that analysis of both unreasonable conduct and proximate cause was based on foreseeability. He concluded that scholarship in the field of negligence has been unable to resolve this dilemma because there is no analytically sound way to distinguish between the two. Despite this dilemma, many struggle in vain to distinguish between the appropriate foreseeability tests to apply in determining unreasonable conduct and proximate cause, generally under the presumption that the two must be distinguishable because they are separate elements of negligence. Furthermore, unreasonable conduct involves a question of law, and proximate cause involves a question of fact. As a result, post-Green scholarship and subsequent case law are replete with illusory terms that attempt to make the distinction, such as general foreseeability, specific foreseeability, forward-looking foreseeability, backward-looking foreseeability, foreseeability as it should apply in duty, foreseeability as it should apply in breach, and foreseeability as it should apply in proximate cause.

This article contends that foreseeability is an immutable concept that cannot change in shape or meaning as it is applied across the spectrum of the elements of negligence. This contention confronts an issue that post-Green scholarship has largely avoided and proposes a modified negligence liability

consequences and cannot be recovered for. In other words the “probability,” “foreseeability” or “anticipation” test is seemingly used both in determining the existence of negligence, and in determining for what consequences of such negligence a recovery may be had. It might be more accurate to say that our courts have apparently treated the existence of negligence and causal relation as the same problem, to be solved by the same formula.

Id. (emphasis in original).

5. Id.

6. Daniel Weddle of UMKC and I, in a rather cathartic manner, both admitted over beers and bratwurst at a conference in Madison, Wisconsin that we suspected this was true, but could never previously admit it because we had to teach it as if there was a difference.

7. See Schwartz et al., supra note 1, at 132 (noting that the negligence formula is useful in segregating and drawing attention to the elements of the cause of action, but that it is misleading to view the elements as discrete because none of the elements can really be defined except by reference to the others).

8. See Mark A Esquada, Note, Tort Law--Legal Malpractice--In a Suit Claiming Attorney Negligence in Handling an Appeal, the Question of Whether the Appeal Would Have Prevailed but for the Alleged Attorney Malpractice is an Issue of Law for the Court Rather Than One of Fact for the Jury. Charles Reinhart Co. v. Winienko, 513 N.W.2d 773 (Mich. 1994), 74 U. Det. Mercy L. Rev. 151, 155 (1996) (explaining that “the vast majority of cases have held that proximate cause is a question of fact for the jury”). See, e.g., Riedel v. Sheraton Bal Harbour Ass’n, 806 So. 2d 530, 532 (Fla. Dist. Ct. App. 2001) (noting that “duty” is a question of law for the court).

9. See infra Part V.

10. See W. Jonathan Cardi, Reconstructing Foreseeability, 46 B.C. L. Rev. 921, 924 (2005) [hereinafter Cardi I] (explaining the issue of whether foreseeability should be decided by the judge or jury is “a matter that courts are reluctant to discuss and that scholars, at least since Leon Green, have largely ignored”).
regime that abandons the accepted fiction that foreseeability takes on different meanings in the context of the different elements of negligence.  

This article begins with an examination of the development of negligence as a liability regime, focusing on whether unreasonable conduct and proximate cause should have been considered separate elements of the tort of negligence. It continues with an assessment of the impact of the *Palsgraf* decision on negligence formulation and suggests that any emphasis on specifically stated duties further exacerbates the overlap between unreasonable conduct (duty and breach) and proximate cause. The article then examines the failure of post-Green scholarship and judiciary to resolve this overlapping foreseeability problem. Included in this examination are excerpts of pretrial documents and transcripts of arguments from an actual negligence case demonstrating that the existing scholarship’s normative bent fails to adequately consider the actual dynamics of the negligence cause of action. The article then turns to the Restatement (Third) of Torts: Liability for Physical and Emotional Harm (“Restatement (Third)”), scrutinizing its reformulation of negligence and its failure to resolve the problem of overlapping foreseeability. Finally, the article proposes a bifurcated

---

negligence analysis as a solution to the overlapping foreseeability problem,\(^{18}\) which retains foreseeability as an essential element of the tort but does so in an intellectually defensible and empirically supportable manner.

**II. HISTORICAL DIFFERENCES BETWEEN UNREASONABLE CONDUCT AND PROXIMATE CAUSE**

Green’s premise is that unreasonable conduct and proximate cause are identical, but are usually masked as separate elements of the negligence tort.\(^{19}\) Because the existence of a duty is a significant component of determining unreasonable conduct, Green implies that there is no difference between duty and proximate cause.\(^{20}\) Therefore, discussion of the issue will begin with a review of the historical development of duty and proximate cause to determine whether they have ever been distinguishable.

**A. The Development of the Duty to Behave Reasonably**

Originally, the English courts developed a system of pleading based on writs or forms of pleading.\(^{21}\) In particular, the writs of trespass and trespass on the case were relevant to the development of tort law and negligence.\(^{22}\) Trespass writs concerned “punishment and amends for past transgression.”\(^{23}\) Upon issuance of this writ, the sheriff would summon the plaintiff to establish the defendant’s wrongdoing.\(^{24}\) In the early thirteenth century, it was settled that the Chancery clerks would issue trespass writs only for “trespasses done with force and arms against the king’s peace.”\(^{25}\) Furthermore, it was required that the plaintiff state this on the writ to gain access to the royal courts.\(^{26}\)

By 1278, no writs of trespass were issued unless the damage done or goods taken were worth at least forty shillings.\(^{27}\) There was also a requirement that the injury be done with force, which was really nothing more than a jurisdictional restriction on the royal courts.\(^{28}\) By the fourteenth century, it became apparent that there were remedies at common law worth more than

---

18. See infra Part VII.
19. See Green I, supra note 4, at 243-44.
20. Id.
22. See Kelley I, supra note 11, at 61 (explaining that “what was originally trespass on the case for negligence of default became modern negligence”). See also Baker, supra note 21, at 56.
23. BAKER, supra note 21, at 56.
24. Id. at 57.
25. Id.
26. Id.
27. Id. (explaining there were three basic forms of trespass recognized: trespass to land, taking and carrying away goods, and assault and battery).
28. Id. (explaining that this limitation was related to “a policy of discouraging private disputes about mere wrongs in the royal courts” and describing the limitation as a “jurisdictional fetter”).
forty shillings, but the royal courts has no jurisdiction because they were not done forcefully. As a result, in the third quarter of the fourteenth century, the courts of the Chancery began to issue new writs for trespass without the force requirement.

These new writs were required to be more specific because they did not fit within one of the established causes of action that the original trespass writs supported. This broader family of trespass actions was referred to as “trespass on the case,” and it facilitated a reshaping of the common law from a rigid body of law into a flexible body of law. This change in thinking was capable of remedying complex actions that resulted from increasingly diverse relationships between people. For example, under the original forms of pleading, a guest could not recover from an innkeeper for damage or loss to his luggage under bailment because the relationship was not close or formal enough to constitute a bailment. However, the new forms of pleading enabled the guest to claim that the innkeeper violated a custom of the realm. This meant that the innkeeper violated an unspecific, but recognized, duty to take care of the guest’s goods. Eventually, this duty emerged as a basis of affixing liability on those who professed competence in certain callings, such as innkeepers and common carriers. However, the “invocation of the custom of the realm in those early duty of care cases” indicates that negligence began not as a vague standard of ordinary care, but as a general way to refer to “specific preexisting customs, conventions, and coordinating practices of the community.” This was the birth of the modern concept of negligence.

As the law developed, other occupations “had implied in them a duty of carefulness,” which the legal scholarship of the era recognized. In 1785,
William Blackstone explained “that everyone who undertakes any office, employment, trust, or duty, contracts with those who employ or entrust him, to perform it with integrity, diligence and skill.” Sir Edward Coke thought the recognition of these duties was necessary to ensure that only skilled people engaged in these callings. Evidently, negligence developed in the context of relationships and voluntary undertakings giving rise to liability—ideas that continue to permeate the duty requirement.

After initial willingness to find and impose a duty, the action of negligence rapidly developed from about 1825 onward—the time of the industrial revolution. As negligence developed, “a failure to perform a duty imposed by law” remained a necessary element. By the mid-eighteenth century, Lord Bathurst suggested the reasonable person principle, which guides the existence of a duty to this day:

Every man ought to take reasonable care that he does not injure his neighbour; therefore, wherever a man receives hurt throughout the default of another, though the same were not willful, yet it be occasioned by negligence or folly the law gives him an action to recover damages for the injury so sustained . . .

In essence, everyone was deemed to have a “duty” to act reasonably. Thus, the evolution of duty in negligence can be viewed as a means of expanding liability.

As liability extended, duties became more general and no longer depended on the specific relationships that had originally established the parameters of a “duty.” The problem then became deciding when a legal duty not defined by

---

41. Id. at 189.
42. Id. at 188-89.
43. Calista Menzhuber, Torts: In the Absence of Parents: Expanding Liability for Caretaker’s Failure to Protect Minors From Third-Party Harm – Bjerke v. Johnson, 35 WM. MITCHELL L. REV. 714, 719 (2009). See also Kaczorowski, supra note 11, at 1132 (demonstrating the clear delineation between situations where duties were imposed and those where they were not: “[T]he common law imposed on persons engaged in a common calling a duty of reasonable care and a standard of professional competence. A non-professional was not held to these standards unless he expressly undertook to perform with such competence and care.”).
44. See Winfield I, supra note 36, at 195 (“Perhaps one of the chief agencies in the growth of the idea is industrial machinery.”). See also Miller, supra note 11, at 14 (stating that Robert Rabin agreed that “negligence law was the nineteenth century’s progressive reaction to industrialization”) (citing Robert L. Rabin, The Historical Development of the Fault Principle: A Reinterpretation, 15 GA. L. REV. 925, 960-61 (1981)).
45. BAKER, supra note 21, at 346.
46. Id. at 347.
47. Id.
48. See Kelley II, supra note 11, at 1060.
an undertaking or custom should be recognized. This expansion of duties eventually resulted in the development of the reasonable person standard.\footnote{49}{Baker, supra note 21, at 346-47 (explaining “duties of care could not be imposed on everyone” and the development of the reasonable person standard by Lord Bathurst was to address this issue).}

There is agreement that “a general duty of care extending to individuals not in pre-existing relationships did not arise until the second quarter of the nineteenth century.”\footnote{50}{See Stein, supra note 11, at 705 (confirming the lack of scholarly disagreement over Winfield’s assertion that “the year 1837 marked a turning point’ with the cases of Vaughan v. Menlove and Langridge v. Levy”) (citations omitted). See also Coker B. Cleveland, Steamfitters Local Union No. 420 Welfare Fund v. Philip Morris: Is Swift v. Tyson Dead?, 25 Am. J. Trial Advoc. 145, 157 n.85 (2001) (“The general duty of care standard originally arose as a matter of pleading trespass on the case actions against common carriers, and was rapidly extended to other negligence cases.”); Kelley I, supra note 11, at 63 (agreeing that the “general duty of care pleading”- which didn’t require a specific duty to be claimed – seemed to have emerged by 1821).}

According to Prosser and Winfield, the mechanization associated with the industrial revolution sparked the rapid development of negligence during this period.\footnote{51}{Winfield’s theory was that the industrial revolution, and the machines that it produced, increased the need for modern negligence. Winfield I, supra note 36, at 195. Trains, for instance, were “notable neither for speed nor for safety. They killed any object from a Minister of State to a wandering cow, and this naturally reacted on the law.” Id. See also Publix Cab Co. v. Colo. Nat’l Bank of Denver, 338 P.2d 702, 708 ( Colo. 1959).}

In summary, the duty element of negligence derived from specific and recognized customs of the realm.\footnote{52}{See Percy H. Winfield, Duty in Tortious Negligence, 34 Colum. L. Rev. 41, 44-58 (1934) [hereinafter Winfield II] (tracing the development of duty in negligence).}

Prior to the industrial revolution, society was expected to act under the guidance of community standards and practices,
which a person living in that society could readily ascertain. As society became more integrated and diverse, it became harder to state negligence in terms of community standards. As a result, the law was forced to quantify the duty of all individuals as the duty to behave reasonably. This flexible standard of care is necessary today because “diverse economic and social needs have emerged and are in lively competition with each other.”

Today, establishment of the first two elements of negligence, duty and breach, turns on whether the defendant behaved unreasonably under the reasonable person standard. In fact, the formula in the Restatement (Third) of Torts for discerning unreasonable conduct is based entirely on a foreseeability-driven analysis, which was taken, unchanged, from the case of United States v. Carroll Towing Co. In Carroll Towing, prominent Judge Learned Hand enunciated his famous “B<PL” formula: if the burden to take

53. See Kelley II, supra note 11, at 1064 (describing the original conception of duty as tied to recognizable societal mores and asking whether that conception, suitable for 1830 England, is desirable or workable in the culturally and economically diverse United States of today).

54. Id.

55. See, e.g., Kelley II, supra note 11, at 1058.


57. Duty is defined as “[a]n obligation recognized by the law, requiring the actor to conform to a certain standard of conduct, for the protection of others against unreasonable risks.” SCHWARTZ ET AL., supra note 1, at 132. Breach is simply “[a] failure to conform to the required standard.” Id. See, e.g., Riedel v. Sheraton Bal Harbour Ass’n, 806 So. 2d 530, 532 (Fla. Dist. Ct. App. 2001) (confirming an actor is held to the standard of reasonable care). See also RESTATEMENT (SECOND) OF TORTS § 283 (1965) (“Unless the actor is a child, the standard of conduct to which he must conform to avoid being negligent is that of a reasonable man under like circumstances.”).

58. See RESTATEMENT (THIRD) OF TORTS § 3 (explaining that, in considering whether the person’s conduct lacks reasonable care, the following factors must be considered: “the foreseeable likelihood that the person’s conduct will result in harm, the foreseeable severity of any harm that may ensue, and the burden of precautions to eliminate or reduce the risk of harm”) (emphasis added).

59. 159 F.2d 169 (1947).

60. See White, supra note 11, at 103-05.

[In Carroll Towing.] Judge Hand sought to establish criteria for determining when an unattended barge was negligence and when it was not. [In doing so, he] explicitly addressed the issue of weighing the costs of accident avoidance against the expected harm to determine the defendant’s obligation: “[t]he owner’s duty . . . to provide against resulting injuries is a function of three variables: (1) The probability that she will break away; (2) the gravity of the resulting injury, if she does; (3) the burden of adequate precautions.
a precaution is less than the product of the potential magnitude of the harm and the likelihood of that harm if no precaution is taken, then a reasonable person has a duty to take that precaution.\(^\text{61}\) Ascertaining the factors of the “potential magnitude of harm” and the “probability of the harm occurring” both involve “foreseeing” what possible harm can occur from the actor’s conduct. Thus, the determination of whether a duty exists is essentially a foreseeability analysis.\(^\text{62}\) This foreseeability-based determination of duty and unreasonable conduct has enjoyed a resurgence in the law of negligence.\(^\text{63}\)

An important aspect of duty is that its existence is typically a question of law for the court.\(^\text{64}\) Therefore, judges in negligence cases act as gatekeepers or screeners of negligence actions.\(^\text{65}\) If the court determines that no duty of
care is owed to the plaintiff or that no duty of care exists, then the court may dismiss the action as a matter of law.66 This gatekeeper function is an important role of judges in negligence actions.67

However, once a plaintiff establishes unreasonable conduct or a breach of a duty, conventional negligence analysis proceeds to whether the defendant’s breach caused the plaintiff’s damages.

B. The Development of Proximate Cause

Causation in negligence is actually a combination of two elements, cause-in-fact and proximate cause.68 Cause-in-fact is “causation in the sense that term is referred to in every day speech.”69 This is satisfied if, “but-for the defendant’s conduct, the harm would not have occurred.”70 Apart from situations involving multiple breaches of duty, cause-in-fact is a simple test with little room for uncertainty.71

In contrast, proximate cause is not really about causation; rather, it provides a limit on negligence liability72 by answering the following complex question: if the plaintiff satisfies the first three elements of negligence, should the defendant be held liable?73 In the mid-eighteenth century, Lord Bathurst realized the need to limit negligence liability to foreseeable harms: “it is proper in [negligence] to prove that the injury was such as would probably follow from the act done.”74 Currently, all but two states extend liability to harm that was foreseeable, but not if the harm was unforeseeable as a result of the unreasonable conduct that comprises the breach of duty.75 More
specifically, proximate cause exists if the harm suffered was of the “kind/type/nature” that was foreseeable. In addition, unlike cause-in-fact, proximate cause is a question of fact for the jury.

While duty evolved as a way of expanding negligence liability and proximate cause evolved as a means to limit negligence liability, both are based on the same analytical fulcrum of foreseeability. And therein lies the rub.

III. PALSGRAF’S EFFECT ON DUTY AND PROXIMATE CAUSE

A. The Structure of Negligence Pre-Palsgraf; The Hitchhiking Scorpion

In my torts class, I begin the study of negligence by using the following hypothetical to demonstrate the elements of the tort:

I throw a baseball-sized metallic ball at the wall clock in a cavernously large lecture hall full of students as hard as I can, and I break the glass on the clock into many pieces. A splinter hits Student A, who is sitting under the clock, causing a severe cut. Additionally, one of the glass splinters from the shattered clock lands at my feet. The splinter just happens to be negatively charged and my shoe also happens to be negatively charged. As a result of the uniform polarity the splinter is repelled rapidly and it hits a fire ant in the carpet, which in turn is thrown through the air and lands on the top of the ceiling-based projector. It just so happens that a rare scorpion unknowingly attached itself to Student B’s shirt when she visited New Mexico, and it has unknowingly been residing on top of the projector.

See also Dodge v. McArthur, 223 A.2d 453, 454 (Vt. 1966) (“[P]roximate cause relates only to cause-in-fact, with no foreseeability required . . . .”); DOBBS, supra note 69, at 444.

The most general and pervasive approach to proximate cause holds that a negligent defendant is liable for all the general kinds of harms he foreseeably risked by his negligent conduct and to the class of persons he put at risk by that conduct. Conversely, he is not a proximate cause of, and not liable for injuries that were unforeseeable.

Id.

76. Stapleton, supra note 11, at 992 (explaining that a necessary requirement for finding proximate cause, or that the harm suffered was within the scope of liability of the negligent conduct, is proving that the harm suffered was of the kind defendants “should have reasonably foreseen and should have acted to avoid”).

77. Esquada, supra note 8, at 155 (explaining that “the vast majority of cases have held that proximate cause is a question of fact for the jury”).

78. See supra notes 21 - 76 and accompanying text (describing the development of negligence and duty as a way of finding liability in situations not previously recognized under the writ pleading standard). See also Kelley I, supra note 11, at 89 (discussing the emergence of proximate cause as a possible “judicial check” on jury awards in negligence); Kelley II, supra note 11, at 1045-46 (discussing two opposing scholarship views which attempt to explain the genesis, structure and role of proximate cause as a limit on liability).

for many months now. The fire ant in its confusion from the less-than-stellar travel experience bites the scorpion on the main nerve controlling the scorpion’s legs and as a result the scorpion experiences an involuntary muscle contraction which causes it to launch off the projector stand onto Student C, who is seated as far away from the clock as possible in the lecture hall. The scorpion stings Student C causing him physical harm. Student A and Student C sue me in negligence.

In order for each student to recover, the students must establish each of the four elements of negligence: 1) duty, 2) breach, 3) causation, and 4) damages. In this case, the duty and breach are proven if a reasonable person would not throw a metallic ball against a clock in an enclosed room full of students. Because I did, I behaved unlike a reasonable person and breached a duty. Cause-in-fact is established because, if I had not breached the duty to act reasonably, then the ball would not have been thrown and the injuries to both students would not have occurred.

However, proximate cause renders different outcomes for each of the two students’ suits. A cut from a glass splinter to a student close to the clock is generally a foreseeable result of smashing a glass clock in a room full of students. Thus, proximate cause would exist for student A, who was injured in this foreseeable manner. However, Student C would likely be denied recovery because, although the throwing of the ball did factually cause the scorpion sting, a scorpion sting to a student is not a foreseeable result of smashing a glass clock in a classroom full of students. The Palsgraf decision, which had an enormous impact on the law of negligence, threw this idealized version of the tort of negligence into question.

B. Palsgraf’s Conflation of Duty and Proximate Cause

The gist of the Palsgraf case is as follows:

Plaintiff was standing on a platform of defendant’s railroad after buying a ticket to go to Rockaway Beach. A train stopped at the station, bound for another place. Two men ran forward to catch it. One of the men reached the platform of the car without mishap, though the train was already moving. The other man, carrying a package, jumped aboard the car, but seemed unsteady as if about to fall. A guard on the car, who had held the door open, reached forward to help him in, and another guard on the platform pushed him from behind. In this act, the package was dislodged, and fell upon the rails.

81. See Kelley I, supra note 11, at 96-99.
It was a package of small size, about fifteen inches long, and was covered by a newspaper. In fact it contained fireworks, but there was nothing in its appearance to give notice of its contents. The fireworks when they fell exploded. The shock of the explosion threw down some scales at the other end of the platform, many feet away. The scales struck the plaintiff, causing injuries for which she sues.

The plaintiff filed a negligence suit against the railroad seeking to recover damages resulting from the accident. The sole question presented to the jury at trial was "whether the defendant’s employees were ‘careless and negligent in the way they handled this particular passenger after he came upon the platform and while he was boarding the train.’" The jury returned a verdict for the plaintiff. On appeal, the Supreme Court of New York affirmed, but one dissenting judge expressed that the verdict should be reversed for lack of proximate cause. Nonetheless, both the majority and dissent agreed that a duty to handle passengers had been breached and that there was negligent conduct in the way the passenger was handled. The defendants appealed and the New York Court of Appeals reversed. Judge Cardozo authored the decision over Judge Andrew’s dissent:

The conduct of the defendant’s guard, if a wrong in its relation to the holder of the package, was not a wrong in its relation to the plaintiff, standing far away. Relatively to her it was not negligence at all. Nothing in the situation gave notice that the falling package had in it the potency of peril to persons thus removed. Negligence is not actionable unless it involves the invasion of a legally protected interest, the violation of a right. ‘Proof of negligence in the air, so to speak, will not do.’

Cardozo mandated that there must be a specific relationship between the plaintiff and the defendant that gives rise to the foreseeability of harm resulting from defendant’s action or inaction before a duty can be found. Perhaps

84. *Id.* at 413.
85. *Id.* at 414.
86. The Supreme Court of New York is an intermediate appellate court.
88. *Id.* at 414-15.
89. In New York, the Court of Appeals is the highest appellate court.
91. *Id.* at 99, 101.
93. *Id.* See also *Palsgraf*, 162 N.E. at 99-100 (“In every instance, before negligence can be predicated of a given act, back of the act must be sought and found a duty to the individual complaining, the observance of which would have averted or avoided the injury.”) (quoting McSherry, C. J., in *W. Va. Cent. & P. Ry. Co. v. State*, 54 A. 669, 671 (Md. 1903)).
Cardozo was trying to contain the rapidly evolving tort of negligence by reminding us that duties were originally relational concepts. According to Cardozo, an action that constitutes a breach of duty to one person does not necessarily constitute a breach of duty to another person, even if the act and the harm caused are identical in both circumstances.

Cardozo concluded that the issue of proximate cause was not relevant because the elements of duty and breach were lacking:

The law of causation, remote or proximate, is thus foreign to the case before us. The question of liability is always anterior to the question of the measure of the consequences that go with liability. If there is no tort to be redressed, there is no occasion to consider what damage might be recovered if there were a finding of a tort.

However, under Cardozo’s formulation of negligence, had the plaintiff been a person to whom a duty was owed, then the breach would ‘entail liability for any and all consequences, however novel or extraordinary.’

Judge Andrews’s dissent revealed a completely different conceptualization of the elements of negligence:

Negligence he said, does not depend upon a relation between the defendant and the plaintiff; it is a wrong toward anyone in fact injured by the negligent act. There is a duty toward the world at large not to [act unreasonably] toward any person; and when an injury to the plaintiff results from a breach of this duty, she may have a cause of action.

Andrews viewed proximate cause as a practical, multifaceted factor limiting liability for breach of duties owed to the world. According to Andrews, the issue of foreseeability that Cardozo felt was locked up in the “duty” prong of negligence should apply in the “proximate cause” prong of the analysis. His actual definition of proximate cause made no apologies for the fact that it is a somewhat arbitrary concept, driven by public policy that

---

94. See William L. Prosser, Palsgraf Revisited, 52 Mich. L. Rev. 1, 5 (1953) [hereinafter Palsgraf Revisited] (interpreting Cardozo as saying, “Negligence must be a matter of some relation between the parties, some duty, which could be founded only on the foreseeability of some harm to the plaintiff in fact injured.”).

95. Palsgraf, 162 N.E. at 100 (“The plaintiff sues in her own right for a wrong personal to her, and not as the vicarious beneficiary of a breach of duty to another.”).

96. Id. at 101.

97. Id.

98. Id. Does this statement foreclose the existence of a proximate cause analysis in a negligence suit?


100. Palsgraf Revisited, supra note 94, at 6.

101. Palsgraf, 162 N.E. at 104 (explaining that general foreseeability “[m]ay have some bearing” on the issue of what is proximate cause).
tends to limit liability: “What we do mean by the word “proximate” is, that because of convenience, of public policy, of a rough sense of justice, the law arbitrarily declines to trace a series of events beyond a certain point. This is not logic. It is practical politics.”

The fascinating thing about the opinion is that both judges used similar foreseeability-based analyses to decide whether a plaintiff should be compensated. Cardozo used foreseeability to decide if the defendant owed a duty to the plaintiff, but Andrews used foreseeability to determine if the harm to the plaintiff was proximately caused by a breach of a duty owed to the world.

Prosser, in his piece entitled *Palsgraf Revisited*, suggested that Cardozo’s opinion in *Palsgraf* was influenced by the prominent legal realists of the era. Prosser politely described this influence as “one of those accidents which shape the course of the law,” and he noted how the appellate opinion in *Palsgraf* fell in to the hands of Francis H. Bohlen, the reporter for the first Restatement of Torts. Bohlen was “disposed to accept” the views of the legal realists Bingham and Green. Bingham felt that legal duties should be a “concrete, specific obligation in a concrete, specific case, authoritatively determined by the court . . . .” Likewise, Green’s views on the subject were “almost identical to that of Bingham’s,” except that he “emphasized much more strongly than Bingham the positive legal character of the duties involved in torts cases . . . .”

Bohlen prepared a draft of the Restatement to submit to his advisers, which reflected the legal realist view of duty. Because “*Palsgraf* provided a perfect illustration, [he] added a statement of the facts to his explanatory notes,” which he forwarded to the advisers for comment. “Among the advisers who met to consider this draft was Chief Judge Cardozo of the New York Court of Appeals, who found himself suddenly and unexpectedly

---

102. *Id.*
103. *Id.* at 103.
105. *See, e.g., Dobbs, supra* note 69, at 451 (explaining that Cardozo’s conception of the *Palsgraf* foreseeability problem as a duty problem while Andrews believed the issue was one of proximate cause).
107. *Id.*
108. *Id.*
109. *Id.* (viewing regard foreseeability as part of the duty analysis).
110. *Kelley I, supra* note 11, at 82-83.
111. *Id.* at 95.
113. *Id.*
confronted with a discussion . . . of a lower New York decision which might very possibly be appealed to his court.”

Cardozo felt his presence at the discussion was appropriate, but that his active participation and voting were inappropriate. Cardozo ultimately participated in the debate, which centered on the proper role of proximate cause and duty in negligence. The advisers adopted Bohlen’s proposal based on Bingham’s and Green’s formulation of duties. Cardozo agreed with the advisers and his opinion in Palsgraf adopted this view. Not surprisingly, the view that Cardozo adopted “go[es] far beyond the scope of arguments of counsel in the case, who were still foundering in the morass of proximate cause.”

Applying Palsgraf to the hitchhiking scorpion hypothetical produces two correct statements involving different elements of negligence, both based on foreseeability, that preclude liability to Student C. After Palsgraf, Cardozo-leaning jurists would say that Student C cannot recover because harm to her is not foreseeable because she is an unforeseeable plaintiff to whom no duty is owed. However, according to Andrews, because the harm that occurred to Student C was an unforeseeable consequence of a breach of a duty owed to the world in general, there is no proximate cause and Student C is barred from recovery.

Key to understanding this difference is the manner in which the specificity of the stated duty exacerbates the problem of duplicative foreseeability. The more specifically the duty is formulated, the harder it is to distinguish unreasonable conduct from proximate cause. The case of Anonymous Plaintiff v. Anonymous Cruise Lines discussed in the next section perfectly illustrates this issue.

114. Id.
115. Id.
116. Id. at 4-5.
117. Id.
119. Palsgraf Revisited, supra note 94, at 4-5.
120. See Palsgraf, 162 N.E. at 99-105.
121. See id. at 99-101.
122. See id. at 101-05.
123. See Kelley I, supra note 11, at 53 (“Prosser and Keeton acknowledge that the foreseeable risk test of proximate cause has gained widespread acceptance, but criticize the foreseeable risk theory as too flexible to be a meaningful guide to decisions or to predicting results.”).
124. The actual names of the plaintiff and the defendant have been redacted to avoid identification of my former client.
IV. Duplicative Foreseeability as a Function of Duty Specificity

The more specifically the duty is stated, the more acute the overlap in foreseeability is between unreasonable conduct and proximate cause. Interestingly, most scholars fail to acknowledge that the formulation of the duty statement in a negligence case is not something the court divines, nor is it always fixed. In reality, the duty statement is a function of the compromise between the plaintiff’s and defendant’s proposed duty formulations. Therefore, as the following case demonstrates, the duty statement that a court adopts in a particular case is directly related to the extent of duplicative foreseeability analysis and corresponding inability to distinguish between the two.

A. The Case of Anonymous Plaintiff v. Anonymous Cruise Lines

In 2004, I represented a plaintiff, whom I will refer to here as “Anonymous Plaintiff” or “AP,” in a suit against his employer, whom I will refer to here as “Anonymous Cruise Line” or “ACL.” AP alleged in his complaint that the defendant-crew’s negligence caused him to suffer a rape at the hands of a passenger. AP had reported to the relevant authorities on board the ship where he worked that his stateroom door could not be locked. Sometime after he made the report, a drunk male passenger entered AP’s stateroom and performed nonconsensual fellatio on AP while AP was sleeping. The complaint claimed negligence under the Jones Act, alleging that the defendant had a duty to provide a reasonably safe place to work and that ACL breached its duty because it did not fix the lock on AP’s door in a timely manner. The complaint further alleged that the unreasonable conduct was the cause-in-fact and proximate cause of the harm.

Reminiscent of Justice Cardozo’s Palsgraf opinion, ACL cited many examples of specific duties in its motion to dismiss the complaint for failure to state a claim. The motion urged the court to conclude that even though a

125. See infra Parts IV.B, V.B.
126. See infra Part IV.A.
127. See infra Part IV.A.
129. Id. ¶¶ 14-18.
130. Id. ¶ 19.
132. Complaint, supra note 128, ¶ 32.
133. Id. ¶ 33(a).
135. Defendant’s Motion to Dismiss at 2-4, Anonymous Plaintiff (on file with author).
duty of reasonable care existed, it did not extend to the prevention of third parties’ unforeseeable criminal acts.”

The motion noted that no duty is owed in negligence to protect someone from “the willful criminal act of a third person which could not be foreseen or anticipated.” If there was no duty, then there could be no unreasonable conduct or breach, and dismissal would be appropriate as a matter of law.

Mirroring Justice Andrews’s dissent in *Palsgraf*, ACL employed what was essentially the same foreseeability argument to assert that there was no proximate cause, even if a duty existed. ACL argued that even though it may have owed AP the general duty to provide a reasonably safe place to work, there was no proximate cause as a matter of law because the specific type of harm that occurred was not foreseeable:

> It is our contention judge, that under no circumstances could anyone have foreseen this. No one could foresee a passenger coming into a crewmember’s room even if a lock is broken. What you have here is that the door is closed just like all other doors and someone actually had to go in the room and then perform fellatio on another human being against that person’s will while that person is sleeping. Perhaps if this was a case where someone went into the room and grabbed his wallet and left, maybe that is foreseeable.

The court denied ACL’s motion to dismiss because the judge held that the duty element was satisfied by the general duty of maritime employers to provide a safe place to work. The court further agreed that duty is a question of law and that duty is broadly interpreted as the duty to behave like a reasonable person. I had argued that ACL was trying to conflate the four elements of negligence by narrowly stating the duty as the duty to protect crewmembers from third party criminal acts. In doing so, ACL was inviting the judge to engage in a factual analysis under the guise of duty, although it is typically analyzed under the prong of breach or proximate cause. ACL’s position would have required the judge to usurp the jury’s role by considering facts in the duty prong (question of law) that properly belonged in breach or proximate cause prong (questions of fact).

Today, as a torts professor, I realize that, under *Palsgraf*, ACL’s arguments were as correct as mine. The court accepted a view of duty that was

---

136.  *Id.*

137.  *Id.* at 4. See also *id.* at 2 (stating “[ACL] does not have a duty to protect crewmembers from third party criminal attacks”).

138.  See *Palsgraf*, 162 N.E. at 101-05.

139.  Defendant’s Motion to Dismiss, *supra* note 135, at 5-8

140.  Transcript of Hearing on Defendant’s Motion to Dismiss, at 6:5-15, *Anonymous Plaintiff* (on file with author) [hereinafter Motion to Dismiss Hearing].


142.  *Id.* at 13:17-25.
equivalent to Andrews’s general, unspecific view of duty in *Palsgraf*. However, if the court had adopted Cardozo’s view of duty, the defendant’s specific duty statement would have been correct. This would have enabled the judge to consider more facts in deciding whether a duty existed under the guise of deciding the existence of duty as a matter of law. We were both correct because modern courts are split between Cardozo’s conception of duty being specific and relational and Andrews’s view of the general duty of reasonable care owed to the world at large.\(^{143}\) As a result, duties in modern negligence cases run the gamut between the two views.\(^{144}\) *Palsgraf* reinforces the problematic overlap in the use of foreseeability to determine both the existence of a duty and of proximate cause.\(^{145}\) Every torts student should be exposed to this overlap when comparing the formulation of duty in *Carroll Towing*\(^{146}\) and the existence of liability in *Wagon Mound No. 2*.\(^{147}\)

\(^{143}\) See infra Part V.A.

\(^{144}\) Compare *Laster ex rel.* *Laster v. Norfolk S. Ry. Co.*, 13 So. 3d 922, 927 (Ala. 2009) (“A possessor of land owes a duty to exercise reasonable care to eliminate an artificial condition on land that poses a danger to children.”) (emphases omitted); *Elizondo v. Ramirez*, 753 N.E.2d 1123, 1127 (Ill. App. Ct. 2001) (“A landowner or occupier has a duty to protect others against criminal acts where there is a special relationship between the landowner and the injured person . . . . Also, the criminal attack must have been reasonably foreseeable.”); *Roland Bros. v. Direct Gen. Ins. Co.*, 11 So. 3d 1154, 1157 (La. Ct. App. 2009) (“A motorist has a duty not to cause an obstruction to the normal flow of traffic on a highway or to adequately warn approaching motorists of any obstructions he does cause.”); *Hummel v. Cnty. of Nassau*, 869 N.Y.S.2d 169, 170 (App. Div. 2008) (“A parent owes a duty to protect third parties from clearly foreseeable harm arising from the child’s improvident use or operation of a dangerous instrument, where such use is found to be subject to the parent’s control.”); *and Rosengren v. City of Seattle*, 205 P.3d 909, 914 (Wash. Ct. App. 2009) (“[A]n abutting land owner has a duty to exercise reasonable care that the trunks, branches, or roots of trees planted by them adjacent to a public sidewalk do not pose an unreasonable risk of harm to a pedestrian using the sidewalk.”), with *Wash. Metro. Area Transit Auth. v. Barksdale-Showell*, 965 A.2d 16, 24 (D.C. 2009) (“[P]arties are charged with the duty to act with reasonable care under the circumstances.”); *Afarian v. Mass. Elec. Co.*, 866 N.E.2d 901, 905 (Mass. 2007) (“[A] a general principle of tort law, every actor has a duty to exercise reasonable care to avoid physical harm to others.”) (citations omitted); *Lukasik v. San Antonio Blue Haven Pools, Inc.*, 21 S.W.3d 394, 403 (Tex. Ct. App. 2000) (“Every person has a duty to exercise reasonable care to avoid a foreseeable risk of injury to others.”); *Gritzner v. Michael R.*, 611 N.W.2d 906, 912 (Wis. 2000) (“At the very least, every person is subject to a duty to exercise ordinary care in all of his or her activities.”); *and Parrilla v. King Cnty.*, 157 P.3d 879, 884 (Wash. Ct. App. 2007) (“If a risk is foreseeable, an individual generally has a duty to exercise reasonable care to prevent it.”).

\(^{145}\) See, e.g., *Stapleton, supra* note 11, at 988-97 (discussing the overlap in foreseeability as it relates to duty and proximate cause). See also *Palsgraf v. Long Island R.R. Co.*, 162 N.E. 99 (1928).

\(^{146}\) See infra Part VI.A. (discussing *Carroll Towing* as the source of a foreseeability based analysis of duty).

B. Doctrinal Examples of Duplicative Foreseeability

In *Wagon Mound No. 2*, while refueling at Caltex Wharf in Sydney Harbor, a ship (the Wagon Mound) discharged a large amount of fuel oil into the water. The fuel oil accumulated around another wharf where repair work involving oxy-acetylene welding was occurring on two vessels. The wharf manager contacted the Wagon Mound’s manager and expressed concern that the falling hot pieces of metal from the welding might ignite the fuel. The Wagon Mound’s manager assured him that it was safe to continue welding, but, two days later, the oil ignited, presumably because droppings from the welding ignited some “inflammable material” floating on the water. The ensuing fire extensively damaged the two ships that were being repaired.

The owners of the two damaged vessels sued the Wagon Mound, alleging nuisance and negligence. It was already established that the Wagon Mound had breached its duty in discharging the oil. The major issue in *Wagon Mound No. 2* “was whether the duty owed by *Wagon Mound* to the plaintiffs with respect to the spillage of the oil included the risk of injury to their vessels by the burning oil.” This is the identical question that the Restatement (Third) asks in determining whether the harm that occurred was within the scope of the risk created by the negligent conduct. In determining whether there was proximate cause, or whether the harm was foreseeable or within the scope of the risk, the Privy Council restated Learned Hand’s “B< PL” formula in deciding if a duty exists:

But it does not follow that no matter what the circumstances may be, it is justifiable to neglect a risk of such a small magnitude. A reasonable man would only neglect such a risk if he had some valid reason for doing so, e.g., that it would involve considerable expense to eliminate the risk. He would weigh the risk against the difficulty of eliminating it . . . . If it is clear that the reasonable man would

151. *Id.*
154. *See Green II, supra* note 148, at 200 (“The negligence [or breach of duty] of the Wagon Mound engineers in allowing the oil to spill in such quantities with respect to the peril created by fire for the plaintiffs’ vessels was readily resolved by the Privy Council.”).
155. *Id.* at 201.
156. *See infra* Part VI.B.
have realised or foreseen and prevented the risk, then it must follow that the appellant is liable in damages.\footnote{157. Overseas Tankship (U.K.) Ltd. v. Miller Steamship Co., [1966] 1 A.C. 617, 630, 644.}

In a scathing critique, Leon Green described this nonsensical and confusing dual application of foreseeability as “overloading foreseeability.”\footnote{158. Green II, supra note 148, at 203-06 (“This double screening through the bankrupt terminology of negligence law serves to protect these courts from any risk of professional backlash, inasmuch as the gibberish is so massively profound that it has gained a high degree of sanctity.”).} He explained that this problematic duplicate use of foreseeability is exacerbated when duties are stated specifically.\footnote{159. See Green I, supra note 4, at 243-44.} According to Green, this happens because specifically stated duties typically encompass the type of harm, something normally reserved for a proximate cause analysis.\footnote{160. Id.} A specifically stated duty relieves the plaintiff of anything other than a superficial burden in establishing breach, whereas a broadly stated duty shifts most of the foreseeability analysis to determination of breach.\footnote{161. Id. at 744-46.}

The result is restricted utility of normative discussions regarding the elements of negligence, such as those engaged in by Professor Cardi.\footnote{162. Cardi II, supra note 11, at 744-62 (generalizing and oversimplifying the clarity of the role of foreseeability in establishing breach, proximate cause, and duty).} Professor Cardi, for example, states that foreseeability is least uncontroversial when applied in determining breach.\footnote{163. Id. at 744.} However, this view of foreseeability is based on a flawed assumption that duty statements are well-defined reflections of the standard of care as determined by a judge,\footnote{164. Id.} which are universally stated in terms of what a reasonable person would do under the circumstances.\footnote{165. Id.} In a breach analysis, Professor Cardi explains that foreseeability is relevant only to blameworthiness\footnote{166. Id. at 744-46.} and not to the extent of possible harm that the actor’s conduct created. To illustrate, Professor Cardi uses the following example:

> [S]uppose that two drinking companions consume an equal amount of alcohol, then pass out while driving home in their respective cars. One car drifts off the road into a private lawn, knocking down the owner’s mailbox en route. The other car veers into a parked car, killing the two teenagers within. Although the two drivers will be liable for different amounts of damages (assuming the mailbox was not the work of Midas), the drivers are equally negligent, equally
blameworthy. In other words, they each breached their duty of care in the same way and to the same “extent.”

Professor Cardi’s example works because the duty when driving while intoxicated is crisp and precise, making breach a non-analytical issue. In fact, driving while intoxicated might refer to one of those societal conventions Professor Kelley thinks would be clearly identifiable in a “pre-Holmesian” restatement of negligence. Typically, both drivers would be considered legally intoxicated by reference to some easily discernable standard, such as a blood alcohol level (BAC) in excess of .08. Therefore, if the driver had an accident while having a BAC greater than .08, then he would have breached his duty. In such a neat and precise duty situation, breach is largely unconsidered with foreseeability and is “uncontroversial.”

However, dismissing breach as uncontroversial does not reflect the reality of litigation. Consider a case where a tenant sues a landlord for harm caused from a criminal entering the tenant’s apartment through a door that is not functioning because the landlord failed to maintain the door’s lock. The relevant duty may appropriately be defined as the general duty to take reasonable steps to protect a tenant. However, a different lawyer might specifically plead that a landlord’s duty in a high crime area includes properly maintaining the door’s locks.

Professor Cardi’s generalized statement about the uncontroversial nature of breach only works if the duty statement is specific. When the duty is stated specifically, as in the latter formulation in the example above, the breach analysis becomes simple for the plaintiff. Because the specific duty is deemed to exist, the foreseeability analysis is locked up in establishing whether a duty existed to fix the lock. If the defendant had a specific duty to fix the lock, then it breached its duty when it failed to fix the lock. But there is no breach if the defendant maintained the lock. In this situation, the

167. Id. at 746.
168. See Kelley II, supra note 11, at 1064-66 (describing reasons for rules which are “clear definite and certain” reflections of the “community’s applicable safety convention”).
169. See, e.g., OHIO REV. CODE ANN. § 4511.19(b) (LexisNexis 2011) (stating that it is a crime for a person to operate a vehicle with a blood alcohol concentration greater than .08%).
170. Cardi II, supra note 11, at 744.
171. See RESTATEMENT (SECOND) OF TORTS §§ 287-88 (1965) (illustrating that a violation of a statute is not an automatic breach of a duty, for example, when the purpose of the statute is to “protect against other harm than that which has resulted”).
172. See supra Part IV.A, (discussing the Anonymous case and the variety of appropriate duty statements in a particular case post-Palsgraf).
173. See Cardi II, supra note 11, at 744.
174. Id. at 746.
175. Id.
176. Id.
analysis is even less controversial than Cardi states because breach is reduced to nothing more than a simple yes/no question devoid of foreseeability. 177

However, consider what would happen if the court adopted a broad duty statement, such as the duty to take reasonable steps to protect a tenant. What the landlord has to do to comply with this duty becomes a fact-intensive analysis. 178 The court would have to reach beyond the simplistic analysis Professor Cardi thinks is suitable for all breach analyses. 179 In fact, Professor Cardi describes the appropriate analysis when he explains the difference between two drivers who are both under a general duty of care to drive reasonably and fail to do so, but one is driving on the Long Island Freeway and one on a deserted country road. 180 Cardi states that the appropriate analysis in determining reasonableness in this breach situation is exactly the same fact-intensive analysis involving foreseeability that Wagon Mound No. 2 employed to determine the scope of liability. 181

Negligence per se perfectly illustrates how specifically stated duties make the breach element simple to resolve. In negligence per se, specific duties are adopted from statutes to define the conduct of the reasonable person. Because the Constitution requires that statutes are specific in order to comply with the requirements of due process, 182 the adopted duties are clear and precise. 183 Nonetheless, the overlap between duty and proximate cause still is problematic.

In order for a court to accept a statute’s mandate or prohibition as defining what a reasonable person would do, courts apply the class harm test or the statutory purpose doctrine. 184 As a result, negligence per se also illustrates the overlap between proximate cause and unreasonable conduct. The class harm test states that a statute can only be used to define the reasonable person

177. Id.
178. For example, crime rates in the area, whether or not the specific landlord was on constructive or actual notice of the crime rates, etc. would all have to be factored in to determine if the landlord acted reasonably or a breach of duty occurred.
179. See Cardi II, supra note 11, at 744-45.
180. Id. at 745-46 (explaining how pulling your car over would be more unreasonable and have more potential for foreseeable and severe harm if one pulls over on the Long Island Expressway rather than on a deserted road).
182. See, e.g., Connally v. General Constr. Co., 269 U.S. 385, 391 (1926) (explaining that “a statute which either forbids or requires the doing of an act in terms so vague that men of common intelligence must necessarily guess at its meaning and differ as to its application, violates the first essential of due process of law”).
183. Did the defendant violate the statutory mandate? If yes, then there was a breach; if no, then there was no breach.
184. See, e.g., Mike Steenson, A Comparative Analysis of Minnesota Products Liability Law and The Restatement (Third) Of Torts: Products Liability, 24 WM. MITCHELL L. REV. 1, 31 (1998) (“Violation of a statute is negligence per se if the statute satisfies the statutory purpose analysis . . . .”).
standard when: (1) the plaintiff is within the class of persons that the statute is designed to protect; and (2) the plaintiff’s harm is of the type that the statute was designed to address. Under this test, in deciding if the statute defines reasonable conduct, the class harm test requires a class-of-person/type-of-harm analysis that is widely regarded as appropriate for the proximate cause element of negligence.

Similarly, Professor Cardi’s conclusion that proximate cause can easily be separated from the question of whether conduct is reasonable is not buttressed by his reliance on Wagon Mound. Professor Cardi refers to the 1961 case commonly referred to as Wagon Mound No. 1 while ignoring the contrary ruling in the subsequent Wagon Mound No. 2 case. The facts of both cases are identical. In Wagon Mound No. 1, the dock owners who were welding sued the operators of the Wagon Mound. “In New South Wales at the time, contributory negligence was a complete bar to recovery.” Thus, the dock owners were stymied in their proof of proximate cause. If they had tried to prove that the fire was foreseeable, then they would also have been admitting that they acted negligently by welding over the oil slick, and contributory negligence would have barred their recovery. Thus, the lack of a finding of proximate cause was probably due to the dock owner’s reluctance to admit or prove the fire’s foreseeability. As a result of this outcome, Professor Cardi concludes that a finding of duty and breach with no finding of proximate cause is indicative of a clear delineation between duty and proximate cause.

In contrast, in Wagon Mound No. 2, where the owner of the ships tied to the dock sued the Wagon Mound operators alleging negligence, the plaintiffs were not reluctant to prove the foreseeability of harm because they could not be found contributorily negligent. In Wagon Mound No. 2, the court found that liability existed via a separate analysis from whether the

186. See Cardi II, supra note 11, at 748-49 (explaining the various factors courts consider in determining proximate cause).
187. Id. at 749-750.
188. Id. at 749 n.47.
189. See supra notes 148-155 and accompanying text.
190. See Kelley I, supra note 11, at 101-04 (describing in detail the identity of facts but difference in parties for Wagon Mound No. 1 and Wagon Mound No. 2).
192. Id.
193. See supra notes 148-155 and accompanying text (detailing the facts of the Wagon Mound cases).
196. Id.
defendant breached a duty. 197 However, in reaching this conclusion, the foreseeability test adopted for proximate cause became virtually indistinguishable from the foreseeability test commonly applied to duty. 198

From a realist’s perspective, the outcomes in Wagon Mound No. 1 and No. 2 differ only because the plaintiff’s proof of foreseeability was stymied in Wagon Mound No. 1 by contributory negligence. 199 From a normative approach, it is easy to gloss over the tangible mechanics at the evidentiary level that contributed to the outcome. The Wagon Mound cases do nothing to differentiate between duty and proximate cause, as has been suggested by Professor Cardi. 200

Academia is not the only place with fantastic theories abound, as a result of the failure to accept that foreseeability has only one meaning in negligence. Modern courts continue to struggle to distinguish between unreasonable conduct and proximate cause. This is primarily due to the uncertain role foreseeability plays in determining whether a duty existed, whether it was breached, and whether liability should attach to the unreasonable conduct. 201

V. POST-PALSGRAF THEORETICAL MISSTEPS IN DISTINGUISHING BETWEEN DUTY AND PROXIMATE CAUSE

A. Judicial Missteps

In order to distinguish the role of foreseeability in unreasonable conduct, courts employ a directionally based analysis that separates duty from proximate cause. 202 While duty employs a forward-looking foreseeability analysis, proximate cause employs a backward-looking foreseeability analysis:

Foreseeability in the context of proximate cause involves evaluating the particular circumstances of an incident after the incident occurs. According to the American Law Institute, foreseeability for proximate cause purposes is determined from a perspective that is “after the event and looking back from the harm to the actor’s negligent

197. Id.
198. See supra note 157 and accompanying text (discussing the similarity between Carroll Towing and Wagon Mound No. 2). See also Kelley I, supra note 11, at 749-50 (suggesting that the Wagon Mound cases are illustrative of cases adopting and applying the modern theory of proximate cause and concluding that the analysis overlaps with breach of duty analysis).
199. See Prosser Teaching Manual at 276 (stating that the welders’ failure to cease operations in Wagon Mound No. 1 could have found to be contributory negligence).
200. See Cardi II, supra note 11, at 748-50.
201. See infra Part V.A.
202. See SCHWARTZ, ET. AL., supra note 1, at 132 (stating that the courts require segregation of the elements of a negligence cause of action).
Thus, when determining proximate cause, foreseeability is determined based on hindsight, and accounts for the circumstances that actually occurred.\textsuperscript{204}

On the other hand, “[f]oreseeability for purposes of establishing whether a defendant’s conduct created a duty to a plaintiff depends on whether the defendant should have foreseen a risk in a given set of circumstances. In this setting, foreseeability is forward-looking.”\textsuperscript{205} Another common distinction is that duty involves a generalized foreseeability analysis, while proximate cause employs a more fine-tuned analysis focused on the facts of the case.\textsuperscript{206}

Indiana’s well-documented struggle to apply these tests illustrates the courts’ failure to distinguish between duty and proximate cause.\textsuperscript{207} In \textit{Goldsberry v. Grubbs},\textsuperscript{208} the court attempted to distinguish these two elements by relying on two arguments.\textsuperscript{209} The issue in \textit{Goldsberry} was whether a telephone company owed a driver a duty when placing telephone poles near roadways.\textsuperscript{210} The claim arose after Grubbs, the driver, lost control of her car and collided with a telephone pole.\textsuperscript{211} As a result of the accident, the passenger, Goldsberry, sustained serious injuries, which left her a quadriplegic.\textsuperscript{212} Goldsberry sued several defendants, including General Telephone Company of Indiana, Inc. (“GTE”).\textsuperscript{213} The trial court granted GTE summary judgment because it concluded that GTE did not owe Grubbs a duty.\textsuperscript{214}

The \textit{Goldsberry} court had to make sense of conflicting precedent because two previous cases had reached different conclusions on the same issue.\textsuperscript{215} Both cases relied on \textit{Webb v. Jarvis},\textsuperscript{216} which set forth the following three factors to be used when determining the existence of a duty: “1) the relationship between the parties; 2) the reasonable foreseeability of harm to the person injured; and[3) public policy concerns” (collectively, the “\textit{Webb} factors”).

\begin{itemize}
\item \textsuperscript{204} \textit{Id.} (quoting \textit{RESTATEMENT (SECOND) OF TORTS} § 435 (1965)).
\item \textsuperscript{205} \textit{See, e.g.}, Lopez v. Three-Rivers Elec. Coop. Inc, 26 S.W.3d 151, 156 (Mo. 2000).
\item \textsuperscript{206} \textit{Goldsberry}, 672 N.E.2d at 479.
\item \textsuperscript{207} \textit{See, e.g.}, Theodore R. Boehm, \textit{A Tangled Webb--Reexamining the Role of Duty in Indiana Negligence Actions}, 37 IND. L. REV. 1, 5-9 (2003).
\item \textsuperscript{208} 672 N.E.2d 475 (Ind. Ct. App. 1996).
\item \textsuperscript{209} \textit{Id.} at 478-79.
\item \textsuperscript{210} \textit{Id.} at 479-80.
\item \textsuperscript{211} \textit{Id.} at 477. There was a dispute about whether the car flipped and then hit the pole or whether the car flipped because it hit the pole.
\item \textsuperscript{212} \textit{Ibid.}
\item \textsuperscript{213} \textit{Ibid.}
\item \textsuperscript{214} \textit{Goldsberry}, 672 N.E.2d at 477.
\item \textsuperscript{215} \textit{Id.} at 478.
\item \textsuperscript{216} \textit{Webb v. Jarvis}, 575 N.E.2d 992 (Ind. 1991).
\end{itemize}
However, courts have struggled to apply these factors because the 
*Webb* court did not offer any guidance.\(^{218}\)

The *Goldsberry* court determined that the lower courts inconsistently 
employed the *Webb* test because they could not “distinguish between the 
foreseeability component of the duty analysis and the foreseeability 
component of proximate cause.”\(^{219}\) The court determined that “foreseeability 
[in proximate cause] is determined based on hindsight, and accounts for the 
circumstances that actually occurred.”\(^{220}\) This is different from duty, which 
“by logical deduction . . . must be something different than the foreseeability 
component . . . it must be a lesser inquiry.”\(^{221}\)

One commentator suggests that the court forcibly arrived at this distinction 
because of the fear that “[i]f [duty and proximate cause] were not two distinct 
concepts, as Judge Bailey recently put it, deciding the duty question would 
subease the entire law of negligence, i.e., duty breach and proximate cause, 
into the duty question.”\(^{222}\) This conclusion has led to copious scholarship 
dedicated to trying to see the emperor’s new clothes.\(^{223}\) If courts are not doing 
what they state they are in resolving a claim, then the resolution is arbitrary 
and foreign to an orderly judicial system.\(^{224}\)

In a further attempt to distinguish duty from proximate cause, the 
*Goldsberry* court concluded that the three factors mentioned above are 
elements of duty, but not of proximate cause.\(^{225}\) However, more modern 
definitions of proximate cause utilize these same factors to decide whether 
proximate cause exists.\(^{226}\)

Later, in *Hammock v. Red Gold*, the court receded from *Goldsberry* when 
it stated that “[t]he analysis of the element of foreseeability with regard to 
negligence lawsuits requires the making of subtle distinctions between the 
application of the concept as to duty and the application as to proximate 
cause.”\(^{227}\) In a similar case involving a utility pole and a car accident, the 
court listed subjective factors to consider when discussing the foreseeability

\(^{217}\) *Goldsberry*, 672 N.E.2d at 478 (citing *Webb*, 575 N.E.2d at 995).
\(^{218}\) *Id.*
\(^{219}\) *Id.* at 478-79.
\(^{220}\) *Id.* at 479.
\(^{221}\) *Id.*
\(^{222}\) Boehm, *supra* note 207, at 11.
\(^{223}\) See *Kelley I*, *supra* note 11, at 53 (noting that the treatise authors, Prosser and Keeton, 
criticize Harper’s and James’s position regarding duty and proximate cause and vice-a-versa).
\(^{224}\) See Cardi *I*, *supra* note 10, at 933-37.
\(^{225}\) *Goldsberry*, 672 N.E.2d at 478.
\(^{226}\) See, e.g., Snyder v. LTG Lufttechnische GmbH, 955 S.W.2d 252, 256 n.6 (Tenn. 1997).
“Proximate or legal cause is a policy decision made by the . . . courts to deny liability for otherwise 
actionable conduct based on considerations of logic, common sense, policy, precedent and ‘our 
more or less inadequately expressed ideas of what justice demands or of what is administratively 
possible and convenient.’” *Id.* (quoting Bain v. Wells, 936 S.W.2d 618, 625 (Tenn. 1997)).
component of duty. The court noted that “enumeration of such factors might well transgress upon the Goldsberry duty analysis ‘without regard to the facts of the actual occurrence,’” but it refused “to rewrite the holding of Webb to alter the duty-foreseeability portion thereof.”

There is no clear answer from the Indiana Supreme Court. This is especially true because the court relies on the existence of specific facts to determine the existence of both duty and proximate cause. In Lane v. St. Joseph’s Regional Medical Center, the Indiana appellate court addressed the difference between duty and proximate cause in a different factual circumstance. The Lane court noted that “[t]he correct application of the foreseeability element is an issue which has plagued this court since the announcement of the three-factor test in Webb v. Jarvis.” The Lane court also acknowledged that the decision in Hammock might have gone against the holding in Goldsberry, but it did so because of unique circumstances in that case.

Florida court decisions illustrate another example of confusion between duty and proximate cause. In McCain v. FPL, the Florida Supreme Court

228. In this case, Hammock was driving his car and hit a utility pole. Some time later, and about two miles away, the power went out at the Red Gold plant for almost five hours. Id. at 496. Red Gold grew tomatoes, and as a result of the power outage, experienced serious losses, in excess of $44,000. Id. at 496-97. The appellate court reversed the trial court and determined that there was no duty owed to Red Gold. Id. at 506. Some of the factors that the court considered when determining the foreseeability component of duty were: (1) the distance between the accident and the business; (2) the population of the area surrounding the accident; (3) how the electrical service is set up; and (4) which electric pole was struck. Id. at 501.

229. Id. (internal citations omitted).

230. In Estate of Heck ex rel. Heck v. Stoffer, 786 N.E.2d 265 (Ind. 2003), the Indiana Supreme Court did not directly address the confusion surrounding the Webb test. However, it may have added to it. Timothy Stoffer shot and killed Officer Heck with a handgun owned by Timothy’s parents. Id. at 266. The estate of Officer Heck sued Timothy Stoffer’s family (he had been killed during the incident) who were “gun owners and parents of a drug-addicted felon fleeing from police, to whom they gave unrestricted access to their home where they kept a handgun.” Id. at 266-68. When determining, pursuant to Webb, the foreseeability component of duty, the court considered very specific facts, including the following: (1) that Timothy’s father thought Timothy had a death wish when he took the gun; (2) that Timothy had stolen from his parents before; (3) that Timothy had a history of resisting arrest; (4) that Timothy’s father was aware that Timothy would most likely flee to avoid an upcoming hearing; (5) that Timothy had access to their home; and (6) that the Stoffers exercised great care to protect their children and valuables when Timothy was released from prison, but not the gun. Id. at 269. Following the Goldsberry analysis, this evaluation should have been done in the proximate cause element and not under duty. Yet, the Heck court said that it “declined[d] to take a narrow view of Webb’s foreseeability of harm prong and determine[d] that this factor weighs in favor of the establishment of a duty.” Id


232. Lane, 817 N.E.2d at 271.

explained that the difference between foreseeability in the two elements is the general zone-of-risk-foreseeability analysis that is used to determine duty, in contrast to the highly specific foreseeability analysis used to determine proximate cause. 236 However, Florida courts tend to erode the usefulness of the Florida Supreme Court’s distinction under another established theory of proximate cause. 237 In proximate cause the specific harm that occurs need not be foreseeable; the required foreseeability is limited to the general type of harm that occurred. 238 This obliterates the specific/general distinction that the Florida Supreme Court set forth. 239 Likewise, most scholarship avoids pinpointing a place where generality ends and specificity begins. 240

The legal arguments from the Anonymous case illustrate the illusory nature of this distinction. 241 As previously mentioned, in Anonymous, the defendant argued that it had no duty to guard against unforeseeable acts. 242 The defendant also argued that even if it owed a duty to the plaintiff to protect him from the independent criminal acts of a third person, there was no proximate

235. See McCain, 593 So. 2d at 503 (attempting to distinguish the foreseeability analyses of duty and proximate cause).

236. See id.

‘Where a defendant’s conduct creates a foreseeable zone of risk, the law generally will recognize a duty placed upon defendant either to lessen the risk or see that sufficient precautions are taken to protect others from the harm that the risk poses.’

Id. (quoting Kaisner v. Kolb, 543 So. 2d 732, 735 (Fla. 1989)) (emphasis in original).

Thus, as the risk grows greater, so does the duty, because the risk to be perceived defines the duty that must be undertaken. The statute books and case law, in other words, are not required to catalog and expressly proscribe every conceivable risk in order for it to give rise to a duty of care. Rather, each defendant who creates a risk is required to exercise prudent foresight whenever others may be injured as a result. This requirement of reasonable, general foresight is the core of the duty element. For these same reasons, duty exists as a matter of law and is not a factual question for the jury to decide: Duty is the standard of conduct given to the jury for gauging the defendant’s factual conduct. As a corollary, the trial and appellate courts cannot find a lack of duty if a foreseeable zone of risk more likely than not was created by the defendant.

On the question of proximate causation, the legal concept of foreseeability also is crucial, but in a different way. In this context, foreseeability is concerned with the specific, narrow factual details of the case, not with the broader zone of risk the defendant created.

Id.

237. Id. (“[I]t is immaterial that the defendant could not foresee the precise manner in which the injury occurred or its exact extent.”) (emphasis in original).

238. Id.

239. Id.

240. See, e.g., Cardi I, supra note 10, at 971-72 (brilliantly acknowledging that this approach is unworkable and not even attempting to delineate what is appropriately specific in proximate cause and what is appropriately general in duty).

241. See supra Part IV.A.

242. See supra Part IV.B.
cause as a matter of law because the specific type of harm that occurred was not foreseeable. At the hearing, the defendant further explained that the specific harm that occurred was unforeseeable because no one could foresee the precise chain of events as they occurred: “No one could foresee that a breach of duty could lead to fellatio, but “if this was a case where someone went into an unopened door and grabbed his wallet and left, maybe that is foreseeable.” Again, the defendants were unsuccessful because a Florida case on point obliterated the general/specific foreseeability distinction that is often touted as distinguishing duty from proximate cause:

[I]t is not necessary that the tortfeasor be able to foresee the exact nature and extent of the injuries in the precise manner in which they occurred; rather, all that is necessary is that the tortfeasor be able to foresee some injury likely to result in some manner as a consequence of his negligence.

Relying on Florida case law, the plaintiff summed up this argument in his response to the defendant’s motion to dismiss:

Turning to the case at bar, we believe the major, conspicuous deficiency in [ACS’s] foreseeability analysis is its meticulous specificity regarding what is or is not foreseeable. In its brief, it argues that “[ACS had no reason to anticipate or foresee that if [the passenger] was not under constant supervision, he would [engage in fellatio with a Jones Act employee whose Cabin could not be locked.]” The problem with this formulation is that it simply is not necessary “that the tortfeasor be able to foresee the exact nature and extent of the injuries in the precise manner in which they occurred . . . .” Goode v. Walt Disney World Co., 425 So.2d 1151 (Fla. 5th DCA 1982). Thus the exact factual circumstances of the incident need not have been foreseen, only the general type of harm. Here, the incident which gave rise to this suit can be characterized as a physical assault, arguably in the nature of a serious physical attack.

On one hand, courts state that the foreseeability prong of duty is different from that of proximate cause because the former requires only a general-foreseeability analysis, while the latter requires a specific-foreseeability analysis. On the other hand, courts blur this distinction by stating that in order

243. See supra Part IV.B.
244. Case on file with author.
246. Id.
247. See Plaintiff’s Response to Defendant’s Motion to Dismiss at 5, Anonymous Plaintiff v. Anonymous Cruise Line (on file with author) (relying on a verbatim expression of Florida law from Collins, 471 So. 2d at 566).
for proximate cause to exist, the type of harm that occurred to be only the type that was generally foreseeable. Thus, there is no satisfactory judicial method to distinguish between duty and proximate cause in negligence.\textsuperscript{248}

B. Scholarly Missteps

One scholarly approach to distinguishing duty from proximate cause is to state that duties are categorical and transcend single cases, while issues of proximate cause are case-specific.\textsuperscript{249} The following passage is particularly illustrative of the attempted distinction: “the decision to provide a general immunity for parents who negligently supervise their children is purely [a duty question] for the court and that decision does not turn on the facts of individual cases.”\textsuperscript{250} According to Dobbs, proximate cause is distinguishable from duty because proximate cause requires a foreseeability-based analysis that determines whether the type of harm that occurred was foreseeable.\textsuperscript{251}

Based on the cases previously described in Part V.A., this distinction regarding foreseeability is unworkable. The famous “B<PL” theory of duty demonstrates that foreseeability is an element of duty, not proximate cause.\textsuperscript{252} Undoubtedly, the second and third prongs of the “B<PL” test (the gravity of the potential harm and the likelihood of its occurrence) require foreseeing what the harm and type of harm would be in determining whether there is a duty to act in a certain way.\textsuperscript{253}

Additionally, the concept of duties being categorical and devoid of factual considerations is far from reality in negligence law today, as further discussion will show. Many of the old categorical duty rules are no longer valid, nor are they even possible to articulate.\textsuperscript{254} For example, the law traditionally assigned landowners specific duties to protect different classes of people from harm on their land.\textsuperscript{255} However, decades ago, courts realized that the artificial nature

\textsuperscript{248} See supra Part V.A.

\textsuperscript{249} See, e.g., Dobbs, supra note 69, at 451 (explaining Justice Cardozo’s conception of the Palsgraf problem as a duty problem while Justice Andrews believed the issue was one of proximate cause).

\textsuperscript{250} Id. at 450.

\textsuperscript{251} Id. at 448-50.

\textsuperscript{252} See supra Part IV.B. (discussing the Carroll Towing formula and its significant overlap with modern proximate cause theory).

\textsuperscript{253} See supra Part IV.B.

\textsuperscript{254} See, e.g., Adams v. City of Freemont, 80 Cal. Rptr. 2d 196, 236 n.8 (Ct. App. 1998) (expressing the view that there is little valid distinction between acts (malfeasance) and failures to act (non-feasance) in terms of establishing duty in negligence cases).

\textsuperscript{255} See, e.g., Koenig v. Koenig, 766 N.W.2d 635, 638 (Iowa 2009) (“The premises liability trichotomy, which distinguishes between invitees, licensees, and trespassers, finds its roots in the English common law.”); Nelson v. Freeland, 507 S.E.2d 882, 884 (N.C. 1998) (“In a traditional common-law premises-liability action . . . there is a descending degree of duty owed by a landowner based upon the plaintiff’s status.”).
of the broad categorical duties was an unworkable relic of simple feudal society that could not be applied in today’s modern industrialized society.  

Similarly, there is a traditional, categorical “common-law rule, as set forth [in the Restatement (Second) of Torts] that there is no duty to act affirmatively for another’s aid or protection absent some ‘special relation’ which justifies the imposition of a duty.” However, decisions that require courts to engage in complex analyses to ascertain if a duty exists are eroding even this stalwart of tort law. 

In fact, contrary to scholarly attempts to confine considerations of foreseeability to proximate cause and not duty, modern negligence cases hold that the “[f]oreseeability of the risk of harm is the foundational element in the determination of whether a duty exists.” In addition, these cases also tend to find that proximate cause is closely related to duty. 

Other categorical duties found in common law have also been obliterated, such as the concept that there is no duty owed to anyone not in privity of contract with the defendant. Similarly, the categorical common law rule

---


The distinctions which the common law draws between licensee and invitee were inherited from a culture deeply rooted to the land, a culture which traced many of its standards to a heritage of feudalism. In an effort to do justice in an industrialized urban society, with its complex economic and individual relationships, modern common-law courts have found it necessary to formulate increasingly subtle verbal refinements, to create subclassifications among traditional common-law categories, and to delineate fine gradations in the standards of care which the landowner owes to each. Yet even within a single jurisdiction, the classifications and subclassifications bred by the common law have produced confusion and conflict. As new distinctions have been spawned, older ones have become obscured. Through this semantic morass the common law has moved, unevenly and with hesitation, towards ‘imposing on owners and occupiers a single duty of reasonable care in all the circumstances.’  

Id. (citations omitted).


In determining whether a duty is to be imposed, courts must engage in a rather complex analysis that weighs and balances several, related factors, including the nature of the underlying risk of harm, that is, its foreseeability and severity, the opportunity and ability to exercise care to prevent the harm, the comparative interests of, and the relationships between or among, the parties, and, ultimately, based on considerations of public policy and fairness, the societal interest in the proposed solution.  

Id. (citation omitted).

259. Id.

260. Id. at 935.

261. See, e.g., MacPherson v. Buick Motor Co., 111 N.E. 1050, 1053 (N.Y. 1960) (“We have put aside the notion that the duty to safeguard life and limb, when the consequences of negligence may be foreseen, grows out of contract and nothing else.”).
that a failure to act, or nonfeasance, cannot be a breach of duty has been eroded.\footnote{262}

Perhaps the most recent scholarly piece arguing for the retention of categorical duties is Goldberg and Zipursky’s article.\footnote{263} In an earlier article, these authors stated that “without the concept of duty, a cogent general description of the structure of negligence would not be possible.”\footnote{264} Therefore, Goldberg and Zipursky appear to be operating on the assumption that negligence, as a four-element tort, is a workable field of liability that can be described.\footnote{265} But, this article is not bound by any such assumption.

In arguing for the relevance and vitality of duty, Goldberg and Zipursky state the following:

\begin{quote}
[D]uty possesses four distinct meanings, and the challenge of the Restatement project is to separate the “primary,” affirmative, obligatory sense of duty from its alternative meanings. In addition to duty in its primary or “obligation” sense, duty also functions as a “nexus requirement,” masquerades as “breach-as-a-matter-of-law,” and, finally, operates as an “exemption from the operation of negligence law,” or “no-duty.”\footnote{266}
\end{quote}

The first distinct area of duty that Goldberg and Zipursky argue for is the existence of “[d]uty in its primary sense.”\footnote{267} This “duty inquiry concerns ‘whether defendant and plaintiff stood in such a relationship to one another that the law imposed upon defendant an obligation of reasonable conduct for

\begin{quote}
The law (if it has a will) never intended that nonfeasance and misfeasance should be distinguished, since in logic inaction is one form of action. Rather, it planned that one should be liable for the harm he causes by acting in violation of a duty, or by failing to act when action was his duty. It is unfortunate that negligence law, when it first arose in the nineteenth century, should have somehow acquired, adopted, and endorsed the false differentiation between nonfeasance and misfeasance first created four hundred years earlier.
\end{quote}

\textit{Id. See also} Adams v. City of Freemont, 80 Cal. Rptr. 2d 196, 236 n.8 (Ct. App. 1998) (expressing the view that there is little valid distinction between acts and failures to act in terms of establishing duty in negligence cases).

\footnote{263. See generally Goldberg & Zipursky, supra note 11; Kelley II, supra note 11, at 1047 (describing the Zipursky article as “excellent”).


265. Id. at 663 (“Our central critical goal in this Article . . . is to demonstrate that the self-conscious and carefully constructed three-element views put forward in the drafts of the Restatement (Third) do not capture the law, and hence that a sound restatement of negligence cannot simply neglect duty.”).

266. See Kolar, supra note 11, at 250.

Professor Howarth, however, points out that this relational concept of duty is problematic and makes very little sense:

Most defendants have had no relationship at all with their claimants. The driver of a car who carelessly runs down a pedestrian is hardly likely to have met the pedestrian on a previous occasion, and is certainly not allowed to escape liability by saying, ‘Never seen him before in my life’. To describe the act of hitting someone as a ‘relationship’ is simply perverse. The usual explanation is that the parties are ‘fellow road users’. But that is a relationship hardly different from ‘fellow inhabitant of the country’ or even ‘fellow human being’.

Goldberg and Zipursky use the case of *Harper v. Herman* to give credibility to the existence of a primary duty. The authors claim that this case is a robust example of the need for duty in negligence. *Harper v. Herman* involved a suit by an adult plaintiff who was a social guest on defendant’s yacht and sustained injuries when the plaintiff-guest dove into water that was too shallow. Goldberg and Zipursky argue that the court’s decision in refusing to find a duty is an example of the existence of primary duties. However, Professor Kelley challenges this type of argument as conceptual formalism that does not resolve the problem of whether duty is a necessary element of negligence.
Applying modern definitions of proximate cause to the Harper case only renders further attempts to distinguish between duty and proximate cause futile.\(^\text{277}\) The Harper court, in denying the existence of liability, stated:

“There are many dangers, such as those of fire and water, . . . which under ordinary conditions may reasonably be expected to be fully understood and appreciated by any child . . . .” If a child is expected to understand the inherent dangers of water, so should a 20-year-old adult.\(^\text{278}\)

Hence, the same result could have been achieved if the court had concluded that there was a duty to warn of the dangers of shallow water; but, that there was no proximate cause because it was unforeseeable that an adult would jump into water he knew nothing about.\(^\text{279}\) Therefore, the case fails to demonstrate the distinction between duty and proximate cause.\(^\text{280}\)

Another interesting aspect of the Zipursky article is that it implicitly supports Palsgraf’s view that duty is relational and does not exist as a general obligation to the world.\(^\text{281}\) However, Cardozo’s opinion of duty in Palsgraf is a poor case to use as support for a relational view of duty because Mrs. Palsgraf was a passenger of the railroad company.\(^\text{282}\) “From the moment that she bought her ticket the defendant [as a common carrier] did in fact owe her a duty of the highest care, one of the most stringent known to the law.”\(^\text{283}\)

Supporters of the idea that duty is relational espouse the view that for each class of cases, there is a duty that can be discerned and applied.\(^\text{284}\) On the other side of the coin, there are scholars who agree with Andrews’s viewpoint that duty is a general obligation imposed on society to behave reasonably by avoiding conduct that results in harm.\(^\text{285}\) This begs the question of whether the overlap between unreasonable conduct and proximate cause is reduced by stating duties broadly rather than specifically. The Restatement (Third)

\(^{277}\) See, e.g., Snyder v. LTG Lufttechnische GmbH, 955 S.W.2d 252, 256 n.6 (Tenn. Ct. App. 1997).

Proximate or legal cause is a policy decision made by the . . . courts to deny liability for otherwise actionable conduct based on considerations of logic, common sense, policy, precedent and “our more or less inadequately expressed ideas of what justice demands or of what is administratively possible and convenient.” Id. (quoting Bain v. Wells, 936 S.W.2d 618, 625 (Tenn. 1997)).

\(^{278}\) Harper v. Herman, 499 N.W.2d 472, 475 (Minn. 1993) (quoting RESTATEMENT (SECOND) OF TORTS § 339 cmt. j (1965)).

\(^{279}\) Id.

\(^{280}\) Id.

\(^{281}\) See Goldberg & Zipursky, supra note 11, at 708-709 (describing Andrews’ view of duty in Palsgraf as an obligation owed to the world in general as “fallacious”).


\(^{283}\) See Palsgraf Revisited, supra note 94, at 7.

\(^{284}\) See, e.g., Howarth, supra note 11, at 449.

\(^{285}\) Id.
suggested that duties should be broadly stated. As the next section indicates, this does not solve the overlapping foreseeability problem.

VI. THE THIRD RESTATEMENT’S PERPETUATION OF OVERLAPPING FORESEEABILITY

A. The Formulation of Unreasonable Conduct

The Restatement (Third) defines negligence as follows: “A person acts negligently if the person does not exercise reasonable care under all the circumstances.” It also clarifies that reasonable care is determined under the standard of the “reasonably prudent person” and explains that “[a]n actor ordinarily has a duty to exercise reasonable care when the actor’s conduct creates a risk of physical harm.”

In order to determine whether reasonable care was exercised, the Restatement adopts Learned Hand’s “B<PL” formula from Carroll Towing. It also uses this formula to discern whether a duty exists. The Restatement explains that, in considering whether the person’s conduct lacks reasonable care, the following factors must be considered: “the foreseeable likelihood that the person’s conduct will result in harm, the foreseeable severity of any harm that may ensue, and the burden of precautions to eliminate or reduce the risk of harm.”

The Restatement also settles the Palsgraf dichotomy by agreeing with Judge Andrews in the recognition of one general duty of care as opposed to many duties of care. This means that actors are generally charged with the duty to use ordinary care or to behave reasonably, which does not require that the existence of a duty be considered on a case-by-case basis. Thus, the

286. RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 7 reporters’ notes to cmt. a (2010) [hereinafter RESTATEMENT (THIRD)] (referring “to the limited number of situations to which the duty doctrine serves to limit or qualify what would otherwise be the broad scope of negligence liability.”).

287. Id. § 3.

288. Id. § 3 cmt. a.

289. Id. § 7.

290. See White, supra note 11, at 103-05 (explaining the B<PL formula from Carroll Towing).

291. RESTATEMENT (THIRD) § 3.

292. Id. (emphasis added).

293. Id. § 7 cmt. a (stating that in most cases courts need not refer to duty on a case-by-case basis).

294. Id. § 6 cmt. f (“[A]n actor ordinarily has a duty to exercise reasonable care. That is equivalent to saying that an actor is subject to liability for negligent conduct that causes physical harm. Thus in cases involving physical harm, courts ordinarily need not concern themselves with the existence or content of this ordinary duty.”). See also Cardi II, supra note 11, at 779-81 (2005) (claiming the Restatement (Third) disagrees with Cardozo and “cleanses” duty of the concept of the foreseeable plaintiff).
Restatement (Third), unlike Cardozo in *Palsgraf*, attempts to remove the factually-laden issue of whether the plaintiff was foreseeable from the duty prong\(^{295}\) to its equivalent of proximate cause, which is otherwise known as scope of liability.\(^{296}\)

Under the Restatement formulation, once physical harm is likely to result from the defendant’s conduct, the defendant is charged with the duty to behave with ordinary care.\(^{297}\) Procedurally, the defendant could terminate the action by demonstrating that the case falls under one of the enunciated categories where no duty rules govern or where “the court determines that the ordinary duty of reasonable care is inapplicable.”\(^{298}\) As a result of this procedural formulation, some commentators suggest that the Restatement (Third) removes duty as an element of a negligence cause of action.\(^{299}\)

According to the Restatement (Third), a finding of no duty\(^{300}\) or a modification of the general duty of care\(^{301}\) is required in some cases, but these decisions should be made as a matter of law and not as a factually-based foreseeable analysis.\(^{302}\) The Restatement discourages the wide-scale application of foreseeability and favors the “more transparent explanations of the reasons for a no-duty ruling and to protect the traditional function of the jury as factfinder.”\(^{303}\) Exceptions to the duty to behave reasonably exist in a particular category of cases that are governed by broad rules of no duty prescribed by the common law.\(^{304}\)

At first blush, it appears that the Restatement eliminates the confusing dual application of foreseeability in determining whether conduct was reasonable

\(^{295}\) Restatement (Third) § 6 reporters’ note to cmt. f.
\(^{296}\) Id. § 29 cmt. b.
\(^{297}\) Id. § 3.
\(^{298}\) Id. § 6.
\(^{299}\) See, e.g., Daniel More, *Negligence in the Law (Part II) - The Boundaries of Negligence*, 4 Theoretical Inquiries L. 339, 342 (2003) (explaining that duty, according to the Restatement (Third), is no longer an element of a negligence cause of action, but “it refers to the failure of a defendant who is already presumed or found to have committed the tort of negligence to obtain a judicial exemption from the liability that a negligent actor ordinarily incurs”).

\(^{300}\) Restatement (Third) § 7(b).
\(^{301}\) See, e.g., id. § 9 (describing emergency situations).
\(^{220}\) Id. § 7 cmt. j.
\(^{303}\) Id.

A no-duty ruling represents a determination, a purely legal question, that no liability should be imposed on actors in a category of cases. Such a ruling should be explained and justified based on articulated policies or principles that justify exempting these actors from liability or modifying the ordinary duty of reasonable care. These reasons of policy and principle do not depend on the foreseeability of harm based on the specific facts of a case. They should be articulated directly without obscuring references to foreseeability.

\(^{304}\) Id. § 7 cmt. b.
and in determining the existence of proximate cause. In reality, the Restatement only shifts the problem of foreseeability from the element of duty to that of breach.

Courts do appropriately rule that the defendant has not breached a duty of reasonable care when reasonable minds cannot differ on that question . . . . These determinations are based on the specific facts of the case, are applicable only to that case, and are appropriately cognizant of the role of the jury in factual determinations. A lack of foreseeable risk in a specific case may be a basis for a no-breath determination, but such a ruling is not a no-duty determination. Rather, it is a determination that no reasonable person could find that the defendant has breached the duty of reasonable care.

This does not solve the problems created by the overlapping use of foreseeability in determining unreasonable conduct, proximate cause, or scope of liability, as the Restatement (Third) defines proximate cause.

Recognizing the existence of a general duty of care means that litigation in negligence will focus not on the existence of the duty, but rather on whether there was a breach of the duty. The shift of foreseeability from duty to breach is irrelevant and formulaic because a foreseeability analysis is required in deciding whether conduct is unreasonable.

For example, in a grocery-store slip-and-fall case the Restatement (Third) eliminates the need for pleading a specific duty, such as: “grocer defendant owed customer plaintiff a duty to ensure that banana peels were not left on the floor because a general duty of care to act reasonably under all circumstances exists.” However, in order to determine whether there was unreasonable conduct or a negligent act, we must still apply the foreseeability based “B<PL” factors from Carroll Towing. Because the Restatement (Third) redefines what was formerly duty and breach as unreasonable conduct, to ascertain whether there was unreasonable conduct, “the foreseeable likelihood that the person’s

305. Id. § 7 cmt. j (stating that the “reasons of policy and principle do not depend on the foreseeability of harm based on the specific facts of a case”).
306. See supra Part IV.B. (examining Professor Cardi’s work and discussing: (a) the overlap between breach and proximate cause when duties are stated broadly; and (b) the overlap between duty and proximate cause when duties are stated specifically). See also Restatement (Third) § 7 cmt. j.
307. Restatement (Third) § 7 cmt. j.
308. Id. § 29.
309. See supra Part IV (demonstrating that duty, specific or general, determines whether foreseeability is a necessary element of duty or breach, respectively, but that foreseeability is always a necessary element of determining whether unreasonable conduct existed).
310. Id.
311. Restatement (Third) § 3.
conduct will result in harm [and] the foreseeable severity of any harm that may ensue . . .” must be considered.\footnote{312. Id.}

This fact-intensive analysis is perfectly illustrated by the Restatement’s adoption of its formulation of unreasonable conduct from the foreseeability-based analysis engaged in by Learned Hand in \textit{Carroll Towing} to determine if unreasonable conduct occurred:\footnote{313. United States v. Carroll Towing Co., 159 F.2d 169, 173-74 (2d Cir. 1947).}

\begin{quote}
Applied to the situation at bar, the likelihood that a barge will break from her fasts and the damage she will do, vary with the place and time; for example, if a storm threatens, the danger is greater; so it is, if she is in a crowded harbor where moored barges are constantly being shifted about. On the other hand, the barge must not be the bargee’s prison, even though he lives aboard; he must go ashore at times. We need not say whether, even in such crowded waters as New York Harbor a bargee must be aboard at night at all; it may be that the custom is otherwise, as Ward, J., supposed in ‘The Kathryn B. Guinan,’ supra; and that, if so, the situation is one where custom should control. We leave that question open; but we hold that it is not in all cases a sufficient answer to a bargee’s absence without excuse, during working hours, that he has properly made fast his barge to a pier, when he leaves her. In the case at bar the bargee left at five o’clock in the afternoon of January 3rd, and the flotilla broke away at about two o’clock in the afternoon of the following day, twenty-one hours afterwards. The bargee had been away all the time, and we hold that his fabricated story was affirmative evidence that he had no excuse for his absence. At the locus in quo- especially during the short January days and in the full tide of war activity- barges were being constantly ‘drilled’ in and out. Certainly it was not beyond reasonable expectation that, with the inevitable haste and bustle, the work might not be done with adequate care. In such circumstances we hold- and it is all that we do hold- that it was a fair requirement that the Conners Company should have a bargee aboard (unless he had some excuse for his absence), during the working hours of daylight.\footnote{314. Id.}

So, even under the Restatement’s formulation of negligence, determining whether conduct is unreasonable or negligent requires an application of foreseeability-based tests.\footnote{315. \textsc{Restatement (Third)} § 3 (adopting the B<PL test from \textit{Carroll Towing}).}
B. The Formulation of Liability

The Restatement attempts to clarify proximate cause by changing the name “proximate cause” to scope of liability. This is because the term proximate cause is misleading, often applied incorrectly and conflated with cause in fact. Additionally, the term “proximate” implies that there is but one cause—the cause nearest in time or geography to the plaintiff’s harm. As neither implication is correct, the Restatement uses scope-of-risk terminology or a “risk standard” to determine “whether a negligent actor, who was the cause-in-fact of the harm, should be liable for the harm caused.”

The Restatement uses “scope of liability” language instead of “proximate cause” to ensure that the proper analysis is performed when determining liability of a negligent actor. The rationale behind the Restatement (Third)’s change in language is that it encourages focusing “on the appropriate context.” Changing from foreseeability to scope of risk adds nothing to the analysis because, as the Restatement admits, “a foreseeability standard, properly explained,” is the same as the scope-of-risk standard and, “[w]hen properly understood and framed, the foreseeability standard is congruent with the risk standard . . . .” Later, the Restatement admits that changing the name of the analysis from “whether the harm was foreseeable” to “whether it was within the scope of risk” is of little practical utility in resolving the appropriate use of foreseeability in negligence because courts are still likely to end up with the same results.

316. Id. § 29 cmt. b.
317. Id.
318. Id. § 29 cmt. d.
319. Id. § 29 cmt. j.
320. Id.
321. RESTATEMENT (THIRD) § 29 cmt. j (explaining that “[t]he risk standard focuses on the appropriate context, although a foreseeability standard, properly explained, could do this also). See also id. § 29 cmt. e.
322. Id. § 29 cmt. e.

As explained in Comment j, the foreseeability test many courts employ in negligence cases for proximate cause is quite compatible with the standard in this Section. In addition to jurisdictions employing foreseeability, a number of jurisdictions that purport to use a different standard for proximate cause in the end reach results that are entirely consistent with the risk standard of this Section and frequently employ language and analysis virtually identical to the risk standard. See DAN B. DOBBS, THE LAW OF TORTS § 185, at 459–460 (2000) (“The upshot is that less than a handful of courts purport to apply the kind of strict liability entailed in imposing liability beyond the risk the defendant unreasonably created, and even those may covertly or sporadically
Additionally, the illustrations used by the Restatement (Third) to demonstrate the clarity of scope of risk are less than realistic. For example, the Restatement states that no liability exists where a hunter stops at his friend’s house and hands the friend’s nine-year-old daughter a loaded gun, which is neither heavy nor unwieldy, and the child drops the shotgun on her toe breaking it. The Restatement explains that no liability exists because the only harm within the scope of the hunter’s liability is harm caused by a bullet discharging.

Central to this conclusion is the Restatement’s propagation of a well-established rule of liability in negligence: “an actor should be held liable only for the harm that was among the potential harms-the risks-that made the actor’s conduct tortious.” In other words, “an actor is liable only if the harm that was caused was the type of harm that made his conduct unreasonable.”

However, from a realistic perspective, this illustration is not absolute. Before concluding that the only harm within the scope of the risk is that of a discharged bullet, the child’s lawyer would attempt to discover more facts, perhaps including the following: the experience of the nine-year-old with holding guns; the hunter’s special knowledge of the nine-year-old’s strength because she is his friend’s daughter; the size of the nine-year-old relative to a loaded shotgun, and so forth. This heavy, factual, foreseeability-based analysis, which the Restatement is determined to confine only to scope-of-liability issues, would also take place in determining whether the conduct was reasonable. Because the Restatement’s formulation involves only a general duty to act reasonably in all circumstances, the child’s lawyer could plead this duty exactly as described by the Restatement (Third). If any of the facts mentioned above are discovered, the lawyer could use them to assert that the

introduce foreseeability into the liability equation.”); ROBERT E. KEETON, LEGAL CAUSE IN THE LAW OF TORTS 41 (1963) (“[W]hether consciously or subconsciously, a concept of risk is almost exclusively, if not exactly so, the determinant of directness and concurrence.”).

Id. § 29 reporters’ note to cmt. d. See also Stapleton, supra note 11, at 992-94 (arguing that the scope of risk and foreseeability rules of proximate cause are identical).

323. See, e.g., RESTATEMENT (THIRD) § 29 cmt. d, illus. 3.
324. Id.
325. Id.
326. Id. § 29 cmt. d.
327. Id. § 29 cmt. j (“Thus when scope of liability arises in a negligence case, the risks that make an actor negligent are limited to foreseeable ones, and the factfinder must determine whether the type of harm that occurred is among those reasonably foreseeable potential harms that made the actor’s conduct negligent.”).
328. See infra Part VI.C. (explaining that the Restatement (Third) shifts the factually based foreseeability analysis from duty to breach).
329. RESTATEMENT (THIRD) § 3.
hunter breached his duty of reasonable care, which according to the Restatement, is better left as a foreseeability-laden analysis for the jury.330

Because the Restatement’s definition of unreasonable conduct is taken unchanged from Carroll Towing, which employs a foreseeability analysis largely indistinct from what the Restatement proposes should only be done in scope-of-liability considerations, it expressly states, albeit in a comment, that this type of analysis is precisely what it envisions taking place in the breach stage.331

C. The Continued Existence of Overlapping Foreseeability

The Restatement fails to clarify the appropriate roles of foreseeability in the supposedly separate inquiries of whether conduct is unreasonable and whether liability should be imposed for unreasonable conduct.332 An unresolved issue remains—to distinguish between what needs to be foreseen when determining whether conduct was unreasonable and what should be considered when analyzing the scope of the risk or type of harm.333

No test or scholarship clearly illustrates what the difference should be.334 In this regard, the unanswered question is whether there is a difference in the appropriate generality of foreseeability at the unreasonable-conduct stage and at the scope-of-liability stage. The Restatement likewise provides no answer to this central question; it merely glosses over it in a comment.335

330. Id. § 7 cmt. j.
331. Id.
332. Id.
333. See supra Part V.A. (discussing unsuccessful judicial attempts to resolve this). See also RESTATEMENT (THIRD) § 29 reporters’ note to cmt. d.
334. RESTATEMENT (THIRD) § 29 reporters’ note to cmt. d.
335. Id. § 29 cmt. i.
Even though the general-duty statement of the Restatement removes foreseeability from the duty prong so that courts no longer are engaged in using foreseeability to determine if a duty existed, like Cardozo did in *Palsgraf*, the jury must consider foreseeability in determining breach, thereby resulting in the use of foreseeability in determining both unreasonable conduct and scope of liability.

Because the Restatement (Third) does not solve the problem of overlapping foreseeability, and neither it nor any of the scholarship adequately describes the difference in the foreseeability tests to be applied in determining unreasonable conduct or scope of liability, the next section proposes a bifurcated negligence test that is realistic, workable, and does not “overload foreseeability.”

VII. A BIFURCATED SOLUTION TO THE PROBLEM OF OVERLAPPING FORESEEABILITY

In order to reform the law of negligence, three propositions must be accepted as true. First, proximate cause or scope of liability is sometimes based on foreseeability. Second, proximate cause focuses on a policy decision to deny liability for otherwise actionable conduct based on considerations of logic, common sense, policy, precedent, and “our more or less inadequately expressed ideas of what justice demands or of what is administratively possible and convenient.” Third, and most controversial, is that clear, concise, and somewhat precise statements of duties exist today for certain factual situations. The Restatement (Third), which recognizes a number of established affirmative duties, supports this proposition.

Professor Kelley plants the seed of modern clear and concise duties in his brilliant suggestion that “[i]t is possible . . . to restate modern negligence in general in the risks that it addresses. Thus, greater uncertainty and difficulty occur in negligence cases in determining whether the harm that resulted arose from the risks that made the actor’s conduct unreasonable.”

336. See Cardi II, *supra* note 11, at 794 (“[T]he most significant effect of the proposed Restatement’s conceptual restructuring is that by casting foreseeability out of duty[, foreseeability] is no longer a presumed matter for the judge . . . .”).
337. *Restatement (Third)* § 7 cmt. j.
338. *Id*.
340. See *supra* note 75 (demonstrating the widespread application of foreseeability in proximate cause analysis).
341. See, e.g., Schwartz et al., *supra* note 1, at 294 (citing Snyder v. LTG Lufttechnische GmbH, 955 S.W.2d 252, 256 (Tenn. 1997)).
342. See, e.g., Goldberg & Zipursky, *supra* note 11, at 698-718 (showing different examples of specific duties in various fact situations).
343. See, e.g., *Restatement (Third)* §§ 37-44 (Proposed Final Draft 2005) (recognizing the existence of at least seven affirmative duties that exist even when the actor’s conduct does not create a risk of physical harm).
pre-Holmsian terms. Pre-Holmsian negligence recognized that the general reasonable person standard was not an open-ended invitation to find unreasonable conduct, but it was tied “to the specific preexisting customs, conventions, and coordinating practices of the community.”

“There is not circularity or judicial ipse dixit here because the courts reason from pre-judicial community-defined obligations, based on the accepted coordination norms of the community, to a conclusion about legally redressing a wrong understood as a breach of that community-defined obligation.”

A large number of these “pre-judicial community-defined customs” exist in today’s society. Examples might include emergency room protocols, seatbelt laws, speed limits, floor-cleaning protocols in grocery stores, antilock brakes, school-bus-stop rules, and deference to emergency vehicles. The list goes on and on. These standards reflect what society thinks is the most beneficial or reasonable behavior in the same way that customs of the realm denoted efficient societal behavior as “community-defined obligations, based on the accepted coordination norms of the community . . . .”

344. See Kelley II, supra note 11, at 1066.
345. Id. at 1058-59.
346. Id. at 1059-60.
348. It would be a fascinating study to see how many of these protocols and standards which guide our conduct were established as a way to reduce negligence litigation in the respective fields, inspired by repeated judicial determinations of what was or was not reasonable in the particular field. If they were born out of negligence litigation, it would further buttress the suggestion that they reflect what a reasonable person would do. Perhaps, this is the subject of a second law review piece on the matter.
349. See Kelley II, supra note 11, at 1060. See also William B. Schwartz & Neil K. Komesar, Doctors, Damages and Deterrence: An Economic View of Medical Malpractice 1282-1283 (Rand 1978) (explaining that the medical malpractice lawsuits result in medical providers acting in a way to prevent their liability in the future).

The negligence system makes a great deal more sense if it is understood primarily as a means to deter careless behavior rather than to compensate its victims. By finding fault and assessing damages against the negligent provider, the system sends all providers a signal that discourages future carelessness and reduces further damages . . . . [Negligence in other words] signals potentially negligent people that it will cost them more to be careless than to invest in an appropriate level of prevention.
Nothing prevents us from adopting these standards based on conduct that a reasonable person would follow, and these standards could provide the basis for a pre-Holmesian restatement of negligence. In situations where they apply, these standards would reflect the reasonable person standard, which would be “equivalent to a pleaded custom of the realm.”

Imagine a suit where it is alleged that a plaintiff’s fall in a grocery store caused him harm that resulted from the grocery store departing from an established floor-cleaning protocol. In this situation, the judge could ascertain whether a duty existed as a matter of law based on the existence of the protocol. After discovery, it would be up to the jury, or in rare cases, the court, to decide breach based on facts demonstrating whether the defendant departed from the standard. This would maintain the important dichotomy between the judge and jury in negligence actions and re-establish the equilibrium that scholarship claims the Restatement disturbs.

Where an established affirmative duty applies, negligence analysis would remain a five-pronged analysis with the same elements of duty, breach, cause in fact, proximate cause, and damages. The judge would determine duty, and the jury would determine breach, as described above. Note that there is no foreseeability rationale employed in the analysis up to this point. Cause-in-fact would be decided as it typically is. In deciding whether there is proximate cause, the jury would apply a foreseeability test to determine if the harm that occurred was the kind that is foreseeable from the breach of the specific duty. This would be an application of pure foreseeability (i.e., whether the type of harm is expected from the breach of the duty), which is already currently in use today in some courts.

---

Id. As a result, for example, “the courts have largely accepted “customary standards of medical practice” as a measure of reasonable investment in mishap avoidance.” Id. at 1283.

350. See Kelley II, supra note 11, at 1063 (reflecting on how juries could apply community standards in negligence cases).

351. Id. at 1058.

352. My former personal injury practice continues to stalk my paradigms.

353. See, e.g., McCain v. Fla. Power Corp., 593 So. 2d 500, 503 (Fla. 1992) (stating that duty exists as a matter of law for the judge to decide).

354. See Cardi II, supra note 11, at 794 (“[I]t is settled law that breach and proximate cause are matters to be decided by the jury.”).

355. See Cardi I, supra note 10, at 925 (“[T]he jury must then decide, in the context of breach, whether the defendant’s conduct failed to conform to that standard.”).

356. See Schwartz, et al., supra note 1, at 132 (explaining the elements of negligence and that causation involves a combination of two elements: (1) causation-in-fact; and (2) legal or proximate causation).

357. See supra notes 354-56.

358. See supra note 75 (demonstrating the widespread use of foreseeability as a determinant of liability in negligence.)
Some might argue that the above is indistinguishable from strict liability. However, two points are worth considering. Similarity to strict liability is historically consistent because the source of the “custom of the realm” standard was originally a strict liability standard. However, strict liability tends to apply to entire activities, regardless of whether the defendant is complying with the established standards or customs. In contrast, liability can attach under this proposed negligence formulation only when the defendant fails to comply with an established standard or protocol.

This formulation raises the question, “how should the decision be made as to whether the custom or duty should be applied in a specific case?” We might glean the answer by borrowing from the statutory purpose test of negligence per se. If the standard’s applicability is in dispute, counsel could argue to the judge during the pleading stage of litigation as to the standard’s applicability in much the same way that attorneys currently argue the statutory purpose doctrine in negligence per se. Then, the judge, just as in negligence per se, should decide whether the standard applies. Additionally, the lawyers would be able to argue the history of the standard and whether it is an incestuous-industry standard, or a standard reflective of societal well being on a broader scale.

Part of the judge’s determination at this stage might include whether the case falls under the categorical “no duty” rules of the Restatement, which the Restatement asserts can be decided as a matter of law. However, categorical “no duty” rules can hardly be applied because they are so riddled with exceptions that depend on the specific facts of the case and

359. See Winfield I, supra note 36, at 186 (describing the original customs of the realm and describing the liability as “strict, but it was probably never absolute”).

360. See SCHWARTZ ET AL., supra note 1, at 686 (describing strict liability as imposing liability on the defendant “although the defendant neither intentionally acted nor failed to live up to the objective standard of reasonable care that traditionally has been at the root of negligence”).

361. See infra notes 380-88 and accompanying text (explaining that under the proposed formulation, the jury would still decide whether the defendant departed from the reasonable person standard).

362. See RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL & EMOTIONAL HARM § 16 cmt. f (2010) (stating that in some instances, a court may conclude that an actor’s compliance with a statute shows the actor did not lack reasonable care).

363. Id. (indicating that a statute does not always set the appropriate standard of care for the defendant).

364. See supra note 354 (noting that the judge, not the jury, determines if the defendant owed a duty to the plaintiff).

365. See, e.g., RESTATEMENT (SECOND) OF Torts § 288(b) (1965) (stating that a “court will not adopt as the standard of conduct of a reasonable man the requirements of a legislative enactment or an administrative regulation whose purpose is found to be exclusively . . . to secure to individuals the enjoyment of rights or privileges to which they are entitled only as members of the public”).

366. RESTATEMENT (THIRD) § 7 cmt. j.
considerations of foreseeability. As a result, the categorical “no duty” rules of the Restatement are not included in this proposed analysis.

The second of two negligence analyses should apply when the judge concludes that an established standard does not apply or when a case involves factual specifics that remove it from the coverage of any established or clearly

367. For example, Restatement (Second) of Torts § 315 states:

There is no duty to control the conduct of a third person as to prevent him from causing physical harm to another unless (a) a special relationship exists between the actor and the third person which imposes a duty upon the actor to control the third person’s conduct, or (b) special relation exists between the actor and the other which gives to the other person a right to protection.

RESTATEMENT (SECOND) OF TORTS § 315 (1965).

Various courts, however, have held that in special circumstances, there are exceptions to this rule based largely on foreseeability. See, e.g., Saccuzzo v. Krystal Co., 646 So. 2d 595, 596 (Ala. 1994) (stating the general rule that a person owes no duty to protect another from third party harm, but noting there is a “special circumstances” exception that exists when the defendant “knew or had reason to know of a probability of conduct by third persons that would endanger the plaintiff”); Bd. of Trs. of Univ. of D.C. v. DiSalvo, 974 A.2d 868, 871-72 (D.C. 2009) (finding that determining whether duty exists “is the result of a variety of considerations and not solely the relationship between the parties. Specifically, consideration of whether a duty exists to protect another from intervening criminal acts includes consideration of heightened foreseeability . . . .”);

Hyde v. City of Columbia, 637 S.W.2d 251, 253-57, 269 (Mo. Ct. App. 1982) (holding that a newspaper that published the name and address of crime victim owed a duty to the victim “to use reasonable care not to give likely occasion for a third party . . . to do injury to the plaintiff by the publication” of personal information; noting that “[t]he judicial determination of the existence of a duty rests on sound public policy as derived from a calculus of factors,” including foreseeability).

But see Corinaldi v. Columbia Courtyard, Inc., 873 A.2d 483, 489 (Md. Ct. Spec. App. 2005) (“[T]he general rule that there is no duty to protect a victim from the criminal acts of a third person” applies even though the party in question had actual knowledge that harm was imminent or the party had knowledge of facts that made harm foreseeable unless a special relationship can be found.).

Additionally, the purported “no duty” rule of § 315 of the Restatement (Second) of Torts may not apply when a plaintiff casts the defendant’s alleged negligence in terms of misfeasance, as opposed to nonfeasance. In such a case, the defendant is held to the ordinary standard of care, and foreseeability is a major factor in determining whether the defendant breached any duty. See Pamela L. v. Farmer, 112 Cal. App. 3d 206, 209 (Ct. App. 1980) (noting that Restatement (Second) of Torst section 315’s general rule of no duty “has no application where the defendant, through his or her own action (misfeasance) has made the plaintiff’s position worse and has created a foreseeable risk of harm from the third person; [i]n such cases the question of duty is governed by the standards of ordinary care.”). See also Doe Parents No. 1 v. State Dep’t of Educ. 58 P.3d 545, 583 (Haw. 2002) (“Absent a duty to adhere to a particular standard of care by virtue of the State and either the plaintiff or the third person sharing a “special relationship” . . . the State is, as is any person, generally required to exercise only “ordinary care” in the activities it affirmatively undertakes to prevent foreseeable harm to others.”) (citations omitted). Tex. Real Est. Holdings, Inc. v. Quach, 95 S.W.3d 395, 397 (Tex. Ct. App. 2002) (noting that although there is “no legal duty to protect another from the criminal acts of a third person,” “one may nevertheless be held liable if the criminal activity is a foreseeable result of one’s own negligence”); Prather v. Brandt, 981 S.W.2d 801, 806 (Tex. Ct. App. 1998) (“Generally, a defendant has no duty to prevent the criminal acts of a third party who is not under the defendant’s supervision or control unless the criminal conduct was the foreseeable result of the defendant’s negligence.”).
defined standards. This creates a situation where, if no existing standard governs the case, the judge will already have made the determination regarding the standard, thereby preserving the important judicial role of gatekeeper.

Professor Barbara White points out that the Carroll Towing foreseeability-based analysis for unreasonable conduct was created to deal with exactly the type of case where no well-defined duties exist to facilitate an easy ascertainment of reasonable conduct, and so it stands to reason that foreseeability should be applied in this type of case.

This analysis resembles the Restatement (Third) in that it presumes the existence of the general duty of care to act like a reasonable person while leaving the jury to decide whether the defendant conformed to that standard. Then, the jury applies an equivalent of the foreseeability-based analysis used in Carroll Towing. Again, as in the Restatement (Third)’s formulation, the judge could function as a gatekeeper by dismissing the case as a matter of law, based on the existence of an articulated no-duty rule as a matter of law.

In such a case, proof of negligence entails three elements. The first is that unreasonable conduct existed. This is accomplished by employing foreseeability as described above. The second element is a simplistic cause-in-fact analysis. If cause-in-fact exists, the third element becomes relevant. This third element is the determination of liability, which constitutes nothing more than asking the jury whether the defendant should be found liable. In charging the jury on this element, the court should instruct the jury to consider logic, common sense, policy, precedent, and “our more or less inadequately expressed ideas of what justice demands or of what is administratively possible and convenient.” As discussed above, some courts openly use this

---

368. See supra note 354 (stating that the judge, not the jury, decides whether the defendant owed a duty to the plaintiff).


370. See White, supra note 11, at 102, 104. Historically, the Carroll Towing formula [was] developed to define a standard of negligence or liability in areas for which traditional analyses were either ambiguous or inadequate. As Judge Hand noted, ‘there is no general rule to determine when the absence of a bargee or other attendant will make the owner of the barge liable for injuries to other vessels if she breaks away from her moorings.’

371. Id.


373. Id. (explaining the breach analysis as a factual one for the jury)

374. Id. § 6.

375. Snyder v. LTG Lufttechnische GmbH, 955 S.W.2d 252, 256 n.6 (Tenn. 1997) (quoting Bain v. Wells, 936 S.W.2d 618, 625 (Tenn.1997)).
to determine liability after cause-in-fact is established. The judge should also instruct the jury not to consider foreseeability in its determination of liability. This is in contrast to the Restatement (Third)’s formulation, in which foreseeability is applied both in determining unreasonable conduct and in proximate cause which is now called scope of liability. Repeated and parallel decisions producing similar outcomes under this analysis might eventually lead to general rules of care, which would then be applicable under the first negligence analysis outlined above.

A popular objection will likely be that the jury is incapable of doing this properly, but this objection can be debated forever. Given the undeniable fact that what the jury is asked to do in this second negligence analysis resembles the empirical process typical of the legislature, is it worse to have the jury, rather than judges, make these decisions? Leaving these decisions to the judge is abhorrent to the notion of separation of powers, while leaving it to the jury is less off an affront because the legislature is supposed to be reflective of the public’s needs and preferences. If we face the fact that courts sometimes legislate, then which aspect of the court, judge or jury, is more suited to do it becomes the only relevant debate. As Professor Cardi points out, this is an unending debate.

When a court determines that an affirmative or recognizable duty governs, the elements of negligence should be as follows:

---

376. See, e.g., id. at 256.
377. See supra Part VI.B. (suggesting that scope of risk is the same as foreseeability).
379. See generally Cardi II, supra note 11 (arguing for the existence of nuanced differences of foreseeability in separate elements of negligence).
380. How else could the following cases be explained? Compare Enright v. Eli Lilly & Co., 570 N.E.2d 198, 199-204 (N.Y. 1991) (finding no liability to anyone harmed by DES beyond those harmed in utero), and Ryan v. N.Y. Cent. R.R. Co., 35 N.Y. 210, 210-17 (N.Y. 1866) (holding that the only damages for which liability would attach as a result of a railroad’s negligent fire was damage to the railroad’s property and not houses in the area), with Kelly v. Gwinell, 476 A.2d 1219, 1220-30 (N.J. 1984) (holding a social host liable for serving a guest alcohol when the guest subsequently drove intoxicated and caused harm).

We judges are strange creatures. It is not that we are less brave than others, but rather by reason of our training, if not our nature, we tend to the conservative. For most of us, prudence and caution are the watchwords. We are rarely rewarded for taking risks. But the rest of the population does not always act the way we do. What may appear strange to judges might seem rather ordinary to others. It thus generally makes sense to have lay people, not judges, make decisions on the question of proximate cause, grounded as that concept is in considerations of foreseeability and fairness.

Id.
382. See Cardi I, supra note 10, at 987 (illustrating the many aspects of the judge/jury debate, albeit in the context of foreseeability).
Is there a duty?
Determined as a matter of law

Is there a breach of the duty?
Determined as a question of fact or a mixed question of law or fact

Is there cause-in-fact?
Determined as a question of fact

Is there proximate cause, or whatever name the powers that be ascribe to this element, which is based solely on foreseeability?
Determined as a question of fact

When the court determines that no existing standard reflects social norms, the elements of negligence should be as follows:
Was there unreasonable conduct?
Determined as a matter of fact based on foreseeability or as a question of law if the judge thinks the facts excuse the defendant from the standard of reasonable conduct

Was there cause-in-fact?
Determined as a matter of fact

Should the defendant be liable?
A policy question which does not involve foreseeability and is question of fact

It is important to note that in both of these formulations of negligence, foreseeability is utilized only once in the negligence analysis. This simultaneously reflects the dual truths that foreseeability can only be one thing and that foreseeability cannot be discarded because it is culturally entrenched in our society. Even if this dual test for negligence does not affect outcomes, it will improve the fidelity of the judicial system because the existence of stated elements, if consistently applied, maintain judicial integrity and the proper functioning of the judicial system.

383. See supra note 227 and accompanying text.
384. See James Jr. & Perry, supra note 11, at 786.
385. See Cardi I, supra note 10, at 934-35. Professor Cardi listed three important reasons for having cause of action divided into elements:

First, a list of elements guides the liability inquiry by providing a structured, ordered method according to which a court or jury may reason through the substantive requirements.

Second, division of a cause of action into elements facilitates decisionmakers' analysis of the legal concepts upon which liability rests. By considering such concepts distinctly and in ordered sequence, courts more capably identify the characteristics of each legal concept, isolate the analysis of each concept from that of others, and determine the relationship between concepts. These acts, in turn, coalesce courts' understanding of the prerequisites to liability and lead to a more correct and consistent application thereof.

Third, setting forth a theory of recovery as a series of elements provides a useful means by which to segregate issues to be decided by the court
VIII. CONCLUSION

Negligence evolved as a way to compensate people whose injuries did not allow them compensation under the old English writ pleading system.\(^{386}\) Originally, compensation was dependant on a showing that an established societal custom, reflected in a statement about the reasonable person, was violated.\(^{387}\) As the industrial revolution increased society’s complexity and diversity, situations arose in which compensation was justified, but lacked an established custom of the realm.\(^{388}\) As a solution, a standard developed that was no longer tied to an identifiable custom of the realm and became known as the reasonable person standard.\(^{389}\) If a person should have foreseen that his behavior would have caused harm, then he had a duty to act accordingly, if the cost was not higher than the harm that would occur.\(^{390}\) Carroll Towing is the ultimate expression of this concept.\(^{391}\)

This led to an expansion of liability that had to be addressed with doctrines, such as proximate cause.\(^{392}\) There could then be no compensation for another’s unreasonable conduct unless the harm caused was foreseeable.\(^{393}\) Foreseeability became a schizophrenic tool for both expanding liability and for limiting liability. The scholarship and judicial decisions attempt in vain to accommodate both roles of foreseeability.\(^{394}\) Palsgraf typifies and exacerbates this schizophrenia.\(^{395}\)

In the meantime, consistent decisions in negligence, based on similar facts, led to the modern equivalents of customs of the realm: distinct, discernable, and tangible approximations of affirmative standards that clearly define what society thinks is optimal.\(^{396}\)

\(^{386}\) See Kelley I, supra note 11, at 61 (explaining “what was originally trespass on the case for negligence of default became modern negligence”).

\(^{387}\) Baker, supra note 21, at 58.

\(^{388}\) Id. at 346-47 (explaining that “duties of care could not be imposed on everyone” and the development of the reasonable person standard by Lord Bathurst to address this issue).

\(^{389}\) Id.

\(^{390}\) See Restatement (Third) of Torts: Liab. for Physical & Emotional Harm § 3 (2010) (indicating that the issue of whether a person’s conduct lacks reasonable care turns on the following factors: “the foreseeable likelihood that the person’s conduct will result in harm, the foreseeable severity of any harm that may ensue, and the burden of precautions to eliminate or reduce the risk of harm”).

\(^{391}\) United States v. Carroll Towing, 159 F.2d 169 (1947).

\(^{392}\) See Dobbs supra note 69, at 443 (describing the main purpose of proximate cause as a way “to limit the defendant’s liability to the kinds of harms he risked by his negligent conduct”).

\(^{393}\) See Baker, supra note 21, at 347.


\(^{395}\) Id.

\(^{396}\) See supra note 341 and accompanying text.
The Restatement ignores these standards and fails to alleviate the problem of duplicative foreseeability, except for an inconsequential and ineffectual change in nomenclature regarding what it formerly called proximate cause.\footnote{RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM (2010).}

The key to bringing negligence back from the brink of absurdity is, as Professor Kelley states, to think about the original historical formulation of negligence, tied to specific customs of the realm, and to develop a negligence structure based on our own current equivalents of customs of the realm.\footnote{See Kelley II, supra note 11, at 1064-66.} My first formulation of negligence reflects this. When there are no equivalent customs of the realm, a different negligence analysis is required. My second negligence formulation addresses this. Together, both of my formulations of negligence can provide a cure to the confusion surrounding modern negligence.
PRESIDENTIAL AMBITIONS OF U.S. SUPREME COURT JUSTICES: A HISTORY AND AN ETHICAL WARNING

William G. Ross*

I. INTRODUCTION

A remarkably large number of United States Supreme Court Justices have had presidential aspirations. In at least three quarters of the presidential elections between 1832 and 1956, one or more justices attempted to obtain a presidential or vice presidential nomination or were prominently mentioned as potential candidates.¹ Several justices conducted covert presidential campaigns, and a few even openly campaigned from the bench.² On numerous occasions, political parties or political leaders encouraged justices to become presidential candidates.³ One Justice, Charles Evans Hughes, resigned from the Court to accept a presidential nomination.⁴

No justice during the past half century, however, appears to have entertained serious presidential ambitions. It is probably no coincidence that it has been nearly a half century since any justice young enough to be elected president has held any significant elected office before taking his or her seat on the Court.⁵ During the past half century, most justices have come to the Supreme Court from the United States Court of Appeals and have had previous careers as federal bureaucrats, private practitioners, or academics.⁶ With the exception of Sandra Day O’Connor, no justice appointed since Earl Warren in 1953 has held any significant elected office, and few have been active in partisan politics.⁷ In contrast, most of the many justices of yore who had presidential aspirations had also served as governors or members of Congress, and many had run for President before ascending to the bench.⁸ Until the 1950s, many justices were

---

¹ Writing in 1958, John Frank accurately concluded that “for at least one hundred and twenty-five years, there has been no ten-year period in which a Supreme Court Justice has not been seriously and soberly considered for the presidential office.” JOHN P. FRANK, THE MARBLE PALACE: THE SUPREME COURT IN AMERICAN LIFE 275 (1958). As Frank observed, “[m]ost glittering of all the baubles for a roving judicial eye has been the greatest honor of all, the Presidency.” Id. at 273.
² See infra Part II.
³ See infra Part II.
⁴ See infra Part II.D.
⁵ See infra Part II.
⁶ See infra Part II.
⁷ See infra Part II.
⁸ See infra Part II.
appointed because they were prominent in national politics. Only a few have become prominent in national politics merely because they were justices.

During recent years, some critics of the Supreme Court appointment process have complained that the career backgrounds of justices have become too narrow and that the Court would benefit from members who have had experience as elected legislative or executive officials. Shortly after David H. Souter announced his retirement in 2009, Senate Majority Leader Harry Reid expressed his preference for the appointment of someone other than a judge — “people with some real-life experiences . . . rather than people who walk around in those black robes all the time.” Similarly, Senator Patrick Leahy of Vermont, a member of the Senate Judiciary Committee, remarked in the wake of Souter’s retirement that he would “like to see more people from outside the judicial monastery, somebody who has had some real-life experience, not just as a judge.” And Vice President Joseph R. Biden, a longtime member of the Senate Judiciary Committee before his election to the vice presidency, declared in 2007 that, “We have enough professors on the bench . . . I want someone who ran for dog catcher.”

Although the Court would probably benefit from justices who had served as elected officials because such persons might have a better understanding of the practical realities of politics and have a more profound appreciation of such issues as separation of powers and federalism, history demonstrates that justices who have extensive political experience may have presidential aspirations. Most scholars agree that presidential ambitions can distract justices and tempt them to trim their decisions to appeal to constituencies that could assist their candidacy. As Felix Frankfurter once asked, “What is more inimical for good work on the Court than for a justice to cherish political, and more particularly Presidential, ambition?” Part II of this article will provide a history of presidential ambitions among Supreme Court Justices. Part III will explain the way in which presidential ambitions among justices are likely to diminish public confidence in the integrity of the Court’s decisions and

9. See infra Part II.
10. See infra Part II.
11. See infra Part II.
15. See infra Part III.
16. See infra Part II.
17. See infra Part II.
19. See infra Part II.
the extent to which such ambitions could influence the votes or opinions of politically ambitious justices.20

II. A HISTORY OF PRESIDENTIAL ASPIRATIONS AMONG SUPREME COURT JUSTICES

During the early years of the Republic, when both the Supreme Court and the federal government were less powerful than they have become, a Supreme Court justiceship was not necessarily regarded as the pinnacle of a career.21 John Jay was a gubernatorial candidate in New York in 1792 and he resigned as chief justice in 1795 to become governor;22 William Cushing was a candidate for the Massachusetts governorship in 1794;23 and John Rutledge resigned as associate justice in 1791 to become chief justice of the South Carolina Court of Common Pleas.24

A. The Early Nineteenth Century

John Marshall appears to have been the first member of the Court to be regarded as a presidential possibility, although there is little evidence that Marshall sought the presidency. Marshall, who served for 14 years in the Virginia legislature and who had briefly served as a member of the U.S. House of Representatives and as Secretary of State, was no stranger to political office,25 but he relished his role as chief justice and left a legacy greater than that of most presidents.26 In 1812, a decade after Marshall became chief justice, many Federalists regarded him as the ideal candidate to defeat James Madison’s bid for re-election as a Jeffersonian Republican.27 As one of Marshall’s biographers has explained, Federalist interest in Marshall was “natural” because

he was well known to all the prominent leaders of the party but still had managed to stay above the local squabbles of the past several years. His refusal to support the Alien and Sedition Laws years earlier was virtually forgotten. Also, it was believed that he could bring in some

20. See infra Part III.
21. See infra notes 22-24 and accompanying text.
26. See generally id. at 180-200.
additional votes because of his nonpolitical stance over the past several years.\textsuperscript{28}

Marshall’s judicial decisions supporting a powerful national government and vested property rights also must have made him attractive to Federalists.\textsuperscript{29}

Federalists also believed that Marshall would help them carry critical states.\textsuperscript{30} In doing their electoral arithmetic, Federalists believed that their party needed to carry either Virginia or North Carolina to win the election.\textsuperscript{31} They could count on the 88 electoral votes of the anti-war New England states and had a reasonable chance of carrying Delaware and Maryland, but they still would need 15 more for victory.\textsuperscript{32} They pinned their hopes on North Carolina and Virginia because they lacked any hope of carrying South Carolina, Georgia, or any western state, and were doubtful about success in the Middle Atlantic states.\textsuperscript{33} Madison was a Virginian, so Federalists believed that they could carry Virginia only if a Virginian were on the ticket.\textsuperscript{34} Some Federalists supported Marshall because they regarded him as the only Federalist who could carry Virginia,\textsuperscript{35} or because they regarded him as a stronger candidate than the front-runner, DeWitt Clinton, who eventually received the nomination.\textsuperscript{36}

Many Federalists, however, opposed Marshall’s nomination because they believed that he could serve the Federalist cause more effectively on the Court, which, as a Marshall biographer has observed, was “a more accurate assumption than people realized at the time.”\textsuperscript{37} Some Maryland Federalists wanted to run Marshall as a peace candidate for president after the declaration of war in June, but he spurned the overtures and urged Federalists to make common cause with antiwar Republicans.\textsuperscript{38} Meanwhile, Marshall removed himself from the political arena by spending the summer of 1812 on an expedition that explored the western wilderness of Virginia.\textsuperscript{39} Marshall’s biographer, Albert J. Beveridge, believed that Marshall would have been a much stronger candidate than Clinton because “[t]he lawyers generally would have worked hard for him,” and he would have had more appeal to the mercantile classes.\textsuperscript{40} Beveridge also concluded that “Federalists, who accepted Clinton with repugnance, would have

\begin{itemize}
  \item \textsuperscript{28} Id. at 533.
  \item \textsuperscript{29} See id. at 531-534.
  \item \textsuperscript{30} Id. at 533-534.
  \item \textsuperscript{31} Id. at 534.
  \item \textsuperscript{32} Id.
  \item \textsuperscript{33} BAKER, supra note 27, at 534.
  \item \textsuperscript{34} DAVID LOTHI, CHIEF JUSTICE JOHN MARSHALL AND THE GROWTH OF THE REPUBLIC 281 (1949); BAKER, supra note 27, at 534.
  \item \textsuperscript{35} LOTHI, supra note 34, at 281.
  \item \textsuperscript{36} BAKER, supra note 27, at 534-35.
  \item \textsuperscript{37} Id. at 535.
  \item \textsuperscript{38} JEAN EDWARD SMITH, JOHN MARSHALL: DEFINER OF A NATION 410 (1996).
  \item \textsuperscript{39} LOTHI, supra note 34, at 281.
  \item \textsuperscript{40} ALBERT J. BEVERIDGE, THE LIFE OF JOHN MARSHALL, VOL. IV, THE BUILDING OF THE NATION 1815-1835 47 (1919).
\end{itemize}
exerted themselves to the utmost for Marshall, the ideal representative of Federalism.  

In the election of 1816, Marshall received four electoral votes for vice president, the smallest number of votes received by any of the running mates of the unsuccessful Federalist candidate, Rufus King. All four votes came from Connecticut, and there is no apparent evidence that Marshall sought the vice presidency or the presidency during that year.

Shortly before the end of Marshall’s tenure, Associate Justice John McLean commenced his perennial campaigns for the presidency. McLean holds the dubious record among justices for the most protracted campaign for the presidency – he was a candidate of one type or another during at least seven of the eight presidential election campaigns between 1832 to 1860, which spanned nearly all of his thirty-two years on the bench. McLean, who had served two terms in the U.S. House of Representatives and six years as postmaster general, had been politically prominent before becoming a justice, but little in his background made him an obvious candidate for the presidency.

At various times during this period, McLean was considered as a candidate by at least eight political parties. Like Salmon P. Chase later in the century, McLean sometimes was available to more than one party during the same election year. In 1832, he hoped at various times to become the standard bearer of the National Republicans, the Democrats, or the Anti-Masons. In advance of the 1836, 1844, and 1848 elections, he pinned his hopes on the Whigs, and in 1848 and 1852 he was mentioned as a possible candidate of the Free Soil party. In advance of the 1856 election, McLean angled for the presidential nomination from the nativistic American (“Know Nothing”) Party while focusing his attention on the newly formed Republican Party. At the first Republican convention in 1856, McLean received 190 votes – about one third of the total--on a first informal ballot for the presidential nomination, and thirty-seven votes

41. Id.
43. Connecticut split its nine electoral votes for vice president in 1816 - with four going to Marshall, and the remaining five for former Senator James Ross of Pennsylvania, who like Marshall, received no electoral votes from any other state. Id.
45. See id. at 69-80, 81-98, 103-105, 105-119, 136-137, 146-152, 211-215.
46. Id.
47. Id.
48. See id. at 74, 76, 79.
49. Id.
McLean quietly sought the Republican nomination again in 1860, a year in which he received twenty-two votes at the convention of the Constitutional Union Party. Meanwhile, McLean spurned overtures for consideration for vice presidential nominations, not because he had any scruples about politicking from the bench, but rather because he regarded the vice presidency as too passive an office.

McLean’s presidential ambitions may have distracted him from his judicial duties, which included not only his service on the Supreme Court, but also his work as a member of the circuit court on which all Supreme Court justices served until 1891. After reviewing McLean’s papers in the Library of Congress, John Frank wryly reported that he (Frank) had “found there substantially nothing on the business of being a judge,” but that he had uncovered “an endless stream of observations” that concerned his perennial candidacy for the presidency.

McLean made no apologies for his presidential ambitions. In a letter written shortly after he became a Supreme Court justice, McLean defended the propriety of presidential candidacies by Supreme Court justices, expressing disbelief that any reflective person “could entertain the least apprehension of any improper influence being used by a Judge who comes before the people in a popular election, and especially that it could lend to corrupt the Bench.”

McLean contended that a judgeship “is without exception the most unfavorable post a candidate can occupy” because a judge has no patronage to dispense and in every judicial decision

disappoints one party and his counsel, who, though restrained in their expressions of resentment against him as Judge, would gladly in the exercise of their right of suffrage show their disapprobation. And if the Judge can be supposed to be influenced in any decision by popular consideration, his popularity is at once destroyed.

53. WEISENBURGER, supra note 44, at 213.
54. Id. at 104-05.
56. FRANK, supra note 1, at 273. McLean’s political correspondence was so voluminous that Congress in 1849 debated the question of whether Supreme Court justices should continue to enjoy the franking privilege. Id.
58. Id.
McLean expressed his political views in many newspaper articles and letters that quickly were released for publication. In 1847, for example, he wrote a letter excoriating the Mexican War, and in 1848 wrote a letter expressing his view on one of the most controversial political and constitutional issues of the day – the power of Congress to prohibit slavery in the territories. Eight years later, when public debate over this issue had reached an even more intense pitch, McLean again publicly discussed the issue of slavery on the eve of the Court’s decision in *Dred Scott*.

McLean had long been an outspoken opponent of slavery, so there is little reason to suppose that his political ambitions motivated his dissent in *Dred Scott*. Because McLean’s fellow justices recognized that *Dred Scott* offered what Don E. Fehrenbacher has described as “a rare opportunity for judicial heroics” that might advance McLean’s presidential ambitions, the Court may have postponed its decision until after the election to dash his candidacy. McLean’s dissent may very well have enabled him to obtain the Republican presidential nomination in 1856 if the Court had decided the case before the Republican convention.

**B. The Lincoln Appointees**

Presidential fever on the Court became even more intense after McLean’s death in 1861, for four of the five Justices appointed by Abraham Lincoln – Chase, Stephen J. Field, Samuel Miller, and David Davis – had strong presidential ambitions and were actively considered as candidates during their years on the Court.

With the exception of McLean, Chase holds the Court’s record for the most vociferous quest for the presidency. Chase’s ambitions were more significant

---

59. *Id.*
60. *Id.* at 270-71.
61. *Id.*
63. Professor Fehrenbacher has questioned whether McLean’s “pathetic whispers of hope” for attaining the presidency in 1860, when he would be seventy-five years old, “actually governed his judicial conduct” in *Dred Scott*. DON E. FEHRENBACHER, THE DRED SCOTT CASE: ITS SIGNIFICANCE IN AMERICAN LAW AND POLITICS 310 (1978). Moreover, Fehrenbacher pointed out that “one may well suspect bias in a historical explanation that attributes selfish personal motives only to the two dissenting justices,” McLean and Benjamin Curtis. *Id.*
64. *Id.* at 289.
65. *Id.* at 202, 290; WEISENBURGER, supra note 44, at 197.
66. Finkelman, *supra* note 50, at 564-565. Professor Finkelman has pointed out the irony that the desire of Republican leaders to nominate a more politically experienced candidate in 1860 may have caused them to ignore “McLean’s solid and powerful response to Taney’s majority opinion.” *Id.* at 565.
67. See infra Part II B.
than McLean’s, however, because they were more realistic. Chase’s presidential ambitions preceded his tenure on the Court, for he was one of the leading contenders for the Republican presidential nomination in 1860.\textsuperscript{68} Lincoln’s presumed motive for appointing Chase both as Secretary of the Treasury and as Chief Justice was to co-opt a dangerous political rival.\textsuperscript{69} Chase also had sought the Republican nomination in 1856 and had hoped to obtain it in 1864 if the Republicans refused to re-nominate Lincoln,\textsuperscript{70} who regarded Chase’s presidential ambition as “mild insanity.”\textsuperscript{71} Like other political activists who have become justices, Chase was bored by legal technicalities and often found his judicial duties unduly dry and tedious.\textsuperscript{72} As Chase admitted privately in January 1866, “I think I have a good deal of executive faculty and often wish I were in some more active employment than hearing causes which take up my whole time.”\textsuperscript{73}

Because Chase already was a perennial presidential candidate when he became chief justice, it is hardly surprising that his critics from the beginning suspected that his every move was calculated for political advantage.\textsuperscript{74} A passionate abolitionist before the war, Chase ardently advocated black suffrage after emancipation, and his efforts to promote civil rights sometimes became entangled with his political ambitions.\textsuperscript{75} For example, he toured the South in May 1865 partly for the purpose of encouraging better conditions for the freedmen, but his tour also appears to have enabled him to assess his political prospects in that region.\textsuperscript{76} As one scholar has written, “Chase corresponded familiarly and indiscreetly with men of the widest possible differences of political opinion” in his efforts to win the presidency.\textsuperscript{77}

\begin{itemize}
  \item \textsuperscript{68} John Niven, Samuel P. Chase: A Biography 212-21 (1995).
  \item \textsuperscript{69} Frederick J. Blue, Salmon P. Chase: A Life in Politics 225 (1987).
  \item \textsuperscript{70} Id. at 111-15, 212-26.
  \item \textsuperscript{71} James Ford Rhodes, History of the United States: From the Compromise of 1850 to The Final Restoration of Home Rule at the South in 1877, Vol VI 168 (1906).
  \item \textsuperscript{72} As one of Chase’s biographers has explained, “[t]hough supremely industrious, Chase had little liking for reading and digesting hundreds of pages of briefs, depositions, and lower court testimony, where one had to be scrupulously careful about trivial points of language . . . He was working as hard as he had ever worked, but not for great issues or worthy causes.” Niven, supra note 68, at 376.
  \item \textsuperscript{73} Charles H. Coleman, The Election of 1868: The Democratic Effort to Regain Control 70 (1971 repr., 1933). Coleman believed that Chase probably “was sincere when he considered the Presidency primarily as an opportunity for service and a chance to show his powers to the best advantage.” Id. at 72.
  \item \textsuperscript{74} See Blue, supra note 69, at 283. In a letter appraising Chase’s character, Justice Miller remarked that “I doubt if for years before his death, his first thought in meeting any man of force, was not invariably how can I utilize him for my presidential aspirations.” Charles Fairman, Mr. Justice Miller and the Supreme Court 251-52 1862-1890 (1939)(quoting letter of Oct. 15, 1876).
  \item \textsuperscript{75} Blue, supra note 69, at 283.
  \item \textsuperscript{76} Id.
  \item \textsuperscript{77} Coleman, supra note 73, at 72.
\end{itemize}
Because Chase was a Republican, he naturally first sought the nomination of his own party. Still smarting from the party’s preference for the upstart Lincoln in 1860, Chase believed that he once again was the most experienced and prestigious member of the party. Chase’s potential candidacy attracted at least some interest among Republicans during the summer of 1867, when “Grant had not yet broken with Johnson, and the Radicals doubted the General’s soundness on reconstruction” and the rights of African Americans.

Chase’s impartiality in presiding over the impeachment trial of Andrew Johnson early in 1868 further diminished his slim chances of obtaining the Republican nomination. Republican Radicals accused Chase of bias in favor of Johnson, alleging that he had slanted his procedural rulings and even lobbied senators to acquit Johnson in an attempt to curry favor for the Democratic presidential nomination. Meanwhile, the Democratic revival in 1868 convinced many Republicans that only the nomination of a military hero would ensure victory. Early in 1868, Chase emerged as a possible Democratic candidate. This was not as farfetched as it seemed, for he had consistently adhered to the egalitarian principles of the original Democratic Party, and he had opposed radical Reconstruction measures. There is evidence that some Democratic leaders cynically encouraged Chase’s candidacy, even though they had no intention of nominating him, because they hoped that his reputation for probity and his role in helping to win the war would cloak the party with greater moral respectability and help the party win votes among moderates who otherwise were repelled by the party’s overt racism and hostility toward Reconstruction.

Although Chase consistently disavowed any interest in becoming a candidate, he fanned interest in his presidential prospects by providing prominent Democrats with a detailed account of his political opinions, a virtual platform that seemed to signal his availability, and by offering advice about how the Democrats might secure the presidency. Chase urged the party to accept universal African American suffrage, but he also advocated amnesty for all former Confederate leaders and an end to military reconstruction of the South. Despite his role as Treasury Secretary in raising taxes and issuing the first paper money in order to finance the war, Chase also appealed to Democrats by
advocating lower taxes and “good money” in the form of hard currency.\textsuperscript{88} As one historian has observed, Chase “was a Republican where the negro was concerned and a Democrat on other issues.”\textsuperscript{89}

In a private letter in June 1868, Chase admitted that he would “not feel at liberty to decline” a nomination representing “the wishes of the masses.”\textsuperscript{90} By the middle of June, Chase began to retreat from his long-cherished wish for universal suffrage, informing Democratic leaders that although universal suffrage remained a democratic principle, the Constitution allowed the states to decide who should be permitted to vote.\textsuperscript{91} Chase included this in a fifteen-point proposed Democratic platform that he permitted to be published days before the convention began.\textsuperscript{92} Although Chase appears to have allowed his presidential ambition to interfere with his long-held support for universal suffrage, some historians have pointed out that Chase’s retreat gave him an opportunity to move the Democratic Party forward on this issue.\textsuperscript{93}

Chase won the support of many prominent Democrats, including Governor Horatio Seymour of New York, who later became the Democratic presidential nominee in 1868.\textsuperscript{94} Additionally, many friends and supporters of Chase went to the Democratic convention in New York City in July 1868 to lobby for his nomination.\textsuperscript{95} A Chase Executive Committee of 100 prominent men worked openly for his nomination, while a smaller circle of friends pulled strings behind the scenes.\textsuperscript{96} Meanwhile, Chase’s daughter, Kate Chase Sprague, “was tireless in her efforts for her father’s nomination” and used her Manhattan home as “a center of Chase agitation.”\textsuperscript{97}

\textsuperscript{88}. Id. 107, 121.
\textsuperscript{89}. Id. at 117.
\textsuperscript{90}. Id. at 115.
\textsuperscript{91}. Id. at 121-26.
\textsuperscript{92}. Blue, supra note 69, at 290; Coleman, supra note 73, at 136-39.
\textsuperscript{93}. See Coleman, supra note 73, at 126. Coleman charitably speculated that Chase may have “thought that if he were once in the White House with the backing of the Democratic party, he could use the executive power to foster a broadening of the suffrage by those white men who would be in control in the South in the event of Democratic victory.” Id. Similarly, another historian has suggested that Chase’s new position provided Chase with an opportunity to make Democrats more amenable toward black suffrage. Blue, supra note 69, at 290. The same historian has pointed out that Chase’s consignment of the issue to the states was not hostile toward blacks in the context of the times because it was consistent with the understanding of the Framers of the Fourteenth Amendment and with the Republican platform in 1868. Id. Professor Niven, however, has written that Chase’s ambition caused him to ignore his long commitment to human equality and to swing “slowly to the dark side of political expediency.” Niven, supra note 68, at 429.
\textsuperscript{94}. Niven, supra note 68, at 428; Coleman, supra note 73, at 126-27, 131-32, 223.
\textsuperscript{95}. Niven, supra note 68, at 430; Coleman, supra note 73, at 215.
\textsuperscript{96}. See Coleman supra note 73, at 221-23.
\textsuperscript{97}. Coleman, supra note 73, at 215; Blue, supra note 69, at 292. Sprague may have been particularly eager for her father to become president because he was a widower and she, his eldest daughter, would have served as White House hostess and “First Lady.” Coleman, supra note 73, at 221. During the convention, Chase expressed fear to her that “you are acting too much the politician. Have a care. Don’t do or say anything which may not be proclaimed from the house tops.” Id. at 220.
Chase’s supporters agreed that they would not immediately present his name at the convention but rather would wait for a deadlock.98 When the convention deadlocked after several days, however, Chase’s supporters failed to persuade party leaders to swing their support to Chase.99 On the twenty-third ballot, Seymour suddenly emerged as the compromise candidate.100 Chase received a one-half vote on five of the twenty-three ballots, and four votes on another ballot.101

Early in 1868, Chase also had been prominently mentioned as a possible leader of a new third party that presumably would have drawn support from moderate Democrats and conservative Republicans.102 After the convention, there was more speculation that Chase might lead a third party.103 As the American correspondent of a London newspaper wrote, Chase “really seems ready to accept a nomination from anybody on any platform.”104 Shortly after the convention and again when Democratic prospects wilted in October, various Democrats urged their party to dump Seymour in favor of Chase or another candidate.105 Although Chase privately urged his supporters to remain loyal to Seymour, he could not resist pointing out that the conservatism of the Democratic platform had encouraged wayward Republicans to return to the Republican party.106

Chase’s presidential ambitions in 1868 were hobbled by many factors, including the failure of supporters to effectively organize and coordinate their efforts, the superior organizational skills of his opponents, his inability to openly promote his own candidacy, his failure to obtain solid support in his home state of Ohio, and his disagreements with both parties regarding key political issues.107

Shortly after the Democratic convention of 1868, The Nation declared that Chase had “destroyed popular confidence in his decisions” and that “no weight whatever will hereafter attach to any judgment of his on any one of the great constitutional questions arising out of the rebellion and reconstruction.”108 One

99. Id. at 294-295.
100. Id. at 295.
101. Coleman, supra note 73, at 382.
102. Id. at 112-16.
103. Id. at 275.
106. Blue, supra note 69, at 299. Blue has explained that “Chase and his partisans did not instigate the movement to change the Democratic ticket, although some of the Chief Justice’s more zealous followers did try to capitalize on it.” Id.
107. Id. at 284, 295, 296. Professor Blue believes that the fact “[t]hat Seymour’s name rather than Chase’s was placed in nomination suggests that the Chief Justice and his supporters were out of touch with the mood of the convention and that his nomination had never been realistic.” Id. at 295.
108. Fairman, Reconstruction and Reunion, supra note 105, at 548 (quoting The Nation,
of Chase’s recent biographers has concluded that Chase’s vacillations during the campaign left him “identified as a weak and contradictory person before the public. His reputation had suffered badly.”\textsuperscript{109} Similarly, another historian has concluded that Chase “lost something of the respect of his contemporaries and consciously lowered the dignity of the office” as the result of his 1868 presidential aspirations.\textsuperscript{110} Writing early in the twentieth century, the historian James Ford Rhodes expressed a view that appears to have withstood the test of time:

The nomination of Chase would have been the condonation of a grossly improper action. That a Chief Justice . . . should have gone a-begging and a-scheming after political power is enough in itself to sicken recollection . . . Men at the time said that he had dragged his silk gown in the mire. And once muddied, he must proceed to wallow. If these Republicans won’t have me for a President – Oh, very well, come on you others; I’ll be a Democratic candidate . . . Nothing in his life has so detracted from his fame.\textsuperscript{111}

In 1872, Chase was widely regarded as a potential presidential nominee of the newly formed Liberal Republican Party, and Chase probably would have attracted substantial support if his health had not been in visible decline.\textsuperscript{112} Although Chase once again disclaimed ambition, he privately acknowledged that he would not decline the nomination.\textsuperscript{113} At its convention in May, Chase received votes on several ballots, receiving a high of 32 votes on the sixth ballot before the convention nominated Horace Greeley.\textsuperscript{114}

Chase’s presidential aspirations naturally have generated suspicion that Chase allowed his ambition to influence his performance as chief justice.\textsuperscript{115} Chase’s efforts to obtain the presidential nominations of as many as four political parties makes these charges more difficult to assess than if Chase had hankered after the nomination of only one party.\textsuperscript{116} On the whole, however, there is little evidence that Chase’s intense political ambitions directly affected his work on the Court.\textsuperscript{117} Chase’s votes and his majority and dissenting opinions

\textsuperscript{109} Niven, supra note 68, at 432.
\textsuperscript{110} Coleman, supra note 73, at 72. Coleman has observed that Chase “realized the impropriety of his ambition for the Presidency and struggled against the Presidential disease – but in vain.” Id.
\textsuperscript{111} Rhodes, supra note 71, at 168.
\textsuperscript{112} Blue, supra note 69, at 316-17; Niven, supra note 68, at 447.
\textsuperscript{113} Blue, supra note 69, at 317.
\textsuperscript{114} Id. at 317-18.
\textsuperscript{115} See infra note 126.
\textsuperscript{116} See supra Part II.B.
\textsuperscript{117} See infra notes 121-127 and accompanying text.
are largely consistent with his long-term political philosophies, and sometimes even may have been at odds with his short-term political advantage.118

In cases concerning Reconstruction, the dominant issue that came before the Court during Chase’s judicial tenure, the Chief Justice’s opinions generally favored the positions and principles of the Republican Party in a manner that was consistent with Chase’s ardent support of the Union cause as Secretary of the Treasury during the Civil War and as a lifelong opponent of slavery.119 In cases that came before the Court before he became a candidate for the Democratic nomination in 1868, Chase dissented from decisions of the Court that invalidated state and federal loyalty oaths for former Confederates,120 and he wrote the Court’s opinion denying an injunction to restrain President Johnson from enforcing the Military Reconstruction Acts.121 When the Court held that military trials of civilian Confederate sympathizers were unconstitutional in areas in which civil courts were functioning, Chase wrote a dissenting opinion arguing that Congress at least had the power to authorize military courts under such conditions.122 In 1869, Chase wrote the opinion for the Court in Ex parte McCardle, which held that Congress could withdraw at least some of the Court’s habeas corpus jurisdiction,123 and his opinion for the Court in Texas v. White in 1870 further validated congressional policies of Reconstruction.124 As chief justice, Chase also remained a champion of the rights of African-Americans. Chase incurred the wrath of Democrats when he admitted the first black to practice before the Court in 1865, and he was the sole dissenter in an 1872 decision sustaining a contract for the sale of a slave before the Civil War.125

The only case in which Chase’s behavior might have been questionable was Hepburn v. Griswold,126 in which Chase wrote the Court’s opinion that the statute authorizing the issuance of paper money during the Civil War did not require the acceptance of paper for debts incurred before the statute’s enactment.127 Because Chase, as Treasury Secretary, had urged the enactment,
and because the Court’s decision was highly popular among Democrats and was widely criticized by Republicans, Chase has been vulnerable to allegations that he allowed his presidential ambition to influence his vote.128 Chase, however, always had been an advocate of “hard money” and only the extreme exigencies of the war had persuaded him to advocate the issuance of greenbacks when he was Treasury Secretary.129 Moreover, by 1870, Chase probably was pinning any remaining political hopes more on the Republicans than on the Democrats.130

Next to Chase, the most politically ambitious Gilded Age justice was the rambunctious Stephen J. Field, who was a candidate for the presidency at least three times during his thirty-four year tenure on the Court.131 Although Field’s ambitions generated considerable support, he never came close to receiving a presidential nomination.132 As one historian has observed of Field, “[w]hile maintaining a proper appearance of aloofness, he really cherished expectations greater than the event warranted.”133 Like most justices of his era, Field had been active in politics before taking his seat on the Court, having served in the California assembly and unsuccessfully sought seats in the California Senate and the United States Senate.134 He was an elected member of the California Supreme Court when Lincoln appointed him to the United States Supreme Court in 1864.135

In 1868, Field received fifteen votes from the California delegation at the Democratic convention, where the California delegation hailed Field as “the guardian of the Constitution of his country against all the power of the Radical party.”136 In 1880, Field contrived to win the Democratic nomination, engaging in one of the most egregious examples of politicking from the bench in Supreme Court history.137 Both of Field’s biographers agree that his political ambitions may have influenced his opinions, if not his votes, in important cases.138

Field’s motives for seeking the nomination in 1880 are unclear. Like a number of other justices who have sought the presidency, Field was dissatisfied on the Court because he was increasingly in the minority in cases in which the Court decided to expand the scope of federal power and sustain the

128. BLUE, supra note 69, at 304.
129. NIVEN, supra note 68, at 297.
130. See supra note 112 and accompanying text.
132. FAIRMAN, MR. JUSTICE MILLER, supra note 74, at 297.
133. Id.
134. See McCloskey, supra note 131, at 529-550.
135. Id.
136. FAIRMAN, RECONSTRUCTION AND REUNION, supra note 105, at 545-46.
137. See McCloskey, supra note 131, at 540-42.
constitutionality of economic regulatory legislation; Field’s theory of substantive due process did not achieve its partial triumph until after Field had retired. Field also may have wanted to block the candidacy of Samuel J. Tilden, who had been a political rival of one of Field’s brothers, David Dudley Field, and a business rival of another, Cyrus Field. Another motivation may have been Field’s desire to restructure the Supreme Court and to appoint justices who would promote Field’s theories of constitutional interpretation. Field’s legendarily large ego also must have stoked his presidential dreams.

Like so many other justices who campaigned for the presidency from the bench, Field lacked effective campaign organization. Like other judicial aspirants for the presidency, Field also lacked a coherent platform. As Paul Kens has explained, “A political record did not flow naturally from his duties as a judge. The high court was an awkward place from which to take a stand on current issues, attack opponents, or try to fit into the popular mood.” Moreover, Field, like other justices, was crippled in his presidential ambitions by the widespread belief that justices should not involve themselves in politics. As Kens has explained, “the sense of impropriety lingered,” despite Chase’s recent campaign for the presidency, and “it was employed against Field as the campaign progressed.” Accordingly, Field, like virtually all other Supreme Court justices who have sought the presidency, did not formally announce his candidacy. This did not, however, prevent Field from obtaining campaign funding. His brothers and other members of his wealthy family provided substantial financing, and Field also may have received funding from wealthy businessmen and financiers, including railroad barons Jay Gould and Leland Stanford. Field’s backers flooded the nation with copies of campaign biographies and various pamphlets touting his candidacy and they even started a Richmond newspaper to act as a cheerleader for Field in the South.
Field apparently hoped that southern support would propel his campaign, and participated in several civil rights cases during the months preceding the Democratic convention that enabled him to demonstrate his commitment to states’ rights and his disregard for the rights of African Americans. In two cases, he dissented from the Court’s decisions sustaining the convictions of state election officials who violated federal law by engaging in fraud to help Democratic candidates who sought to deprive African Americans of their civil liberties. In lengthy dissents, Field passionately argued that Congress had no power over these state officials even though the officials interfered with federal elections.

Several days later, Field again demonstrated his solicitude toward southern whites and the Democratic Party’s states’ rights principles in three cases involving the exclusion of African Americans from juries. In the most famous of these cases, Strauder v. West Virginia, Field and Nathan Clifford were the only justices who dissented from the Court’s opinions that deliberate exclusion violated the Fourteenth Amendment’s equal protection clause. Field contended that jury service was not protected by the equal protection clause because it was a political right rather than a civil right. Field and Clifford likewise dissented from a companion case in which the Court held that systematic exclusion of blacks from juries gave rise to an action under the Civil Rights Act of 1875. Field also vigorously championed the right of states to exclude blacks from juries in a long concurring opinion in Virginia v. Rives, in which the Court held that exclusion of blacks was constitutional so long as it was not deliberate. In that opinion, Field narrowly defined state action to exclude the actions of state officials who were not acting pursuant to the express authorization of the state. Professor Kens has concluded that “Field’s 1880 opinions definitely were intended to provide a boost to his political campaign,” although Kens acknowledges that it is unclear “[w]hether they were merely a political expedient or part of a broader theory of jurisprudence.”

---

152. See Ex parte Siebold, 100 U.S. 371 (1880); Ex parte Clarke, 100 U.S. 399, 404-22 (1880).
153. Ex parte Siebold, 100 U.S. 371 (1880); Ex parte Clarke, 100 U.S. 399, 404-22 (1880).
154. Ex parte Siebold, 100 U.S. 371 (1880); Ex parte Clarke, 100 U.S. 399, 404-22 (1880).
155. See infra notes 156-160.
157. Id.
158. Ex parte Virginia, 100 U.S. 339 (1880).
160. Id. at 324-38.
Field’s well-timed opinions in these cases involving the civil rights of blacks appear to have helped Field win support in the South. It is therefore ironic that his various decisions protecting the rights of Chinese workers in the West may have cost him critical support in his home state of California, where opposition to rising Chinese immigration was a potent political issue. One biographer has explained that “it is commonly thought that Field’s Chinese decisions doomed his bid for the presidency.”

In particular, many opponents of Chinese immigration were infuriated by Field’s 1879 opinion as a circuit judge nullifying a San Francisco ordinance that directed jailors to crop the hair of prisoners, a law presumably intended to humiliate Chinese men who for religious reasons wore their hair in long braids. Field contended that the discriminatory ordinance violated a treaty between the United States and China and that it deprived Chinese of substantive due process in violation of the Fourteenth Amendment. Outrage over Field’s opinion was exacerbated by his position in two earlier cases. In an 1875 U.S. Supreme Court decision, Field had voted with the majority to grant habeas corpus to a Chinese woman who challenged the constitutionality of a state statute permitting the state commissioner of immigration to deny entry to foreigners on vessels coming into the state if he believed that they were likely to become public charges or if they were diseased or debauched. Miller’s opinion for the Court held that the state had intruded upon federal power over international relations. In a decision as a circuit judge the previous year, Field had granted habeas corpus to another Chinese woman on the ground that the statute interfered with a treaty with China and with congressional power to regulate commerce with foreign nations.

Field apparently hoped, in vain, that his ardent opposition to Chinese immigration would counteract the unpopularity of these decisions. Field declared in one of his campaign statements in 1880 that he always had “regarded the immigration of Chinese in large numbers into our state as a serious evil, and

162. Hailing Field for carrying on the states’ rights tradition of Thomas Jefferson, The Washington Post praised his opinions as “like a bugle blast of the Monticello statesman.” Id. at 188. Meanwhile, the Richmond State reported that “Democrats are delighted with the able and exhaustive opinions of Justices Field and Clifford” and that several Democratic congressmen believed that they would provide Field with “a tremendous boom for the presidency.” Id. at 216. As another biographer pointed out, “[i]t was a discouraging task” for Field’s supporters “to play up the qualifications of a man whose own state was against him.” Swisher, supra note 138, at 292.

163. See e.g., Ah Kow v. Nunan, 12 Fed. Cas. 252 (C.C.D. Cal. 1879); Chy Lung v. Freeman, 92 U.S. 275 (1875).

164. Id. at 256-57.


166. Id. at 280-81.
likely to cause great injury to the morals of our people,” and that Californians had a duty to “defend American institutions and republican government from the oriental gangrene.”

In a possible reference to his judicial decisions, however, he warned that “no good can come of a resort to small vexations against the Chinese.”

Field also aroused antagonism in California by supporting the position of the Union Pacific and Central Pacific railroads in a dissent from an 1879 decision in which the U.S. Supreme Court sustained the constitutionality of a federal statute that unilaterally restructured the process by which the railroads repaid their debt to the government for the loans that they received in connection with their construction of the first transcontinental railroad. Although Field largely based his dissent upon the same sanctity-of-contract theories that he espoused in so many other cases during his long tenure on the Court, his rhetoric was so extravagant that critics then and now have concluded that his opinion was calculated to stimulate political support among powerful corporate interests. Field, also, however, may have written his opinion with at least one eye toward populist sentiment insofar as he also warned that “[t]he law that protects the wealth of the most powerful, protects also the earnings of the most humble.”

As Kens has observed, Field’s characteristic “penchant for sweeping language and visionary reasoning” make it “difficult to say where in the Sinking Fund Cases his judging may have ended and his campaigning begun. Nevertheless, several long paragraphs at the end of his dissent bore the markings of a campaign speech.” Similarly, many contemporary commentaries castigated Field for allowing politics to affect his dissent. As The San Francisco Examiner complained, “[t]he prurient desire for the presidency is a bee in a bonnet, and wildly afflicts the head it attacks. No man who aspires to the exalted station ought ever to be upon the Supreme Bench.” Field, however, had little need to assure corporate interests of his support, and his opinion clearly antagonized liberals and workers who opposed the monopolistic practices of railroads and other corporations.

Because Field’s opinions in the cases involving African Americans, Chinese immigrants, and railroads were consistent with his opinions in earlier and later...
cases, there is little reason to suppose that Field’s presidential ambitions changed his votes.\(^{180}\) Professor Kens believes, however, that Field’s ambitions affected the manner in which he wrote his opinions and that his “campaign efforts undoubtedly had a long-term effect on constitutional law” involving state action, equal protection, and the rights of corporations.\(^{181}\)

Field’s supporters arrived with much fanfare at the Democratic national convention in Cincinnati in June 1880, touting their candidate with newspaper advertisements, bright banners, horse-drawn wagons, and a brass band in a campaign that was much more visible and colorful than that of any other candidate.\(^{182}\) At the convention, his name was placed in nomination but he never ranked higher than fifth among the candidates, receiving only sixty-five of the 728 votes on the first ballot and sixty-five and a half on the second ballot before the convention stampeded to Winfield S. Hancock.\(^{183}\)

Field conducted no serious campaign in 1884, although his rousing support for substantive due process in his concurring opinion in *Butcher’s Union Co. v. Crescent City*,\(^{184}\) two months before the Democratic national convention, appears to have been calculated to win political support.\(^{185}\) Professor Fairman has observed that this paean to liberty of contract, along with Field’s implied criticism of the Fourteenth Amendment’s imposition of legal disabilities on some former Confederates, “expressed just the sentiment that Democrats delighted to hear.”\(^{186}\) Field’s decisions protecting the rights of the Chinese and the prerogatives of the powerful railroads, however, continued to rankle the dominant populist sentiment of the California Democratic Party.\(^{187}\) As Kens has explained, “with anti-monopolists in control of the party, Field’s presidential aspirations were resoundingly crushed before his campaign even left California.”\(^{188}\) One month before the Democratic national convention, the state convention formally repudiated Field’s “presidential aspirations” by a vote of 453 to 19.\(^{189}\) Moreover, at age sixty-eight, Field was perhaps too old by the

\(^{180}\) See McCloskey, *supra* note 131 at 529-550.

\(^{181}\) Kens, *supra* note 138 at 232. As Kens explains, Field’s narrow construction of the state action doctrine provided a constitutional basis for political opposition to the rights of African Americans. *Id.* Field’s opinions in the Chinese cases, according to Kens, “defined both class and privilege in a new way that eventually would give federal courts oversight of most types of legislation.” *Id.* at 233. And his *Sinking Funds Case* opinion, in Kens’s view, “with its emphasis on economic rights, the limits of state power, and an introduction of the notion that corporations had the same rights as persons – gave some sense of the direction judicial oversight of economic regulation would take.” *Id.*


\(^{183}\) *Id.* at 298-99.

\(^{184}\) Butcher’s Union Co. v. Crescent City, 111 U.S. 746, 754-60 (1884).

\(^{185}\) Fairman, *Reconstruction and Reunion*, *supra* note 107, at 1382.

\(^{186}\) *Id.* at 1384.

\(^{187}\) See Kens, *supra* note 138, at 234.

\(^{188}\) *Id.* In advance of the national Democratic convention, California’s state Democratic convention expressed explicit opposition to Fields’s nomination. *Id.*

\(^{189}\) Swisher, *supra* note 138, at 304-07.
standards of nineteenth century presidential candidates.\textsuperscript{190} Field later privately admitted to a friend that he regretted that he did not openly seek the presidency earlier in his life.\textsuperscript{191}

Another of the most politically ambitious justices was Samuel Freeman Miller, another Lincoln appointee who was mentioned as a possible presidential nominee in several elections during the 1870s and 1880s, and who was active as a candidate at least once.\textsuperscript{192} Although Miller never held elective office, he was an unsuccessful Republican candidate for the state senate in Iowa and later made a failed attempt to obtain the Republican Party’s gubernatorial nomination in that state.\textsuperscript{193}

Miller’s presidential ambitions were stoked by his growing popularity among critics of the untrammeled capitalism of the Gilded Age, who admired Miller’s judicial decisions favoring populist causes as the Court grew more conservative during the years following the Civil War.\textsuperscript{194} As a recent biographer has written, “Miller, with his critical view of eastern financiers, became a minor folk hero in the West.”\textsuperscript{195} Some liberal Republicans who had grown disillusioned with President Grant hoped that Miller would challenge Grant for re-election in 1872, but Miller rejected their overtures because he remained supportive of the President.\textsuperscript{196} Miller’s friends issued a statement in March 1872 that denied he was a candidate for president and emphasized that Miller believed that his position on the Court demanded the subordination of all political ambition.\textsuperscript{197}

In 1880, Miller neither encouraged nor discouraged a plan by Iowa Republicans to nominate him as a compromise candidate if the Republican convention deadlocked between Grant and Senator James G. Blaine.\textsuperscript{198} Before Miller’s advocates were able to generate enough support to place his name into nomination, the Wisconsin delegation, which had leaned toward Miller, voted for James Garfield, starting a steamroller that ended in Garfield’s nomination.\textsuperscript{199}

Finally, in 1884, Miller appears to have encouraged his own candidacy, although by then he was sixty-eight years old and he was concerned that he

\begin{itemize}
\item \textsuperscript{190} McCloskey, \textit{supra} at 542.
\item \textsuperscript{191} Swisher, \textit{supra} note 138, at 318.
\item \textsuperscript{192} See William Gillette, \textit{Samuel Miller, in 2 The Justices of the United States Supreme Court, Their Lives and Major Opinions} 504, 504-18 (Leon Friedman and Fred L. Israel eds., 1995) (discussing the life and career of Samuel Miller).
\item \textsuperscript{193} Fairman, \textit{Mr. Justice Miller, supra} note 74, at 28-32, 36-37.
\item \textsuperscript{194} See Michael A. Ross, \textit{Justice of Shattered Dreams: Samuel Freeman Miller and the Supreme Court During the Civil War Era} 235-37 (2003).
\item \textsuperscript{195} \textit{Id.} at 236.
\item \textsuperscript{196} \textit{Id.}
\item \textsuperscript{197} Fairman, \textit{Mr. Justice Miller, supra} note 74, at 297 (quoting \textit{N.Y. Times}, Mar. 29, 1872).
\item \textsuperscript{198} Ross, \textit{Justice of Shattered Dreams, supra} note 194, at 237.
\item \textsuperscript{199} \textit{Id.; Fairman, Mr. Justice Miller, supra} note 74, at 302.
\end{itemize}
would suffer financially if he resigned from the bench and lost the election. He received the ardent support of many western Republicans, who hoped that Miller “could reclaim the soul of their party from corporate conservatives,” as well as support from some eastern conservatives who believed that he could unite various Republican factions. Miller’s hopes had been raised in 1883 when Blaine told Miller that party leaders favored his nomination because he was the only Republican who was strong among all wings of the party. Like so many other justices, Miller professed to be content on the Court, but admitted that he would not refuse the nomination.

After Blaine again became a candidate, at least some supporters of Miller and Blaine agreed that Miller’s supporters would support Blaine, and that Blaine’s delegates would throw their support to Miller if Blaine’s candidacy collapsed, as it had in 1876 and 1880. This agreement became moot when Blaine received the nomination on the fourth ballot. According to one historian, Blaine never actually intended to throw his support to Miller or any minor candidate if he could not himself be nominated. In correspondence shortly after the convention, Miller revealed that he would have been willing to accept the nomination if, as he believed, Blaine would have thrown his support to him if Blaine could not have been nominated. Miller, who had refused to allow his supporters to place his name before the convention, explained that his candidacy “had but that one chance and it was not to be frittered away. No one can say now that I have ever sought the place or brought reproach or folly to the judicial ermine.” He suggested, however, that he had taken satisfaction in the honor of his candidacy, which he described as “the nicest and quietest little scheme.”

A fourth Lincoln appointee, David Davis, was also involved in presidential politics during his years on the Court, although his level of activities and expectations, if not his hopes, were less than those of his brethren Chase, Field, and Miller. Davis, a Republican activist, had helped secure the Republican nomination for Lincoln. After the National Labor Reform Committee nominated Davis for the presidency at its convention in February 1872, Davis

---

200. ROSS, JUSTICE OF SHATTERED DREAMS, supra note 194, at 237.
201. Id. at 237.
202. FAIRMAN, MR. JUSTICE MILLER, supra note 74, at 303 (quoting letter from Miller to William Pitt Ballinger, May 9, 1883).
203. ROSS, JUSTICE OF SHATTERED DREAMS, supra note 194, at 238.
204. Id.
205. Id.
206. FAIRMAN, MR. JUSTICE MILLER, supra note 74, at 306.
207. ROSS, JUSTICE OF SHATTERED DREAMS, supra note 194, at 238.
208. FAIRMAN, MR. JUSTICE MILLER, supra note 74, at 306 (quoting letter from Miller to William Pitt Ballinger, June 15, 1884).
209. Id.
210. Id. at 293-96.
211. See id. at 293.
responded with a cryptic assertion that the presidency should neither be sought nor declined by any citizen. Two months later, Davis’s friends attempted to secure his nomination by the Liberal Republican Party, but he ranked only fifth on the first ballot and his support fizzled on later ballots, as the number of votes for Chase increased.

Later in 1872, Davis became the only sitting justice ever to receive a presidential electoral vote, obtaining one vote in Missouri. The state was ultimately won by Greeley, who died after the election but before the electoral votes were cast. Since Davis resigned from the Court in 1877 after the Illinois legislature elected him to the United States Senate, he presumably would have had little compunction about surrendering his Supreme Court seat for the presidency. In 1880, Davis was a “dark horse” candidate for the presidency.

Like other presidential aspirants on the Court, Davis provoked criticism. A famous Thomas Nast cartoon, published on the cover of Harper’s Weekly in April 1872, mocked the presidential ambitions of Chase and Davis. Labeled “The Presidential Fever on the Supreme Bench,” it depicted a chastened Chase admonishing a fidgety Davis to “fling away” the ambition that had so embarrassed Chase.

During the same years that Chase, Field, Miller, and Davis entertained presidential ambitions, Justice Samuel Nelson, an older justice who had served since 1845, also may have harbored presidential hopes. Nelson downcast efforts by friends to encourage him to seek the Democratic nomination in 1868 not from fear of diluting the Court’s integrity, but from his own lack of interest and fear that he would be a mere pawn of party leaders. At age 76, however, it is unlikely that Nelson would have attracted much support even if he had wanted the nomination.

C. The Late Nineteenth Century

At least three other members of the Court during the late nineteenth century, Chief Justices Morrison R. Waite and Melville W. Fuller and Associate Justice
John Marshall Harlan, were prominently mentioned as presidential candidates.\textsuperscript{223} Unlike several other justices of this era, however, Waite, Fuller, and Harlan strongly discouraged such speculation.\textsuperscript{224}

Shortly after Waite became Chief Justice in 1874, newspapers throughout the eastern and mid-western states began to mention him as a possible Republican nominee in 1876, when Grant was not expected to seek a third term.\textsuperscript{225} Although Waite had little political experience,\textsuperscript{226} his political obscurity was an asset because, as one of his biographers has explained, “he was not plagued by a record of past commitments” and was not “tainted by the scandals which left a cloud of doubt hanging over the careers of so many public men in the 1870’s.”\textsuperscript{227} Waite also was attractive to many Republicans because he favored sound money and supported the Civil War amendments to the Constitution.\textsuperscript{228}

In various correspondences during 1875 and 1876, Waite disavowed any interest in becoming a presidential candidate.\textsuperscript{229} In November 1875, in a long letter to his nephew, who was a Republican congressman, Waite explained that “one who occupies this position should keep himself above suspicion . . . it is dangerous to have a judge who thinks beyond the judicial in his personal ambitions.”\textsuperscript{230} “The Court,” Waite observed, “is now looked upon as the sheet anchor. Will it be if its Chief Justice is placed in the political whirlpool?”\textsuperscript{231} Waite declared that the office of Chief Justice “came to me covered with honor. When I accepted it, my duty was not to make it a stepping stone to something else, but to preserve its purity . . . .”\textsuperscript{232} In another letter written on the same day, Waite declared that “I am one of those who know that a Chief Justice cannot be a candidate for the Presidency, without damaging the office he holds and himself too.”\textsuperscript{233} Similarly, Waite privately expressed his belief that such a candidacy would remove “at least one stone from the foundations that uphold the government.”\textsuperscript{234}

\textsuperscript{223} See infra Part II.C.
\textsuperscript{224} See infra Part II.C.
\textsuperscript{226} Waite had never served in elective office, aside from a brief stint on the Toledo City Council and a single term in the Ohio legislature. See LOUIS FILLER, MORRISON R. WAITE, in 2 THE JUSTICES OF THE UNITED STATES SUPREME COURT, THEIR LIVES AND MAJOR OPINIONS 610, 610-26 (Leon Friedman and Fred L. Israel eds., 1995) (discussing the life and career of Morrison R. Waite). He was twice an unsuccessful candidate for the U.S. House of Representatives. Id.
\textsuperscript{228} Id. at 283.
\textsuperscript{229} See, e.g., MAGRATH supra note 227, at 280 (quoting Waite to H.L. Bond, Nov. 7, 1875).
\textsuperscript{230} TRIMBLE, supra note 225, at 142 (quoting letter from Waite to John Turner Waite, Nov. 7, 1875).
\textsuperscript{231} Id.
\textsuperscript{232} Id.
\textsuperscript{233} MAGRATH, supra note 227, at 280 (quoting letter from Waite to H.L. Bond, Nov. 7, 1875).
\textsuperscript{234} Id. at 285, (quoting letter from Waite to Clark Waggner, Dec. 4, 1875).
Waite believed that he needed to unequivocally withdraw his name from contention, lest suspicions about his political ambitions discredit his judicial decisions. He believed that Chase had “detracted from his name by permitting himself to think he wanted the presidency” insofar as “[w]hether true or not, it was said that he permitted his ambitions...to influence his judicial opinions.”236 In addition to fearing that the political ambitions of a justice could diminish public faith in the integrity of the Court, Waite also believed that he would demean the Chief Justice position if he were an unsuccessful president. 237 Although Waite originally expressed these sentiments only in private, fearing that public statements would fan interest in his potential candidacy, Waite eventually allowed his letters to be published to quash growing speculation about his availability for the Republican nomination.238

Waite’s refusal to become a presidential candidate generated widespread plaudits. The Nation, for example, hailed Waite for “pointing out the gross impropriety of making the bench of the Supreme Court a stepping-stone to something else.”239 Waite’s resistance of political intrigue continues to inspire admiration. As Chief Justice William Rehnquist pointed out, “Waite felt that it would sully the office he held to aspire to any other one, and this firm determination on his part helped to restore the prestige of the Court”240 which was still recovering from the ignominy of its Dred Scott decision.

Waite’s steadfast refusal to become a presidential candidate was emulated by his successor, Melville W. Fuller. Although Fuller never held any elective office except for a brief stint in the Illinois legislature, his stature as chief justice helped to ensure that he, like so many other chief justices, would be mentioned as a possible presidential candidate.241 In spring 1892, four years after Grover Cleveland had appointed Fuller as Chief Justice, several of Fuller’s politically powerful friends in Chicago urged him to seek the Democratic nomination. 242 Although former President Cleveland was the front-runner, many Democrats feared that Cleveland’s quarrel with Tammany Hall and New York’s Governor Hill would prevent Cleveland from carrying the critical state of New York. 243 In contrast, Fuller had no apparent political enemies, and his mid-western ties could

---

235. See TRIMBLE, supra note 225, at 141-42.
236. Id. (quoting letter from Waite to John Turner Waite).
237. Id. at 142 (quoting letter from Waite to John Turner Waite).
238. MAGRATH, supra note 227, at 282-84.
239. TRIMBLE, supra note 225, at 146.
243. Id. at 162.
help the ticket. By May 1892, there were reliable reports that Democratic leaders had persuaded Cleveland to withdraw in favor of Fuller. When Fuller learned this, he immediately wrote to Cleveland to assure him that “you are the only nominee absolutely certain of success.”

Shortly after the 1892 election, Cleveland asked Fuller to resign from the Court to become Secretary of State. Although Fuller seems to have given the offer serious consideration, his stated reason for rejecting the offer closely resembled the reasons why other justices have refused to become presidential candidates. Fuller explained to Cleveland that “the effect of the resignation of the Chief Justice . . . would be distinctly injurious to the court. The surrender of the highest judicial office in the world for a political position, even though so eminent, would tend to detract from the dignity and weight of the tribunal. We cannot afford this.”

Felix Frankfurter once described this letter as “[a]n important document in the history of the judiciary, and . . . in the history of the law.”

Like Waite and Fuller, John Marshall Harlan seemed determined to avoid any appearance of presidential ambition. Although Harlan had been active in Kentucky politics before becoming a Justice, he never held any significant elective office other than state attorney general, and his presidential prospects appear to have emanated from his eminence as a justice. In 1888, when Harlan was mentioned as a presidential candidate, he was so fearful of encouraging speculation about his candidacy that he deliberately refrained from traveling to Chicago during the Republican convention there, even though he wished to meet in Chicago with counsel in a case in which he was presiding as circuit judge.

D. The Twentieth Century

The Supreme Court’s role as a recruiting ground for presidential candidates faded for two decades after Fuller resisted overtures in 1892. Speculation about
the nomination of a justice reemerged in 1912, when many Republicans urged the dynamic young Charles Evans Hughes to seek the presidency.254

As governor of New York in 1908, Hughes had been a candidate for the Republican nomination for president.255 In 1912, two years after Hughes became a justice, many Republicans hoped that the Republican Party could unite behind Hughes to avoid the continuation of the internecine battle between William Howard Taft and former President Theodore Roosevelt that paved the way for Wilson’s election.256 On the eve of the Republican convention, Taft privately admitted his willingness to stand aside for Hughes to prevent a rift in the party.257

When New York Republican leaders urged Hughes to allow the convention to nominate him, Hughes flatly refused to become a candidate.258 With pressure mounting on Hughes during the acrimonious convention, he issued a public statement declaring that “no man is as essential to his country’s well being as is the unstained integrity of the courts.”259 Hughes explained that a justice’s resignation from the Court to seek political office would impair judicial independence and public confidence in the courts because political parties would make partisan use of judicial decisions and judges might be tempted to render decisions to promote their political ambitions.260 In a letter to Elihu Root, the chair of the convention, Hughes explained that “[t]he highest service that I can render in this difficult situation is to do all in my power to have it firmly established that a Justice of the Supreme Court is not available for political candidacy. The Supreme Court must be kept out of politics . . . even if nominated, I should decline.”261 Observing Hughes from a close distance on the Court, however, Oliver Wendell Holmes was not so certain that Hughes was so averse to presidential ambition.262 In a private letter in April 1912, Holmes remarked that while “he says he won’t be a candidate unless the President asks him to be – or something like that,” Holmes had “noticed a certain reserve in his

255. Id. Although he refused to actively campaign, he did not disavow attempts to promote his candidacy by supporters who found his passivity increasingly frustrating. Id. Hughes refused to release New York delegates pledged to him even after Taft’s nomination had become inevitable. Id.
256. MERLO J. PUSEY, CHARLES EVANS HUGHES, VOL. I 300 (1951).
257. Id.
258. Id.
259. Id.
260. Id. at 300.
261. Id. at 301. After the convention, Root assured Hughes that “you took a wise and patriotic course” because “keeping the Supreme Court out of politics is more important than the Presidency.” Id.
speech that seems to leave possibilities open.” Holmes averred that “a judge should extinguish such thoughts when he goes on the Bench.”

Four years later, however, Hughes did not resist the call of his party. He accepted the Republican nomination and narrowly lost the election to Woodrow Wilson for reasons that appear to have little to do with any revulsion over his willingness to leave the bench. Historians generally have accepted Hughes’s protestations that his nomination was a genuine draft. As Hughes declared in his autobiographical notes,

> I did not wish to be nominated. In view of the war in Europe, the responsibilities of the Presidency were heavier than ever and, knowing well what they meant, I had no desire to have the burden placed on my shoulders. The idea of a Justice of the Supreme Court taking part in politics, promoting in the slightest degree his selection as a candidate for the office of President, was abhorrent to me. I strongly opposed the use of my name and the selection or instruction of any delegates in my interest.

Torn between what he described as “two profound desires, one to keep the judicial ermine unsullied, and the other not to fail in meeting what might be a duty to the country,” Hughes finally felt that he “could not refuse” his party’s call to serve his country.

Hughes left the Court with the blessings of his colleagues. Holmes privately wrote that “[i]t was not preference but simply and solely, as I believe, a sense of duty that led Hughes to accept the nomination.” Although Chief Justice White expressed fear that Hughes’s nomination was a great blunder from which it would take the Court years to recover, he nevertheless blamed the Republican Party rather than Hughes, and he encouraged Hughes to respond to his party’s call.

---

263. Id.
264. Id.
265. See Pusey, supra note 256, at 315-334 (detailing Hughes’ 1916 presidential nomination).
266. Since the election was so close that it ultimately hinged upon a few thousand votes in California, however, it may be possible that Hughes’s willingness to doff his judicial robes may have been one of many factors that explain his defeat. Complaining that Hughes’s candidacy threatened to corrupt the integrity and reputation of the Court, The New York Times interpreted Hughes’s defeat as “a guarantee that the freedom of the Supreme Court from playing to the galleries . . . is to be maintained.” The Supreme Court and the Presidency, N.Y. TIMES, Nov. 11, 1916, at 8.
267. See, e.g., Pusey, supra note 256, at 315-334.
269. Id. at 180-81.
270. Letter from Oliver Wendell Holmes to Frederick Pollock, July 12, 1916, in Howe & Palfrey, supra note 262, at 237.
271. Pusey, supra note 256, at 332.
Hughes’s resignation to accept the nomination at first generated understandable indignation. Deploring “the celerity with which . . . Hughes has ceased to be the Supreme Court Justice and become Busy Charlie the Candidate,” the New York Evening World observed that “[t]he nation has grown to think of its highest judicial tribunal as something stable and permanent, wherein men sat as with a final dignity befitting ultimate honor. The new view is something of a shock. Not all of us will get used to it – or even wish to.”

Within days of Hughes’s resignation, Democratic senators introduced resolutions for constitutional amendments to prevent Supreme Court justices from becoming candidates for public office within five years after leaving the Court. In a speech that coincided with Hughes’s acceptance of the nomination, Democratic Senator Thomas J. Walsh of Montana alleged that Hughes’s election would forever cause Americans to distrust the motives of federal judges.

Ultimately, however, the circumstances of Hughes’s resignation from the bench and his judicial record did not become campaign issues. Wilson refused to permit the Democratic platform to condemn Hughes for leaving the Court to become a candidate, and even many liberals refused to assail Hughes’s motives. As The Nation observed, “[n]o man can point to a single decision, a single gesture, of his on the bench and say that it was animated by political ambition.” This probably was wise because Democratic attacks on Hughes could have seemed hypocritical after Alton B. Parker had resigned from the New York Court of Appeals in 1904 to accept the Democratic presidential nomination. Moreover, the circumstances of Hughes’s resignation and his judicial record did not emerge as major campaign issues insofar as the principal issue was whether the United States should remain neutral in the First World War.

As one historian has concluded, the hazards that Hughes’s candidacy posed for the Court were “negotiated with singular success and luck.” There is little evidence that Hughes’s transition from Supreme Court justice to presidential candidate significantly exacerbated hostility toward the Court among

276. Id.
278. Ross, A Muted Fury, supra note 275, at 162.
progressives and trade union leaders who objected to its decisions striking down economic regulatory legislation, or that it encouraged the burgeoning movement to curb the Court’s powers. Although some opponents of Hughes’s nomination to the chief justice position in 1930 complained about his resignation from the Court in 1916 to run for the presidency, this was not a major issue in the substantial controversy over his fitness for the chief justice position. 282

After Hughes’s resignation from the Court, another two decades passed with little or no speculation about presidential ambition among the justices. Although the relatively scant controversy over Hughes’s nomination might have encouraged such ambition, the men who served on the Court during these years generally were too old to aspire to the presidency or lacked any political constituency. Perhaps the most obvious choice during this period would have been Brandeis, who was politically involved during his years on the Court even though he never had held public office, and who was held in high esteem by progressive activists. In 1924, Brandeis was Senator Robert M. LaFollette’s first choice as his vice presidential running mate in LaFollette’s bid for the presidency as an independent progressive. Even though Brandeis rejected LaFollette’s invitation, he may have given it at least some consideration.

The Court re-emerged as a potential recruiting ground for presidential candidates during the middle 1930s and remained so for the next two decades. Two Justices, Harlan F. Stone and Owen Roberts, were widely touted as Republican presidential prospects after the G.O.P. debacle in the 1934 mid-term elections, which swept many Republican congressmen and governors from office. Bereft of any obvious front-runner, various Republican elders who hoped to nominate an experienced and prestigious candidate set their sights on these two distinguished members of the Court. Although both justices

---

281. Ross, A Muted Fury, supra note 275, at 163.
283. See 3 The Justices of the United States Supreme Court, Their Lives and Major Opinions 969, 969-1098 (Leon Friedman and Fred L. Israel eds., 1995) (examining the lives and careers of Supreme Court Justices who served on the Court between 1916-1926).
284. See id.
285. Ross, A Muted Fury, supra note 275, at 255.
286. Id. at 255 (citing Belle Case LaFollette and Fola LaFollette, Robert M. LaFollette, vol. 2 1115 (1953)). Predicting that LaFollette would mount “a grand fight,” Brandeis privately acknowledged that “[i]f I had several watertight compartment lives, I should have been liked to be in it. The enemies are vulnerable and the times ripe.” Melvin Urofsky and David W. Levy, eds., Letters of Louis D. Brandeis, vol. 5 134 (1978) ( Louis D. Brandeis to Alfred Brandeis, July 19, 1924).
288. Id. See Lippmann, supra note 279, at 278.
289. See Leuchtenburg, supra note 287, at 43; see Lippmann, supra note 279, at 278.
consistently disclaimed any presidential ambitions, it is not clear that these jurists, especially Roberts, would have resisted a draft.290

The hazards of plucking a candidate from the Court, however, has rarely been more acute than it was in 1936, because the Court itself was involved in political controversy that was as intense as at any time in the Court’s history.291 Although public respect for the judiciary made the Democrats loath to openly criticize the Court during the 1936 presidential campaign, the nomination of Roberts by the Republicans almost certainly would have transformed the Court’s decisions and its institutional powers into an election issue.292 Stone’s nomination probably would have had much the same effect, although Stones’ votes on the Court to sustain the constitutionality of New Deal legislation made his nomination by Republicans improbable.293

Because the Court’s decisions striking down New Deal legislation during 1935 and 1936 had generated numerous proposals to curb the Court’s powers, the resignation of a justice to run for the presidency would not have occurred at a worse time for the Court.294 The candidacy of any former justice would have bolstered public support for the multitude of Court-curbing proposals that percolated in Congress and in academe by enhancing the perception that the Court was unduly political.295 Assuming that Roosevelt had defeated Roberts or Stone, the politicization of the Court during the campaign might have encouraged Roosevelt to propose legislation to restrict judicial review rather than the measure to pack the Court that he proposed three months after his re-election.296 Because the ultimate defeat of Roosevelt’s Court packing plan was not inevitable,297 the candidacy of either Roberts or Stone might have provided the additional anti-Court animus that would have enabled it to succeed. The controversy over the Court packing measure must have discouraged Republican leaders from encouraging Roberts and Stone to seek the presidency in 1940, because Republicans in opposing the Court packing bill had complained so vociferously about Roosevelt’s attempt to politicize the judiciary.298 Even in politics, there are limits to hypocrisy.

Roberts’s principal allure was that he was widely regarded as a moderate who could appeal to both moderate and conservative factions in the G.O.P., insofar as he had been the “swing vote” in many of the Court’s rulings on the constitutionality of social and economic regulatory legislation since he became a

290. See Leuchtenburg, supra note 287, at 43-45.  
291. See id. at 43-49.  
292. Id. at 43.  
294. See Leuchtenburg, supra note 287, at 49-51.  
295. See id.  
296. See id.  
297. See id.  
298. See id.
justice in 1930.\textsuperscript{299} As the columnists Drew Pearson and Robert S. Allen pointed out shortly after the 1936 election, interest in Roberts also stemmed from “[h]is handsome, virile appearance, his impressive record as Teapot Dome prosecutor,” and “the favorable geographic position” of his home state, Pennsylvania, which at that time commanded more electoral votes than any state other than New York.\textsuperscript{300} Although Roberts never had held elective office,\textsuperscript{301} his very lack of political experience may have been an asset because Republicans could present him as a non-politician who was above the fray of petty politics.

Roberts’s status as a justice also boosted his appeal insofar as many Republicans wanted to generate and exploit public indignation over what they perceived as the constitutional depredations of the New Deal.\textsuperscript{302} As Walter Lippmann explained in a column in June 1935, “[t]he idea is that Mr. Justice Roberts, having decided against New Deal measures, is to run as the savior of the Constitution.”\textsuperscript{303} Lippmann pointed out that this was “certain to cause acute embarrassment to the Supreme Court as a whole and to Mr. Justice Roberts in particular,” especially because Roberts would continue “to sit in judgment upon much of the New Deal.”\textsuperscript{304} Praising Roberts as “a man of the highest judicial integrity and sensibility,” Lippmann warned that “if the boom is not destroyed first by responsible Republican leaders, his choice will be to destroy it himself or to resign from the bench.”\textsuperscript{305} Writing at a time when widespread dismay over the Court’s invalidation of New Deal legislation had generated a plethora of proposals to curb the Court’s power of judicial review, Lippmann declared that “[n]ow is a very good time to put an end once and for all to the idea that justices of the Supreme Court are available candidates for public office. This would make a very good first plank in a program to defend the Constitution and the Supreme Court as its final interpreter.”\textsuperscript{306}

In 1954, nine years after his retirement from the Court, Roberts publicly disavowed having had any interest in the presidency.\textsuperscript{307} Testifying before the Senate Judiciary Committee in favor of a constitutional amendment that would have barred members of the Court from serving as president or vice president within five years of termination of their service on the Court, Roberts acknowledged that “enthusiastic friends” had importuned him to become a

\textsuperscript{300} Pearson & Allen, supra note 293, at 161.
\textsuperscript{301} See David Burner, Owen J. Roberts, in 3 The Justices of the United States Supreme Court, Their Lives and Major Opinions 1117, 1117-29 (Leon Friedman and Fred L. Israel eds., 1995) (examining the life and career of Owen J. Roberts).
\textsuperscript{302} Leuchtenburg, supra note 287, at 43; Baker, supra note 299, at 123-24.
\textsuperscript{303} Lippmann, supra note 279, at 278.
\textsuperscript{304} Id. at 279.
\textsuperscript{305} Id. at 279.
\textsuperscript{306} Id. at 279-80.
candidate twice, presumably in 1936 and 1940. “Of course, I turned a hard face on that thing,” Roberts testified; “I never had the notion on my mind.” Although Roberts expressed hope that he would be excused from “naming names,” he explained that “it is a matter of common knowledge that ambition to go from the Court to the Chief Executive...has hurt the work of a number of men on the Court.” Roberts testified that a number of chief justices and associate justices

have had in the back of their minds a possibility that they might get the nomination for President. Now, that is not a healthy situation because, however strong a man’s mentality and character, if he has this ambition in his mind it may tinge or color what he does, and that is exactly what the Founding Fathers wanted to remove from the minds of the Supreme Court, to make them perfectly free knowing that there was no more in life for them than the work of the Court.

Pearson and Allen, however, reported that “Roberts considered the possibilities most seriously,” and that he ultimately rejected overtures from party leaders because his wife, Elizabeth Rogers Roberts, was unwilling to risk the loss of her cherished role as a leading Washington socialite. An ardent opponent of the New Deal, she also reportedly told members of her family that her husband should not hazard his opportunity to continue to stabilize the country. Moreover, both Roberts and his wife may have felt that his position as the swing vote on the Court made him nearly as powerful as the President.

Unlike Roberts, Stone appears to have been sincere in his renunciation of political ambition. Like Roberts, Stone had never held elective office. Before his appointment to the Court, Stone had served briefly as attorney general and as dean and professor at Columbia Law School. His easy-going, bluff personality seems to have been more suited to the contemplative life of academia and the judiciary than to the hurly burly life of politics. Stone does not appear to have had any particular desire to be President, although he maintained

308. Id.
309. Id.
310. Id. at 8.
311. Id. at 9. Similarly, Roberts testified before the House Committee on the Judiciary that “I think it is much better for a judge not to have any thought that he would ever be anywhere except where he sits, in the Supreme Court. Hearing Before Subcomm. No. 4 of the House Comm. on the Judiciary on S.J. Res. 44, H.J. Res. 194, H. J. Res. 27, and H.J. Res. 91, 83rd Cong., 2d Sess.22 (1954).
312. PEARSON & ALLEN, supra note 293, at 161-62.
313. Id. at 162.
314. Id. at 161.
316. Id. at 1100.
317. See id. at 1099-1116.
personal friendships with three Presidents, the strange trio of Calvin Coolidge, Herbert Hoover, and Franklin Roosevelt. Although he had no compunction about providing private advice to these presidents on sundry affairs of state, Stone had an especially ardent belief that judges should avoid public extra-judicial activities that distracted them from their judicial duties and cast doubts on their objectivity.

As early as 1927, two years after Stone became an associate justice, at least one of Stone’s friends, Sterling Carr, privately suggested that he should set his sights on the presidency. In reply, Stone expressed reluctance to leave a position for which he was “supremely interested.” In 1929, when President-elect Hoover invited Stone to become Secretary of State, Carr admonished Stone that he was “too young, too energetic, too full of leadership . . . to remain in the cloister.” Stone replied, “I owe something to the institution of which I am a part, and there are battles to be carried on here and won, as elsewhere.”

During 1935 and 1936, many liberal Republicans urged Stone to become a candidate for the G.O.P. presidential nomination. The chair of New York’s Republican Party advocated Stone’s candidacy, the popular columnist Mark Sullivan declared that “[t]here is not . . . a better man for President than Justice Stone,” and the editor William Allen White publicly proclaimed that “Stone would make an ideal candidate.” Stone steadfastly resisted their encouragement, explaining to one friend who urged his candidacy that justices “should keep out of politics.” Stone refused even to address the American Bar Association at its 1935 convention because he feared that his appearance would stimulate speculation that he would become a candidate. In 1940, many Republicans again urged Stone to become a candidate, but Stone adamantly refused to seek the nomination.

Franklin Roosevelt deliberately appointed highly political persons to the Court to help ensure that they would provide judicial support for his New Deal policies. It therefore is not surprising that several of Roosevelt’s nominees

318. See id.
319. See id.
321. Id.
322. Id. at 268.
323. Id. at 269.
324. See id. at 402-04.
325. See id. at 402-04.
327. Mason, supra note 315, at 402.
328. Id. at 705.
would harbor presidential ambitions or would be mentioned as possible candidates. At least five – Hugo Black, James F. Byrnes, William O. Douglas, Robert H. Jackson, and Stanley Reed – were mentioned to some extent as potential nominees.

Jackson was mentioned as a candidate only before he became a justice.\[330\] Shortly after Jackson became U.S. Attorney General in January 1940, he was prominently touted by various journalists as a contender for the 1940 Democratic presidential nomination.\[331\] According to various press reports, Jackson was Roosevelt’s favorite choice as a successor.\[332\] Jackson did not discourage such speculation, but he made no serious effort to seek the nomination and whatever hopes he entertained were dashed when Roosevelt decided to seek a third term.\[333\] In planning his bid for a third term, Roosevelt may have been instrumental in preventing the selection of Jackson to serve as a delegate to the Democratic national convention from New York.\[334\] Jackson, who was regarded as a potential vice presidential nominee during the Democratic convention, made a series of public speeches during the autumn of 1940 in which he vigorously urged Roosevelt’s re-election.\[335\] After he became a justice in 1941, Jackson was not prominently mentioned as a presidential candidate despite his stature.\[336\]

In addition to considering Jackson as a running mate in 1940, Roosevelt also may have given at least passing thought to the selection of Reed, whom he had appointed to the Court in 1938.\[337\] Roosevelt ultimately did not select Reed, apparently because Reed lacked a vibrant public personality.\[338\]

Black, who had served as a senator from Alabama for a decade before Roosevelt appointed him to the Court in 1937,\[339\] appears to have had presidential aspirations. During the year before Roosevelt decided to seek a third term in 1940, Black apparently considered becoming a presidential candidate if Roosevelt did not run again, and for a brief time in 1940 he may even have thought about challenging Roosevelt for the Democratic nomination.\[340\] After Roosevelt obtained the nomination, Black quietly let


\[332\] Id. at 199, 201.

\[333\] Id. at 199, 201, 204-05.

\[334\] Id. at 205.

\[335\] Id. at 205, 206-09.

\[336\] See KURLAND, supra note 330, at 1301-11 (examining the life and career of Robert H. Jackson after he became a Justice).


\[338\] Id. (quoting letter from Josephine Black to Hugo Black).

\[339\] See JOHN P. FRANK, Hugo L. Black, in 3 THE JUSTICES OF THE UNITED STATES SUPREME COURT, THEIR LIVES AND MAJOR OPINIONS 1152, 1159-63 (Leon Friedman and Fred L. Israel eds., 1995) (examining Justice Black’s career as an Alabama senator).

\[340\] NEWMAN, supra note 337, at 306-07.
prominent Democrats know that he was available for the vice presidential nomination, and he even wrote to Roosevelt in July 1940 to offer to resign from the Court “should you conclude you need my services,” presumably as a running mate.341

In 1944, when Vice President Henry A. Wallace’s increasingly erratic personal and political behavior jeopardized his chances of re-nomination as Roosevelt’s running mate, Black angled again for the vice presidency and received encouragement from some prominent Democrats. Former Roosevelt advisor Thomas Corcoran went so far as to assure Black that “[w]e’ve got labor and the liberals lined up for you” and that “[t]he nomination is yours if you want it.”342 Black reportedly declined Corcoran’s offer of support because of his wife Josephine’s poor health.343 However, according to Roger Newman, one of Black’s biographers, Black “never lost his desire to be president.”344

Early in 1948, when President Harry Truman’s prospects for election were bleak and there was widespread uncertainty about whether he would even run, Black discreetly contacted political leaders in his home state of Alabama to determine whether they would support him if he sought the presidency.345 As Newman explained, “[t]he odds were small, but Black’s dream was too strong not to try . . . The responses were uniformly negative.”346 Although some prominent Alabama liberals, particularly Senator Lister Hill, believed that Black would have made an excellent president, conservative forces in the state made his candidacy impossible.347

Black recognized that his brief membership in the Ku Klux Klan in Alabama during the 1920s impeded his presidential prospects.348 His association with the Klan created a furor when it was publicly revealed shortly after his nomination as a justice in 1937, and his opponents would have used it as a political weapon against him if he sought either the presidency or the vice presidency.349 Black’s candidacy also might have encountered indifference if not hostility from liberals and labor leaders who appreciated Black’s support for many of their causes, but who remained suspicious of his rural populist roots.350 Black, therefore, may have hoped that his many decisions in support of the rights of African Americans and organized labor would enhance his presidential prospects.351 Because there is powerful evidence that Black’s decisions were motivated by long-standing

341. Id. at 307.
342. Id. at 309.
343. Id.
344. Id. at 306.
345. Id. at 383.
347. Id. at 383-84.
348. Id. at 307.
349. Id.
351. See id.
convictions, there is no reason to suppose that his political ambitions affected his votes.\footnote{352}  His presidential aspirations, however, may occasionally have influenced the way in which he wrote majority or dissenting opinions, or perhaps encouraged him to wangle writing assignments.\footnote{353}  

After Hughes, William O. Douglas was the only justice who became a significant contender on presidential tickets. Roosevelt appears to have seriously considered offering Douglas the vice presidential nomination in 1940 and 1944, and Truman invited Douglas to run with him in 1948.\footnote{354}  Douglas also was prominently mentioned as a presidential candidate in 1948 and 1952.\footnote{355}  Although Douglas had never held elective office, his keen political instincts and interests and his youth during his early years on the Court – he was only forty when appointed – made him in many ways a natural prospect for the presidency.\footnote{356}  

One of Douglas’s biographers probably is correct that Douglas “cared much more about politics than about his work on the Court” within a year of his appointment in 1939.\footnote{357}  Nonetheless, Douglas made no apparent effort to seek the 1940 Democratic nomination for vice president even though he was widely mentioned as a possible running mate in Roosevelt’s bid for a third term, after it became clear that Vice President John Nance Garner would not be on the ticket a third time.\footnote{358}  Douglas privately disclaimed any interest in the nomination.\footnote{359}  In a letter to Felix Frankfurter, he described talk about his candidacy as “disturbing because I want none of it,” urging Frankfurter to help “scotch it” because “I want to stay where I am.”\footnote{360}  To a friend, he insisted that “I have no political ambitions.  My sole desire is to give a life time of service here in an endeavor to fill the shoes of my illustrious predecessor,” Brandeis.\footnote{361}  After Roosevelt reportedly told his advisors shortly before the Democratic convention that Douglas and Secretary of Agriculture Henry A. Wallace were becoming his favorite choices, Democratic leaders urged Wallace’s selection because Douglas lacked a political base and experience as a campaigner.\footnote{362}  After Roosevelt selected Wallace, Douglas expressed relief that the nomination passed him by.\footnote{363}
Douglas came much closer to the vice presidency four years later, when Roosevelt expressed enthusiasm for Douglas after Democratic leaders convinced Roosevelt that Democrats needed to deny the vice presidential nomination to the increasingly left-wing and eccentric Wallace. 364 A week before the national convention, Roosevelt mentioned Douglas as a potential running mate, and there is evidence that he preferred Douglas to Missouri Senator Harry Truman, who was emerging as the favorite of party leaders. 365 When Democratic National Chairman Robert Hannegan and other party leaders expressed their preference for Truman at a meeting with Roosevelt on July 11, Roosevelt gave Hannegan a handwritten note stating that he would be pleased to run with either Douglas or Truman, in that order. 366 Four days later, at the request of Hannegan, who argued that Douglas was not a viable candidate, Roosevelt re-wrote his note to place Truman’s name first. 367

Hannegan kept the letter private until the eve of the balloting in order to prevent the Douglas forces from organizing. 368 This was wise because Douglas’s nomination would have been difficult to arrange without the open candidacy of Douglas, who probably would have felt obliged to resign from the Court if he had publicly sought the nomination. He would not have done so without believing that his chances were excellent. 369 Although Douglas’s supporters worked vigorously behind the scenes at the convention to woo delegates, the release of the letter produced a stampede for Truman, and Douglas won only two votes. 370 Although Douglas later disavowed having had any interest in the nomination, there is considerable evidence that he actively encouraged efforts to secure his selection. 371 Assured by his political allies that they could secure his nomination, Douglas appears to have been keenly disappointed, particularly because he knew that Roosevelt’s health was so poor that he had little chance of surviving a fourth term. 372

365. Id. at 348; Edward L. & Frederick H. Schapsmeier, Prophet in Politics: Henry A. Wallace and the War Years, 1940-1965 104 (1968). Roosevelt told Hannegan and Edwin C. Pauley, the treasurer of the Democratic National Committee, that he believed that Douglas had greater public appeal. Id. (citing authors’ interview with Pauley on Sept. 3, 1963).
367. Id. at 275-77. The note stated that: “You have written me about Harry Truman and Bill Douglas. I should, of course, be very glad to run with either of them and believe that either one of them would bring real strength to the ticket.” Id. at 276. After Roosevelt wrote the note in longhand, his secretary typed it on White House letterhead and the President signed it. Id.
368. Murphy, supra note 357, at 223.
369. Id. at 228.
370. Id. at 228-29.
371. Id. at 228-32.
372. Id. at 218, 228-32.
Despite strong evidence to the contrary, Douglas claimed in his memoirs that “I never had the Potomac Fever and could not be excited about catching it.”\textsuperscript{373} Douglas contended that Roosevelt never discussed the 1944 nomination with him and that before the convention he gave a letter to Senator Frank Maloney refusing the nomination.\textsuperscript{374} He was in the Oregon mountains at the time of the convention and claimed that he did not know about Roosevelt’s note to the convention.\textsuperscript{375} He admitted, however, that “[a]n FDR ‘draft’ would have been difficult to resist. I am glad I never had to face up to it in 1944.”\textsuperscript{376}

In 1948, Douglas worked privately with members of the liberal Americans for Democratic Action (ADA) to promote his candidacy at a time when nomination of the unpopular President Truman remained uncertain.\textsuperscript{377} Douglas privately criticized Truman for surrounding himself with too many conservatives and for appointing administrators who had emasculated federal regulatory agencies.\textsuperscript{378} Douglas shared Truman’s fear of Soviet expansionism, but, as one historian has explained, Douglas’s “public addresses carried the promise of a more inspiring, positive, and democratic foreign policy, oriented to the common people of the world and seeking, as he put it, to win the Cold War in the rice fields rather than on the battlefields.”\textsuperscript{379} A prominent journalist argued that Douglas, who had been a “tough, resourceful, and courageous” Securities and Exchange Commission (SEC) director when enthusiasm for the New Deal was flagging, “would restore direction to the liberal ranks.”\textsuperscript{380}

Convinced by his advisors that he needed Douglas on the ticket to win support among liberals, Truman offered Douglas the vice presidential nomination on the eve of the convention.\textsuperscript{381} Douglas at first may have hesitated because he had not yet given up hope that the convention would nominate him for president.\textsuperscript{382} After the ADA threw its support to Truman, Douglas finally told Truman, who had ended up pleading with Douglas to serve as his running mate, that he refused to “use the Court as a stepping stone.”\textsuperscript{383} In fact, however, Douglas probably rejected the nomination because he accepted the conventional
wisdom that Truman couldn’t win. According to some accounts, Douglas, whose finances were troubled, was particularly reluctant to give up the security of his judicial position. Although he could make more money in private practice, being a lawyer had little appeal for him. According to another account, however, Douglas said that he expected Truman to win and that he would have run if he thought that Truman would lose, thereby enabling him to avoid the indignity of four years in the vice presidency and positioning him for a good shot at the top of the ticket in 1952. Douglas himself explained in his memoirs that he believed that Truman could win but that he wanted to remain on the Court in order to continue to work with Murphy, Rutledge, and Black in support of civil liberties. “It seemed to me,” Douglas recalled, “that only very special reasons would justify breaking up that foursome.”

Only forty-nine years old when he rejected Truman’s offer, Douglas appears to have entertained hope that he could be elected president in another four or eight years. In 1952, several prominent Democrats encouraged Douglas to seek the presidency, and Douglas was the first choice of an astonishing 67 percent of 3,600 readers of the liberal Nation who responded to a poll in April 1952. Expressing the view of many liberals who wanted Douglas to seek the presidency in 1952, Yale Law Professor Fred Rodell praised Douglas for having “etched out a clear and militantly liberal political credo” in his “ringing defense of civil liberties at home” and “his castigation of a military-minded foreign policy that would try to stop communism with guns (and dollars) alone.” Denying that Douglas had subtly crafted his judicial opinions to win political support, Rodell argued that Douglas had expressed “precisely the views that no elective office-seeker in his right mind would utter in the face of the conform-or-else anti-Communist atmosphere of the times.”

Douglas concluded in 1952 with much regret that his plan to divorce his wife made his prospects too slim. Corcoran, who warned Douglas in 1952 that his divorce would ruin his political prospects, believed that Douglas would have been president if he had not divorced his wife of twenty-nine years in 1953. Similarly, another biography of Douglas concludes that any political ambitions

384. SIMON, supra note 358, at 274.
385. MURPHY, supra note 357, at 262.
386. Id.
387. Id.
388. DOUGLAS, supra note 373, at 290.
389. Id.
390. FRANK & COUNTRYMAN, supra note 354, at 1240.
391. MURPHY, supra note 357, at 314.
394. Id.
395. MURPHY, supra note 357, at 314.
396. SIMON, supra note 358, at 286.
he might have had were destroyed by this divorce and by his public support for U.S. diplomatic recognition of Communist China. 397

In 1960, bitten again by the presidential bug, Douglas told Lyndon Johnson that he was willing to resign from the Court to promote Johnson’s candidacy for the Democratic nomination, apparently believing that Johnson would offer the vice presidential slot to him. 398 When Johnson himself became the vice presidential nominee after the Democrats nominated John F. Kennedy for the presidency, Douglas lapsed into a severe mental depression because he knew that his last chance had passed. 399 He appears to have remained embittered until the end of his days, for, as one biographer has explained, “he had now been buried alive, trapped forever on the Supreme Court.” 400 At age seventy, he shocked a group of law clerks by emphatically declaring that he regretted his service on the Court, which he described as “peripheral” and as a “backwater” that was “too far out of the action.” 401

In addition to Douglas, at least one other justice appointed by Roosevelt had serious political ambitions. James R. Byrnes, who served on the Court from October 1941 to October 1942, was a member of the House of Representatives for fourteen years and spent ten years in the Senate, where he was serving when Roosevelt placed him on the Court. 402 Byrnes, who later was Secretary of State from 1945 to 1947 and governor of South Carolina from 1951 to 1955, was an intensely political man who found his judicial duties boring and confining. 403 He frequently conferred with Roosevelt on affairs of state even while he was a justice, and he resigned to lead the wartime Office of Economic Stabilization, a post in which he was often called the “assistant president.” 404 Byrnes, who harbored presidential ambitions for decades, tried to obtain the Democratic vice presidential nomination in 1944, but he was unacceptable to Roosevelt because he was an ardent racial segregationist. 405 Temperamentally unsuited to the Court, Byrnes had the good grace to resign after only one term to return to active

397. HOWARD BALL & PHILLIP J. COOPER, OF POWER AND RIGHT: HUGO BLACK, WILLIAM O. DOUGLAS, AND AMERICA’S CONSTITUTIONAL REVOLUTION 82 (1992). The assessments of the impact of Douglas’s divorce may be exaggerated. As early as 1920, the Democrats had nominated a divorced man, James M. Cox, and the 1952 Democratic presidential candidate, Adlai Stevenson, was divorced. Although both Cox and Stevenson were defeated in landslides, there is little reason to suppose that their marital status was a major factor in their defeat. Unlike Stevenson and Cox, however, Douglas may have divorced his first wife to marry another woman.
398. MURPHY, supra note 357, at 350.
399. Id.
400. Id.
401. Id. at 351.
403. Id. at 304-07.
404. Id. at 311-21.
405. Id. at 332-62.
Another justice who was considered for the presidency was Fred Vinson. Truman, who had offered the vice presidency to Douglas, later tried to anoint Chief Justice Vinson as his successor. During his final years as president, Truman repeatedly urged Chief Justice Vinson to seek the presidency and assured Vinson of his full support if he became a candidate. Truman, who believed that he had a duty to recommend a successor, was a personal friend of Vinson and had appointed him to the Court in 1946. Truman admired Vinson’s record as a member of Congress during the 1920s and 1930s and as a high ranking administrator in federal agencies during World War II, and he believed that a Vinson presidency would remain faithful to Truman’s own policies and politics. Like Truman, Vinson was moderately liberal on domestic issues and supported a strong American role in international affairs.

Truman broached the idea with Vinson as early as 1949 or 1950, but Vinson reportedly rejected Truman’s overtures. In meetings in October and November 1951, several months before Truman announced that he would not seek another term, Truman again exhorted Vinson to become a candidate for the Democratic nomination in 1952. When Vinson expressed reluctance to use the chief justice position as a stepping stone to the presidency, Truman urged Vinson’s wife, Roberta, to encourage her husband to run. Vinson waited several weeks before informing Truman that he definitely would not become a candidate. Although the sixty-one year old Vinson told Truman that his “physical condition” was why he was rejecting Truman’s proposal, Vinson also appears to have worried that his candidacy would harm the Court’s reputation. Vinson’s refusal to run surprised Truman, who had no second choice.

406. Id. at 318-320.
407. HAMBY, supra note 377, at 480.
408. Id. at 480-81.
409. Id. at 480.
410. Id. at 480-81.
411. Id. at 480.
413. Id.
414. Id.
415. Id. at 195-96.
416. HARRY S TRUMAN, MEMOIRS, VOL. 2 490 (1956).
417. ST. CLAIR & GUGIN, supra note 412, at 195-96. Although Vinson died of heart trouble the following year at age 63, he apparently was not aware that his health problems were serious when he rejected Truman’s of fer of support. Id. at 196, 336.
It is not surprising that Truman is the only President to offer the vice presidency to a justice or to importune a justice to run for President, \(419\) for few Presidents have displayed less respect for the Court’s institutional integrity. None of the four men he appointed to the Court had particularly outstanding qualifications for the high bench, and historians rank at least three of them as poor or mediocre justices.\(420\) Three were Truman’s personal friends, and all continued to act as political advisors to one extent or another after they joined the Court.\(421\) When one of them, Tom Clark, voted with the majority to invalidate Truman’s seizure of the steel mills during the Korean War, Truman appears to have felt personally betrayed, although he resumed his friendship with Clark after fulminating against the Justice for a while.\(422\)

Aside from Douglas, Chief Justice Earl Warren probably was the member of the Court who came the closest after Hughes to resigning from the Court to become a presidential candidate. Warren was a natural prospect for the presidency even after he became Chief Justice in 1953 because he had been such a prominent politician before he joined the Court.\(423\) Warren was elected three times as governor of California, the post from which he resigned to take his seat on the Court.\(424\) Widely commended for his honest, progressive, and efficient governorship during a time of explosive growth in California, Warren was a leading candidate for the Republican presidential nomination in 1948 and 1952.\(425\) Warren nearly was elected to the vice presidency in 1948 when he was the running mate of Thomas E. Dewey on a ticket that voters defeated in one of history’s closest elections.\(426\) In 1952, Warren stayed in the race until the Republican convention, when he threw his support to Dwight D. Eisenhower, who, according to some accounts, promised to appoint Warren to the Court in exchange for his support.\(427\)

Although Warren had substantial legal experience as a former district attorney and former attorney general of California, he was widely regarded at the time of his appointment as much more of a politician than a judge.\(428\) Even Warren himself was not sure how well he would adapt to the cloistered life of the Court.\(429\) Less than a year after Warren became Chief Justice, the Court’s decision in Brown v. Board of Education generated support for another

\(419\). See supra Part II.
\(420\). See DONOVAN, supra note 418, at 375.
\(421\). Id. at 387.
\(422\). See id. at 391.
\(424\). See id.
\(425\). See id.
\(426\). See id.
\(427\). See id.
\(429\). See id. at 315.
presidential candidacy by Warren because Americans outside the South widely hailed *Brown* and appreciated Warren’s important role in shaping it.\footnote{See id. at 313.} For example, on the day after the *Brown* decision, the *Sacramento Bee* editorialized that “the Republican Party will consider him among the very eligible persons when the time comes to nominate a successor to President Eisenhower.”\footnote{Id.} According to one account, a frustrated Warren exclaimed at a dinner party, “Do people have so low a regard for the Supreme Court that they think all they need do is tap a justice on the shoulder and he will run for the presidency? If so, it’s time they learned better.”\footnote{Id.}

Speculation continued to grow because of uncertainty about whether Eisenhower would run again. The prospect that a chief justice would resign to seek the presidency generated considerable dismay.\footnote{See Leslie Dubeck, *Understanding “Judicial Lockjaw”: The Debate Over Extrajudicial Activity*, 82 N.Y.U. L. REV. 569, 600 (2007).} For example, *The Washington Post* warned that Warren’s candidacy “would inflict incredible harm upon the Supreme Court as an institution.”\footnote{Id. (quoting *Voters Eye the Chief Justice*, WASH. POST & TIMES HERALD, Apr. 14, 1955, at 16).} Warren publicly disavowed presidential aspirations on April 15, 1955 in the wake of a Gallup poll that indicated that Warren was the favorite choice of Republicans if Eisenhower did not seek re-election.\footnote{George Gallup, The Gallup Poll: Public Opinion 1935-1971, VOL. 2 1325 (William P. Hansen and Fred L. Israel, eds.) (1972). In this poll, Warren was the choice of 25 percent of Republicans and 31 percent of independents. *Id.* Vice President Richard Nixon ranked second, as the choice of 19 percent of Republicans and 11 percent of independents. *Id.*} Warren declared that speculation about his candidacy had become

a matter of embarrassment to me because it reflects upon the performance of my duties as Chief Justice of the United States. When I accepted that position, it was with the fixed purpose of leaving politics permanently for service on the Court. That is still my purpose. It is irrevocable. I will not change it under any circumstances or conditions.\footnote{Jim Newton, *Justice for All: Earl Warren and the Nation He Made* 334 (2006).}

Warren’s statement received widespread praise.\footnote{Id.} As one historian has explained, “Warren had raised the stature of the Court by removing it and himself from electoral politics and by making it clear that the presidency was not an office worth leaving the Court to pursue.”\footnote{Id.}

Speculation about a Warren candidacy became more intense after Eisenhower suffered a heart attack in September 1955.\footnote{See Cray supra note 428, at 314.} During the following
month, 62 percent of persons polled by George Gallup said that they did not expect Eisenhower to run for re-election in 1956. 440

Despite renewed speculation that Warren would become a candidate, Warren did not reiterate his refusal to seek the presidency. 441  Apparently interested in the possibility, Warren informed a friend, the prominent syndicated columnist Drew Pearson, that he might run. 442  Meanwhile, Warren reportedly told a longtime aide in December 1955 that he would run for President if he thought that his old nemesis Richard Nixon otherwise would become President. 443

Public opinion surveys conducted by George Gallup during late 1955 and 1956 continued to indicate that Warren would command impressive support for the Republican nomination if Eisenhower did not seek re-election. 444  In four surveys between October 1955 and January 1956, support for Warren among Republicans ranged from 23 percent to 30 percent. 445  Warren was the most popular choice in polls conducted in December 1955 and January 1956, when he out-placed even Vice President Nixon. 446  In one poll during the autumn of 1955, 38 percent of respondents expressed an expectation that Warren would be the Republican nominee in 1956, and 50 percent — including 58 percent of Republicans — said that they would like to see Warren run for president in 1956. 447

When a reporter asked Eisenhower at a press conference on January 25, 1956 whether it would be “bad policy” for Warren to seek the presidency if he chose not to seek re-election, Eisenhower replied that “[w]e shouldn’t get too great a confusion between politics and the Supreme Court” and that maintenance of the distinction would enable each branch to remain “respected in the eyes of the people.” 448  Eisenhower also hinted that Warren should resign from the Court if he became a candidate. 449

A few days later, Warren surprised Eisenhower’s press secretary, James Hagerty, by expressing annoyance over Eisenhower’s remarks when Warren saw Hagerty at a party. 450  Eisenhower observed in his personal diary that

440. Gallup, supra note 435, at 1367.
441. Newton, supra note 436, at 334.
442. Newton, supra note 436, at 334-35.
445. Id.
446. Id. at 1385, 1395.
447. Id. at 1374-75.
449. Id.
the first time he indicated himself as receptive to the nomination, he
should resign from the Supreme Court. I think it would be possibly
feasible and ethical for him to say nothing until the time of the
Convention. At that time, if literally drafted, he could submit his
resignation and accept the nomination.\footnote{451}

Eisenhower mused, however, that he could not step aside in favor of Warren
unless he had assurances from Warren in advance of the convention that he
would respond to a draft.\footnote{452}

When asked at a press conference on February 8, 1956 whether his
comments at the previous press conference indicated that he would oppose
Warren’s nomination if he did not run again, Eisenhower replied, “[o]pposed?
For goodness sake, I appointed him as Chief Justice of the United States; and
there is no office in all the world that I respect more.”\footnote{453} Privately, Eisenhower
told his press secretary that he did not believe that Warren would seek the
presidency because he was happy to serve as Chief Justice, did not want to give
up a position that had life tenure, and wanted to become one of the great chief
justices.\footnote{454}

During the early weeks of 1956, when Eisenhower considered dumping
Nixon from the ticket if he (Eisenhower) sought re-election, there also was
widespread speculation that Eisenhower might offer Warren the vice presidential
nomination.\footnote{455} Although Eisenhower might have preferred Warren to any other
running mate, Eisenhower was convinced that Warren would not resign the chief
justiceship to seek the vice presidency.\footnote{456} Speculation about a Warren candidacy
finally ended on February 29, 1956, when Eisenhower announced that he would
seek a second term.\footnote{457}

Warren does not appear to have been widely considered as a potential
nominee in 1960, partly because Nixon and Nelson Rockefeller had emerged as
strong contenders, and also because Warren would have been nearly 70 years old
by inauguration day.\footnote{458} Even so, the United Automobile Workers President,
Walter Reuther, stated in a television interview in October 1958 that organized
labor would support Warren for the presidency, perhaps because Reuther feared
Nixon’s candidacy.\footnote{459}

\begin{thebibliography}{99}
  \bibitem{451} Id.
  \bibitem{452} Id.
  \bibitem{453} The President’s News Conference of February 8, 1956, THE AMERICAN PRESIDENCY
  PROJECT. (John Wooley & Gerhard Peters, eds.), http://presidency.ucsb.edu/ws/index.php?
  pid=10726 (last visited Feb. 7, 2011).
  \bibitem{454} CRAY, supra note 428, at 314.
  \bibitem{455} STEPHEN E. AMBROSE, EISENHOWER, VOL. 2, THE PRESIDENT 293 (1984).
  \bibitem{456} Id.
  \bibitem{457} Id. at 296.
  \bibitem{458} See POLLACK, supra note 443, at 10.
  \bibitem{459} Id.
\end{thebibliography}
During the past half century, Sandra Day O’Connor is the only justice who has been mentioned even half seriously as a possible presidential or vice presidential candidate. Speculation that Republicans might nominate O’Connor for the presidency or vice presidency occurred mostly in 1988, when Ronald Reagan was retiring from the presidency. In a written statement released on May 27, 1988, O’Connor quashed most speculation by declaring that ‘I most assuredly am not considering any other position in or out of government and do not expect to do so in the future.’ Although O’Connor was the only justice at the time who had been closely connected with partisan politics and who had elected office – she had been the majority leader of the Arizona senate and an elected trial court judge – she appears to have been considered for the presidency more because she was a prominent personality rather than because she once had been a prominent politician. As the first female member of the Court, she naturally attracted much publicity, and she probably was more familiar to voters than any other Justice at the time. O’Connor does not appear to have been mentioned for the presidency or vice presidency in 1992, when George H. Bush and Dan Quayle were incumbents. By 1996, when she was sixty-six years old, her age may have excluded her from serious consideration.

III. HAZARDS OF PRESIDENTIAL AMBITIONS AMONG SUPREME COURT JUSTICES

There is an obvious danger that presidential aspirations among Supreme Court Justices will encourage justices to change their votes or at least craft their opinions in a manner to curry favor with voters or political king makers. As Gary Stein observed in assessing the extent to which presidential ambitions may have influenced Black’s performance on the Court, “the influence of ambition on reason is usually . . . subtle. It can be the puff of wind that imperceptibly pushes analysis in one direction rather than another, the unseen pebble that tilts the scales this way or that.”

The perception that any justice was trimming her opinions to fit her political ambition easily could erode public faith in judicial independence. This would tend to diminish public confidence in the integrity of judicial opinions, which is the ultimate source of the Court’s power and influence. As the syndicated

461. Id.
464. Stein, supra note 350, at 433. Stein aptly remarked that “[o]f all the many factors that tug at a judge’s conscience in deciding cases, personal political advantage should not be among them.” Id.
465. In assessing the political ambitions of Chase and Field, for example, Professor Fairman observed that, “Chase’s pursuit of the Presidency had been notorious and enduring, and had with good reason caused his action in significant cases to be suspect. Field’s aberration was short-lived,
columnist Arthur Krock warned in 1940 when Douglas was mentioned as a possible vice presidential candidate, any opinion written by Douglas “will be weighed on the scale of political ambition. This will damage him and hamper the Supreme Court.” 466 Similarly, Professor Peter Bell has pointed out that “Justices who run for office subject themselves and the Court to the traditional attacks and slurs of campaign politics.” 467 As Walter Lippmann observed in 1935 in response to speculation about the possibility that Justice Roberts would become a presidential candidate, justices “should be regarded as having forever renounced all other worldly ambitions” because the Court’s prestige “depends upon the people’s own belief that the Court is above all ordinary worldly temptation.” 468

Although even occasional forays by justices into presidential politics may harm the Court, persistent politicking among justices could have catastrophic consequences. Yale Law Professor Alexander Bickel warned that “the recurrence of justices with manifest political aspirations would in time destroy an institution whose strength derives from consent based on confidence.” 469 Bickel pointed out that, paradoxically, the Court needs to stand above political involvement precisely because the constitutional issues that it addresses so often have political dimensions. 470

Life tenure, which was intended to help federal judges preserve their independence, 471 also should discourage Supreme Court Justices from seeking the presidency insofar as justices do not need to stray from the bench in order to enjoy the permanent security of a powerful position. As a Tennessee delegate to the 1880 Democratic national convention explained his opposition to Justice Stephen J. Field’s presidential candidacy,

The fathers decided that it was wise to place one branch of the government above temptation and therefore made the tenure of office of the judges of the courts a life one. I am not in favor of changing that and giving the judges reason to expect political reward for their decisions. 472

466. SIMON, supra note 358, at 260.
467. Peter Alan Bell, Extrajudicial Activity of Supreme Court Justices, 22 STAN. L. REV. 587, 595, 595 n.52 (1970). Bell has pointed out that “[a] Justice in the political arena will probably provoke attacks on the Court as well as on himself by opponents who seek to counteract the institution’s prestige, which apparently has been thrown into battle against them.” Id.
468. LIPPMANN, supra note 279, at 278-79.
470. Id. at 134.
Justices are understandably reluctant to lay down the security of a powerful lifetime appointment for the uncertainties of electoral politics, and this presumably is one of the reasons why even more have not sought the presidency.

In addition to eroding public confidence in herself and the Court itself, any justice who would allow political ambition to influence his or her work on the Court also might violate Canon 1 of the American Bar Association’s *Model Code of Judicial Conduct*, which provides that a judge shall uphold the independence, integrity, and impartiality of the judiciary. As the Preamble to the Code points out, “An independent, fair, and impartial judiciary is indispensable to our system of justice. A judge should participate in establishing, maintaining and enforcing high standards of conduct, and shall personally observe those standards so that the integrity and independence of the judiciary will be preserved.” At the very least, any activities by a justice in anticipation of a presidential candidacy, or any encouragement that a justice might provide to speculation that he or she would available as a candidate, would surely violate the rest of Canon 1, which declares that a judge “shall avoid impropriety and the appearance of impropriety.” Preparations for a presidential campaign also might violate Canon 2, which provides that “[a] judge shall perform the duties of judicial office impartially, competently and diligently.”

Although political ambition may influence a justice’s votes in actual decisions, an even more insidious danger is that Justices will allow their presidential aspirations to influence their votes to grant or deny writs of certiorari. Because cert votes usually are confidential unless a Justice chooses to dissent, a member of the Court could conceal his or her political motivations much more easily in voting on a cert petition than on a vote in a case on which the Court renders a decision. A politically ambitious justice might oppose a writ of certiorari to avoid having to vote on a politically controversial issue. On the other hand, such a justice might vote to grant certiorari in order to enable him to cast a vote and write an opinion for the Court or in concurrence or dissent that would curry favor with individuals, interest groups, or voting blocs that would facilitate the justice’s political ambitions. Because the Court’s decisions about what to decide are tremendously important in shaping the law, a justice who permits political ambition to affect the performance of his or her gate-keeping function has committed an ethical violation that potentially is as serious as allowing presidential aspirations to influence her vote or opinion in a decided case. Insofar as it takes the votes of only four justices to grant a writ of

474. *Id.*
475. *Id.*
476. *Id.* at Canon 2.
certiorari,\textsuperscript{478} improper political motivations can have an even greater impact upon the Court if a justice votes to grant certiorari for political reasons than if a justice permits politics to affect his or her vote on the merits.

Another hazard is that efforts by justices to promote their political prospects also will almost inevitably distract them from their arduous duties on the Court. Even though the Court during the past twenty years has reduced the number of cases it hears from an average of 150 per year\textsuperscript{479} to about 85,\textsuperscript{480} justices must spend as many or more hours than ever on the time-consuming task of reviewing certiorari petitions and on various collateral responsibilities. Because there are so many activities that a justice has no choice but to perform, such as attendance at conferences and oral arguments, politicking is likely to reduce the time that a justice has for the duty that is the most important but also the most easily dispensable – careful deliberation and reflection. As Felix Frankfurter once observed, the Court’s judgments

\begin{quote}

presuppose ample time and freshness of mind for private study and reflection in preparation for discussion at Conference. Without adequate study there cannot be adequate reflection; without adequate reflection there cannot be adequate discussion; without adequate discussion there cannot be that fruitful interchange of minds which is indispensable to thoughtful, unhurried decision and its formulation in learned and impressive opinions.\textsuperscript{481}
\end{quote}

Preparation for a presidential candidacy might therefore violate Rule 2.1 of Canon 2 of the \textit{Model Code of Judicial Conduct}, which provides that the judicial duties of a judge take precedence over all the judge’s personal and extrajudicial activities.\textsuperscript{482} As we have seen, presidential aspirations have distracted some Justices from their judicial duties.\textsuperscript{483} The danger of such distraction may be even greater today, when the nomination process has become even more complex. In particular, even a justice whose only hope was to become a compromise candidate might perceive that he or she needed to line up potential donors in order to convince the party that he or she would be a viable candidate in the general election. Even a surreptitious, discrete, or half-hearted campaign could therefore seriously distract a justice from his or her duties.

Presidential fever also might encourage an aspiring candidate to make public pronouncements about issues that the Court later might later need to adjudicate.

\begin{itemize}
\item \textsuperscript{478} Williams v. Chrans, 50 F.3d 1358, 1360 (7th Cir. 1995).
\item \textsuperscript{480} See Supreme Court of the United States, \url{http://www.supremecourt.gov/opinions/opinions.aspx}.
\item \textsuperscript{482} \textit{MODEL CODE OF JUD. CONDUCT}, Canon 2 R. 2.1 (2010).
\item \textsuperscript{483} \textit{See supra} Part II.
\end{itemize}
McLean, for example, was rightly criticized for publishing a letter in 1848 in which he discussed his views about congressional power to enact legislation concerning slavery in the territories because his contemporaries recognized that this issue could very well come before the Court\(^{484}\) – as it did nine years later in *Dred Scott*. Such extrajudicial comments by judges naturally create suspicion of judicial bias, either because the judge appears to have pre-judged an issue or because a judge is loath to contradict his earlier expression of his or her views. Although a justice who created such suspicion could recuse herself from a case involving an issue about which she had spoken,\(^{485}\) recusal would deprive the Court of her services, thereby foisting more work onto the other justices. Moreover, the political motivation to make extrajudicial comment about political issues could cause a justice to violate Rule 2.10(B) of Canon 2, which prohibits a judge from making “pledges, promises, or commitments” regarding “cases, controversies, or issues that are likely to come before the court.” Such comments could raise questions about whether the justice has an “open mind.”\(^{486}\)

Even if a justice who had presidential ambitions refrained from making statements about issues or cases that might come before the Court, she might be tempted to defend her votes or opinions in cases that the Court already had decided. Moreover, a former justice who campaigned for the presidency also might feel the need to publicly defend his or her judicial votes and opinions or even to boast of them. In addition to generating questions about whether the justice had allowed politics to influence the justice’s work on the Court, such public comments also might detract from the integrity of the justice’s written opinions by creating confusion about what the justice – and perhaps the Court – actually intended to say.\(^{487}\) As the columnist Walter Lippmann observed in 1935 when some Republicans were advocating the presidential candidacy of Justice Roberts, “what could be less edifying than a political campaign in which a justice of the Supreme Court was defending his judicial opinions from the end of a railroad car?”\(^{488}\)

The presidential ambition of a justice also is likely to create discord and disharmony on the Court.\(^{489}\) Acutely aware that they lack the powers of the

---

486. The Terminology of the *ABA Model Rules of Professional Conduct* and *Code of Judicial Conduct* defines “impartiality” within the meaning of Canons 2(10) and 3(1) as denoting “absence of bias or prejudice in favor of, or against, particular parties or classes of parties, as well as maintenance of an open mind in considering issues that may come before a judge.”
487. Justice Brennan aptly observed that “the reasons behind the social policy fostering an independent judiciary also require that the opinions by which judges support decisions must stand on their own merits without embellishment or comment from the judges who write or join them.” Alan F. Westin, *Out-of-Court Commentary by United States Supreme Court Justices, 1790-1962: Of Free Speech and Judicial Lockjaw*, 62 *COLUM. L. REV.* 633, 633 (1962) (quoting Address by William J. Brennan, Jr., Student Legal Forum, Charlottesville, Virginia, Feb. 17, 1959).
488. *Lippmann*, *supra* note 279, at 278.
489. For example, Black, despite his own presidential ambitions, was dismayed by the persistent
purse and the sword, justices tend to be fiercely protective of the Court’s institutional integrity and independence. They are ever mindful that the perception that the Court is above politics is essential in maintaining the public respect that the Court must have in order to retain its authority. Any justice who threatens to erode this respect by abusing her office as a platform for higher office is likely to encounter stinging disapproval from her colleagues. Any provocation of ill-will, friction, resentment, or suspicion among the Nine cannot help but to interfere with the efficient disposition of the Court’s work. At the very least, the perception that a justice harbored political ambition might interfere with the manner in which the Court conducts its business. For example, a chief justice or senior dissenting justice might be reluctant to assign a politically significant opinion to such a justice. Justices also are likely to resent a colleague who might not be carrying his or her full work load because he or she is distracted by presidential ambition.

Because presidential aspirations among justices may so severely compromise the Court’s integrity, Presidents should be circumspect about nominating persons who seem likely to use the Court as a stepping stone to the presidency. As the foregoing history indicates, most of the many Justices who harbored presidential ambitions had served in high ranking legislative and judicial offices before joining the Supreme Court, and a handful also had been prominent state jurists. A President, therefore, should be particularly careful to take the danger of presidential ambition into account when he nominates a politician to the Court. Although no prominent political leader has been appointed to the Court

rumors that Douglas would be willing to become a presidential or vice presidential Candidate. Newman, supra note 337, at 329. Similarly. Jackson appears to have been much annoyed by Douglas’s vice presidential aspirations. Culver and Hyde, supra note 364, at 317. Waite privately made sarcastic remarks about Davis’s presidential ambitions and also appears to have been annoyed by Field’s efforts to wangle the Democratic presidential nomination. Magrath, supra note 227, at 286-87.

490. See, e.g., supra Parts II.C., II.D. (discussing Justice Waite’s and Justices Hughes’s desires to maintain the integrity of the Court).

491. See supra Parts II.C., II.D.

492. As Paul Finkelman has observed, “McLean’s quadrennial quest for the White House surely affected his position on the Court. When Chief Justice Taney handed out assignments for majority opinions, he must have understood that allowing McLean to write an important opinion would only heighten McLean’s political profile. Since McLean and Taney were not political allies, the Chief Justice had little reason to support McLean’s aspirations.” Finkelman, supra note 50, at 525.

493. Bell, supra note 467, at 597.

494. In deploring the political ambitions among Justices, Frankfurter once alleged that “the most numerous and in many ways the worst offenders” were “men who came to the Court from state courts, in some instances with long service on such courts.” Frankfurter, The Supreme Court in the Mirror of Justices, supra note 18, at 787. Frankfurter, however, provided no examples, and there is little empirical support for his thesis. Aside from McLean, who served three years on the Ohio Supreme Court, Field, who was a justice of the California Supreme Court for four years, Davis, who served for 14 years as an Illinois circuit court judge, and Black, who served briefly as a municipal court judge, none of the justices who have harbored serious presidential aspirations have been state court judges.
for nearly sixty years, several such persons are reported to have received consideration for the presidency. In particular, President Clinton reportedly considered the nomination of former Arizona Governor Bruce Babbitt, former New York Governor Mario Cuomo, and Senator George Mitchell of Maine.

The danger of presidential ambition among Supreme Court Justices remains present even though it is virtually unthinkable that any sitting justice today would openly campaign for the presidency. Such a campaign would severely compromise contemporary expectations of judicial behavior, and it also would violate Rule 4.5(A) of Canon 4 of the ABA Model Code of Judicial Conduct, which requires a judge to resign upon becoming a candidate for non-judicial office.

Similarly, it would be almost impossible for any justice today to conduct a covert campaign for the presidency. Such campaigns were possible until the middle of the twentieth century because most delegates until then were selected by party organizations and usually were not pledged to vote for any candidate. Justices such as McLean, Chase, Miller, and Field, therefore could quietly line up support among delegates without attracting public attention. Because, however, most delegates today are selected in party primaries in which voters either vote directly for a candidate or select delegates who are committed to a particular candidate, a justice could not hope to become a serious contender without openly declaring his or her candidacy. At the present time, as in the past, age is another reason why more justices have not sought the presidency. Even at the time of their appointment, many justices have pushed the outer limits of the acceptable age range for presidential timber, and most justices at any given time have been far past the average age of presidential contenders.

If a justice who wanted the presidency were unwilling to resign from the Court, his or her only option therefore would be to surreptitiously build support as a compromise candidate to which a convention might turn if the candidates

495. See Vol. 4-5 THE JUSTICES OF THE UNITED STATES SUPREME COURT, THEIR LIVES AND MAJOR OPINIONS (Leon Friedman and Fred L. Israel eds., 1995) (examining the lives and careers of the Justices appointed to the Supreme Court from 1939 – 1994).


501. See generally id.

502. See supra Part II.
who ran in the primaries could not obtain the majority vote necessary for nomination. A covert campaign of this nature, however, would be unrealistic today. Long gone is the era in which conventions would nominate “dark horse” candidates or draft a nominee who had not sought the nomination. The juggernauts of front-running candidates and the imperative of party unity have ensured that party conventions have become coronations rather than convocations for selecting candidates. Not since 1952 has any major party’s convention needed more than one ballot to nominate a presidential candidate. Moreover, because there is today an unwritten understanding that voters, rather than politicians, should select presidential candidates, it is extremely unlikely that any convention would nominate anyone who had not obtained a mandate from members of his or her political party by winning substantial numbers of votes in presidential primaries, even if a convention were deadlocked.

Despite these formidable odds, however, a justice might still convince herself that a deadlocked convention might seek her services as a candidate. Even a remote prospect might encourage a politically ambitious justice to write opinions in a manner that would not offend powerful constituents of her own party. Unscrupulous interest groups that wanted to influence the justice might help fan the flames of her ambition. As one scholar has observed, “[p]olitical ambitions are of course tempting apples of discord in a judicial Eden. When lobbies form, uninvited but perhaps not sufficiently discouraged, to promote the fortunes of individual justices, additional strain is put upon judicial otherworldliness.” Moreover, there is a very real danger that a highly political justice might try to use the Court as a platform for a future presidential bid.

Even many justices who have no political constituency may tend to harbor secret presidential ambitions because their elevation to the Court has catapulted them into political prominence. If any of the Justices appointed during the past half century had sought the presidency before they became members of the Court, their candidacies would have been jokes. Although none of these Justices, with the possible exception of O’Connor, would likely have had the ability to raise sufficient capital and attract a sufficient number of votes to have a significant chance of winning a presidential nomination, any of these Justices probably would have received reasonably serious consideration by the news media and at least some political operatives and potential donors if they had sent out feelers from the bench or had resigned from the Court to actively campaign. Similarly, their appointment to the bench would have transformed any vice presidential ambitions from the realm of fantasy to at least the level of

503. See supra note 217 and accompanying text.
504. See generally ABRAHAM, supra note 500 (examining the history of Supreme Court appointments).
505. See id.
improbability. The danger that Supreme Court Justices will seek the presidency is therefore not entirely eliminated by the appointment of justices who have not served in elective office.

Another danger of appointing a politician may arise from the politician’s connections to the nominating president. All four of Truman’s nominees were personal friends, and three had been his colleagues in the United States Senate.507 At least three of these Justices receive low rankings from Court historians.508 Although some presidential friends have turned out to be distinguished Justices,509 a President’s appointment of a close political ally can encourage the erosion of separation of powers insofar as some presidents have remained in close contact with “their” Justices.510 The most egregious example was Abe Fortas, who offered political advice to Lyndon Johnson on a virtually daily basis and even is reported to have helped him write speeches and select Vietnam bombing targets.511 Since Johnson, no political buddy of a president has served on the Court.512 Harriet Miers’s close association with George W. Bush as White House counsel helped sink her nomination in 2006.513

If a prominent political figure were appointed to the Court, there also is a danger that the justice would fail to make the adjustment from the robust world of politics to the more hermetic life of the Court. Such a justice might fail to recognize that public respect for the Court and the integrity of its rulings are jeopardized when a justice makes extra-judicial comments about non-judicial issues. Although justices can facilitate public understanding of the law through public comments on judicial issues, a justice who is too accustomed to the limelight of the media might be tempted to comment indiscreetly about the inner workings of the Court or the Court’s opinions, which usually should speak for themselves. Too many American institutions, public and private, have been trivialized by our culture of celebrity. The Court has largely avoided this virus, and all justices would do well to emulate Souter, who shunned unnecessary publicity and faithfully and quietly performed his duties without fanfare or

507. See ABRAHAM, supra note 500, at 157-159 (discussing Truman’s appointments to the Supreme Court).
508. See id.
509. See supra Part II.
510. See FRED GRAHAM, Abe Fortas, in 4 THE JUSTICES OF THE UNITED STATES SUPREME COURT, THEIR LIVES AND MAJOR OPINIONS 1451, 1451-64 (Leon Friedman and Fred L. Israel eds., 1995) (examining the life and career of Abe Fortas).
511. See id.
512. See generally ABRAHAM, supra note 500, at 251-326 (examining Supreme Court appointments from Johnson to Clinton).
This is not too much to ask of a professional politician, but it may be too much to expect.

Of course, the Court might benefit from the appointment of a prominent politician who had extensive experience as in the executive or legislative branches of state or federal government. Executive experience might be especially useful for a chief justice. Warren’s long tenure as governor of California, for example, is credited with helping him achieve a consensus in *Brown v. Board of Education* and helping to maintain harmony on the Court during his years as chief justice. Similarly, Taft’s experience as president enabled him to persuade Congress to enact the Judiciary Act of 1923, which gave the Court much greater discretion over its docket. There is, however, no clear correlation between experience as an elected official and distinction as a Supreme Court justice. Several of the Court’s most distinguished members, including Chief Justice Warren, Marshall, Chase, and Taft, had substantial experience in elective office. On the other hand, many other notable justices, including Holmes, Brandeis, and Cardozo, never were elected to public office.

As various commentators have pointed out during recent years, a president who wanted to diversify the career backgrounds of the Justices could start by appointing a justice who had trial court experience. This would be particularly useful in the many cases involving criminal procedure that come before the Court. O’Connor’s experience as a state trial judge clearly made her a

---

514. See Edward de Grazia, David Hackett Souter, in 5 The Justices of the United States Supreme Court, Their Lives and Major Opinions 1806, 1806-28 (Leon Friedman and Fred L. Israel eds., 1995) (examining the life and career of David Hackett Souter).


516. It is ironic that the only Justice who also served as a President –William Howard Taft – desired a Supreme Court position much more than he ever wanted the presidency. Alpheus Thomas Mason, William Howard Taft: Chief Justice 22-23 (1964). Taft who was president from 1909 to 1913 before serving as Chief Justice from 1921 to 1930, had turned down two offers of an associate justiceship in 1902 because his wife Helen strongly wanted him to remain available for the presidency and in 1906 because he wanted to complete his work as governor of the Philippines. Id. at 27, 30.

517. See supra Part II.

518. See Paul A. Freund, Oliver Wendell Holmes, in, 3 The Justices of the United States Supreme Court, Their Lives and Major Opinions 873, 873-82 (Leon Friedman and Fred L. Israel eds., 1995) (examining the life and career of Oliver Wendell Holmes); Alpheus Thomas Mason, Louis D. Brandeis, in, 3, The Justices of the United States Supreme Court, Their Lives and Major Opinions 1018, 1018-36 (Leon Friedman and Fred L. Israel eds., 1995) (examining the life and career of Louis D. Brandeis); Andrew L. Kaufman, Benjamin Cardozo, in, 3 The Justices of the United States Supreme Court, Their Lives and Major Opinions 1130, 1130-1151 (Leon Friedman and Fred L. Israel eds., 1995) (examining the life and career of Benjamin Cardozo).

more effective Justice.\textsuperscript{520} Hugo Black was derided at the time of his nomination in 1937 because his only judicial experience was as a judge of Birmingham’s lowly night police court, but Black’s tenure on this grassroots court is credited with helping to forge his sensitivity toward the needs of poor persons and African-Americans.\textsuperscript{521} During the past century, only a handful of Justices -- Clarke, William J. Brennan, Sanford, Vinson, and Whittaker -- have served as federal trial judges.\textsuperscript{522} Sonia Sotomayor’s experience as a federal trial judge before she became a Court of Appeals judge presumably was one of the reasons why Obama appointed her to the Supreme Court in 2009.\textsuperscript{523} The danger of political ambition among Supreme Court Justices also provides a sound reason for rejecting proposals for constitutional amendments to limit the tenure of justices. Except for mandatory retirement at an age when most men or women would be too old for consideration for the presidency or vice presidency, any limitation of tenure might encourage politically ambitious justices to trim their judicial opinions in order to promote their political prospects after the expiration of their judicial terms. Limitation of tenure would seem even more likely to encourage public speculation that at least some justices were rigging their opinions, if not their votes, to court political popularity. Since public faith in judicial independence is nearly as important as judicial independence itself, public suspicion that justices were allowing political ambition to influence their votes in anticipation of the day when their terms would expire would have a highly corrosive effect on the integrity of the judiciary, even if the justices were in fact beyond reproach in their refusal to allow political ambition to influence their opinions.

Some critics of justices who aspire to the presidency have proposed constitutional amendments to restrict the ability of justices to seek the presidency. Chief Justice Waite favored such an amendment.\textsuperscript{524} In the wake of the recent presidential bids of Field and Miller, \textit{The Washington Post} in 1884 called for an amendment making any Supreme Court Justice forever ineligible for any elective office.\textsuperscript{525} Hughes remarked privately late in his life that he wished that a constitutional amendment had prevented him from resigning from the Court to accept the Republican presidential nomination in 1916.\textsuperscript{526}

\begin{thebibliography}{1}
\bibitem{520} \textit{Id.}
\bibitem{521} \textit{Id.}
\bibitem{522} \textit{Id.}
\bibitem{524} \textit{TRIMBLE, supra} note 225, at 144 (quoting letter from Waite to H.L. Bond, Nov. 7, 1875).
\bibitem{525} Dubeck, \textit{supra} note 433, at 598 (quoting \textit{Let Us Amend the Constitution} (editorial), \textit{WASH. POST}, Feb. 17, 1884, at 4).
\bibitem{526} Hughes reportedly revealed this in April 1947 to the Special Committee on Federal Courts of the Association of The Association of the Bar of the City of New York, \textit{Senate Hearings, supra} at 38 (testimony of Edwin A. Falk, chairman of the Special Committee on the Federal Courts, The Association of the Bar of the City of New York).
\end{thebibliography}
In 1954, both houses of Congress considered a proposed amendment that would have barred justices from the presidency within five years of leaving the Court.\textsuperscript{527} Senator John Marshall Butler of Maryland, who sponsored the Senate resolution, explained that the legislation “not only would remove possible political temptation from the Justices, but would also remove any suspicion that the judicial deliberation of the Justices was colored by political aspirations.”\textsuperscript{528} Similarly, a prominent New York attorney because “[o]ver the years many presidential bees have been lodged in Supreme Court bonnets. Future bees are most apt to light in the bonnets of those jurists least dedicated to their judicial duties and most open to suspicion of political aspirations.”\textsuperscript{529} Likewise, the president of the American Law Institute, Harrison Tweed, expressed approval of the amendment because it would “protect the high repute of the Court” by helping to keep “politics out of the Court.”\textsuperscript{530}

Restrictions on political availability also might discourage some highly qualified persons from accepting a nomination to the Court. Chase and Warren, for example, might not have been willing to have become chief justice if their appointment had necessarily terminated their presidential opportunities. This is why at least one member of the House of Representatives opposed the proposed amendment to bar former justices from seeking the presidency within five years after leaving the Court.\textsuperscript{531}

Such an amendment also seems unfair to the Justices because it deprives them of a right that other Americans enjoy. An amendment also could have the unintended consequence of actually diminishing public confidence in judicial independence because it would suggest that justices were politically ambitious. Even though former Justice Roberts favored an amendment to prohibit justices from seeking the presidency within five years of leaving the Court, he acknowledged that some opposition to the proposed amendment arose from the belief that it placed justices “in a different class than any other official and that it shows kind of a distrust.”\textsuperscript{532}

527. Id. at 6.
528. Id. Butler believed that the measure was necessary “in order to protect the Justices from any suspicion that they, consciously or unconsciously, might be seeking public favor.” Id. Butler declared that “once a man has become of the judiciary, he should devote his entire thought toward serving our Government without the slightest intimation that he might some day divest himself of his judicial robes to assume a different role in his country’s Government.” Id.
529. Id. at 36 (testimony of Edwin A. Falk, chair of the Special Committee on Federal Courts of the Association of the Bar of the City of New York).
530. Id. at 44 (statement of Harrison Tweed on S.J. Res. 44).
531. Id. at 12 (testimony of Rep. George Anthony Dondero of Michigan). Although Dondero expressed his opinion that Chase “would have rejected the appointment offered to him . . . because Chase had presidential ambitions, and particularly in 1864,” Chase’s presidential ambitions in 1864 could not have influenced his decision to accept the chief justiceship since Lincoln did not nominate Chase until December 6, one month after the election. Chase’s future presidential ambitions, however, might well have caused him to reject the nomination. Whether this would have been good for the Court and country must remain a subject of speculation.
532. House Hearings, supra note 527, at 21. Roberts made clear, however, that he did not
Moreover, such a provision could deprive the nation of much needed leadership in a time of crisis, when a particular Supreme Court Justice might be uniquely able to provide the type of experience or talent that the nation desperately needed. Or a justice might be so respected that only he or she could provide sorely needed party unity or national unity. As we have seen, the perception in 1916 that Hughes was the only person who could unite the Republican Party went far toward ameliorating criticism of his resignation from the Court to accept his party’s presidential nomination.\(^{533}\)

IV. CONCLUSION

Presidential fever rarely was dormant among Supreme Court Justices before the middle of the twentieth century and at times it was rampant.\(^{534}\) Although no justice has sought the presidency during the past half century, the danger of presidential aspirations among justices never disappears.\(^{535}\) Such ambition is hazardous to the Court’s integrity because a politically ambitious judge could alter his votes or opinions to promote his availability for the presidency. Even the mere appearance of such temptation could impair the prestige that the Court must have to command public support for its decisions. Although the Court might benefit from the experience of a career politician, presidents should nominate such persons only if convinced that the Justice lacked any further political ambition.

\(^{533}\) See supra at notes 256-264 and accompanying text.
\(^{534}\) See supra Part II.
\(^{535}\) See id.
LIMITATIONS IMPOSED ON THE DUAL SOVEREIGNTY DOCTRINE BY FEDERAL AND STATE GOVERNMENTS

Dr. Thomas White, J.D., Ph.D.*

ABSTRACT

Because the dual sovereignty doctrine permits multiple prosecutions of an individual by state and federal governments for essentially the same conduct, the increasing federalization of criminal law has marginalized much of the double jeopardy protection afforded by the Constitution. The Supreme Court’s admonitions to the federal government to judiciously exercise the ability to conduct subsequent prosecutions under federal law following state prosecutions led to the Justice Department’s 1959 creation of the Petite doctrine, which limits and prioritizes prosecutions where overlapping jurisdiction exists. Over half of the states, at one time, limited or prohibited subsequent prosecutions and half still do. This article examines the dual sovereignty doctrine by analyzing (1) the jurisprudential bases of the doctrine, as well as two recent practical applications, (2) whether the doctrine continues to be constitutionally viable, in light of the increasing amount of federal criminal law in areas traditionally regarded as matters of state primacy and the rise in intergovernmental law enforcement efforts, and (3) the extent and enforceability of limitations placed by the federal and state governments upon the doctrine.

In 1959, the U.S. Supreme Court held in *Bartkus v. Illinois* that prosecutions for essentially the same violation may be conducted by separate sovereigns, with accompanying punishments imposed upon defendants convicted of violating the laws of each.¹ This seeming violation of the Fifth Amendment prohibition against double jeopardy is made possible by a judicial sleight of hand known as the dual sovereignty doctrine, which allows multiple prosecutions as long as they are conducted by wholly separate governmental entities.² However, as Justice Hugo Black so aptly put the issue in his dissent in *Bartkus*, that a different sovereign is conducting the second prosecution doesn’t matter to the person charged; the defendant only cares that he or she is being made to stand trial twice

---

* Dr. Thomas White is an assistant professor of criminal justice at the University of Texas-Pan American. Dr. White received a J.D. from the University of Houston in 1980 and also a Ph.D. in Political Science from Texas A&M University in 2003.


for the same offense.\footnote{3} Justice Black’s observation has since underscored much of the criticism of the dual sovereignty doctrine.

To ameliorate some of the unfairness inherent in multiple prosecutions by different sovereigns, the federal government and many states have established limitations, or even prohibitions, on subsequent prosecutions after an initial prosecution in which double jeopardy has attached.\footnote{4} Part I of this inquiry discusses the dual sovereignty doctrine, focusing on its relationship to the Fifth Amendment prohibition of multiple prosecutions and punishments.\footnote{5} Part II addresses the greater potential for multiple prosecutions occasioned by the increasing “federalization” of criminal law.\footnote{6} Next, Parts III and IV, respectively, examine how the federal and state governments have addressed (or failed to address) the prospect of multiple prosecutions under the dual sovereignty doctrine.\footnote{7} Part V concludes with suggestions aimed at resolving the issues of double jeopardy and dual sovereignty.\footnote{8}

I. THE DOCTRINE OF DUAL SOVEREIGNTY

The theoretical justification for allowing different sovereigns to pursue criminal prosecutions simultaneously based upon the commission of the same or similar acts is grounded in the notions that (1) each sovereign has a right to enforce its own laws and (2) each sovereign’s laws address different interests.\footnote{9} The central inquiry in determining whether multiple prosecutions may constitutionally be pursued is the source of each prosecution’s power and whether it is, indeed, sovereign.\footnote{10}

In the case of state prosecutions, the Supreme Court recognizes that states possessed sovereign authority to enforce criminal laws well before the creation of the United States, and, in the American federal system, the states continue to be sovereign.\footnote{11} Similarly, the Constitution of the United States affords the federal government sovereignty where the Constitution implicitly or explicitly grants national authority.\footnote{12} Thus, each state government and the U.S. federal government are sovereigns independent from one another.\footnote{13}
The Supreme Court had no reason to address the issue of subsequent prosecutions by state governments prior to 1969 because the Fifth Amendment provisions concerning double jeopardy had yet to be incorporated and applied to the states. In 1937, the Court held in *Palko v. Connecticut* that Fifth Amendment protections against double jeopardy, while no doubt important, did not rise to such a level of fundamental fairness as to require incorporation and application to the states under the Fourteenth Amendment’s Due Process Clause.\(^{14}\) In *Palko*, the defendant, Frank Palko, was initially convicted of second degree murder and sentenced to life in prison after crucial prosecutorial evidence was ruled inadmissible at trial.\(^{15}\) The state appealed the suppression ruling, which resulted in a reversal of the conviction and remand for retrial.\(^{16}\) The ensuing trial resulted in a conviction for murder in the first degree and a sentence of death.\(^{17}\)

Employing traditional double jeopardy analysis, Palko argued that a conviction for second degree murder implicitly acquitted him of any greater offense.\(^{18}\) In an opinion by Justice Benjamin Cardozo, the Supreme Court disagreed that the Fifth Amendment guarantee against being put in jeopardy twice for the same offense should be applied to the states by virtue of the Fourteenth Amendment.\(^{19}\) This classic Cardozo opinion established the principle of “selective incorporation,” under which only those matters in the Bill of Rights that are the “very essence of a scheme of ordered liberty” would be incorporated under the Fourteenth Amendment’s Due Process Clause and applied to the states.\(^{20}\) Because *Palko* refused to extend double jeopardy protections to the

---

16. *Id.* at 321.
17. *Id.* at 321-22.
18. *Id.* at 322.
19. *Id.* at 328.
20. *Palko*, 302 U.S. at 324-25. Discussing Palko’s contention that double jeopardy should be applied to the states through the Due Process Clause of the Fourteenth Amendment, Cardozo concluded:

> Is that kind of double jeopardy to which the statute has subjected him a hardship so acute and shocking that our polity will not endure it? Does it violate those ‘fundamental principles of liberty and justice which lie at the base of all our civil and political institutions’? The answer surely must be ‘no.’
states, the Supreme Court created the dual sovereignty doctrine and applied it to the federal government long before the Court incorporated the protection against multiple prosecutions and applied it to the states in *Benton v. Maryland*\(^2\) in 1969.\(^2\)

The Court first recognized the possibility of overlapping state and federal jurisdiction in 1820 in *Houston v. Moore*,\(^2\) and concluded some twenty-five years later in *Fox v. Ohio*\(^2\) and *United States v. Marigold*\(^2\) that, if the same act violated both federal and state law, the act could be prosecuted by either sovereign. In 1852, two years after *Marigold*, the Court first held in *Moore v. Illinois* that the Double Jeopardy Clause of the Fifth Amendment did not bar federal prosecutions subsequent to state prosecutions.\(^2\) Significantly for dual sovereignty purposes, the Court observed in *Moore v. Illinois*:

\[
\text{Id. at } 328 \text{ (quoting Hebert v. Louisiana, } 272 \text{ U.S. } 312, 316 \text{ (1926)). See also Kanowitz & Kanowitz, supra note 12, § 1.13 (discussing the Due Process Clause of the Fourteenth Amendment, its relation to the Bill of Rights, and the inclusion and application to the states of “rights deemed fundamental to the American conception of justice” through the process of selective incorporation).}
\]

- **22.** Hagburg, supra note 2, at 586 (describing the legal history of the dual sovereign doctrine as applied to double jeopardy).
- **23.** In *Houston v. Moore*, 18 U.S. 1 (1820), a state militiamen was prosecuted for failing to report with his detachment when ordered to do so by the President of the United States. 1820 U.S. LEXIS 244, ***2-3. The defendant maintained that a Pennsylvania statute providing for a state court martial sanctioning such behavior was unconstitutional because the offense was one against the United States. 18 U.S. at 32. The U.S. Supreme Court held that the conduct could be punished under the laws of either the state or the United States, since Congress had not statutorily exercised an exclusive right to punish an offender that would preclude state prosecution. Id. at 33-34, 44-45.
- **24.** In *Fox v. Ohio*, 46 U.S. 410 (1847), the defendant’s conviction for passing and uttering a counterfeit coin under the laws of Ohio was affirmed, id. at 435, despite the defendant’s contention that the U.S. Constitution’s express provision of authority to Congress to coin money under Article I, Section 8, Clauses 5 and 6, made an Ohio law prohibiting counterfeiting beyond the purview of the state, since a conviction for counterfeiting in the state court would bar a federal prosecution under the Fifth Amendment. Id. at 432-33. The Court held that there was no double jeopardy protection, since the Fifth Amendment protected only against the federal government. Id. at 434-35 (refusing to apply the Bill of Rights protections to the states under the Due Process Clause of the Fourteenth Amendment (citing Barron v. City of Baltimore, 32 U.S. 243, 250-51 (1833))). The Court’s subsequent enunciation of the doctrine of selective incorporation in *Palko v. Connecticut* undermined the rationale of the Court’s decision in *Fox*.
- **25.** *United States v. Marigold*, 50 U.S. 560 (1850), is another counterfeiting case in which the defendant contested the ability of the United States to make it illegal for a person to import counterfeit coinage. Id. at 566. The Court held against the defendant, stating that the authority of Congress to outlaw counterfeiting was given to Congress pursuant to its power to coin money under Article I, Section 8 and “the correspondent and necessary power and obligation to protect and to preserve in its purity this constitutional currency for the benefit of the nation.” Id. at 568.
- **26.** Moore v. Illinois, 55 U.S. 13, 19-20 (1852). *Moore* involved a defendant charged with harboring and secreting a slave in violation of state law. Id. at 17. The Court held that the Illinois statute, which had as its purpose the prohibition of introducing fugitive slaves into Illinois, was not inconsistent with the provisions of a federal law which prohibited similar conduct if detrimental to the owner of the fugitive slave. Id. at 18. The Court noted that, while the statutes addressed different purposes (the federal statute addressing interference with recapture and reclamation of fugitive slaves by their owners, and the Illinois statute preventing the immigration of certain classes...
Every citizen of the United States is also a citizen of a State or territory. He may be said to owe allegiance to two sovereigns, and may be liable to punishment for an infraction of the laws of either. The same act may be an offence or transgression of the laws of both. . . . That either or both may (if they see fit) punish such an offender, cannot be doubted. Yet it cannot be truly averred that the offender has been twice punished for the same offence; but only that by one act he has committed two offences, for each of which he is justly punishable. He could not plead the punishment by one in bar to a conviction by the other . . . . 27

Although the philosophical concept of dual sovereignty was thus recognized early on, it was not used to permit multiple prosecutions until the 1920s. 28 In 1922, a unanimous Supreme Court held in United States v. Lanza that the Fifth Amendment’s Double Jeopardy Clause did not bar a federal prosecution for violation of the National Prohibition (Volstead) Act 29 following a successful prosecution under state liquor laws because Congress had the power to enact liquor laws under the Eighteenth Amendment (Prohibition) while states maintained the same power as independent sovereigns. 30 The Court held that while state and federal laws may vary,

[T]his is an inseparable incident of independent legislative action in distinct jurisdictions.

. . . .

We have here two sovereignties, deriving power from different sources, capable of dealing with the same subject-matter within the same territory. Each may, without interference by the other, enact laws to secure prohibition . . . . Each government in determining what shall be an offense against its peace and dignity is exercising its own sovereignty, not that of the other.

. . . .

It follows that an act denounced as a crime by both national and state sovereignties is an offense against the peace and dignity of both and may be punished by each. 31
In 1959, the U.S. Supreme Court decided *Abbate v. United States*, which explicitly reaffirmed *Lanza*’s holding and reasoning that, were an accused able to escape federal prosecution by pleading to a lesser state charge, damage would be done to the interests of the federal government, especially if the accused’s actions imposed more seriously upon federal interests.\(^{32}\) Justice Black also dissented in *Abbate*, joined by Chief Justice Earl Warren and Justice William O. Douglas, both of whom joined his dissent in *Bartkus*.\(^{33}\) Black added a new wrinkle to the argument he advanced in *Bartkus*\(^{34}\) by asserting in *Abbate* that the federal and state governments could not properly be regarded as two separate sovereigns, similar to the relationship of actual foreign regimes.\(^{35}\)

Application of Fifth Amendment double jeopardy protections to the states would not occur until 1969 in *Benton v. Maryland*, thirty-two years after the Court’s decision in *Palko*.\(^{36}\) In *Benton*, the Supreme Court finally determined that the guarantees against being twice placed in jeopardy or twice punished for the same offense were of such constitutional magnitude that they should be incorporated and applied to the states under the Fourteenth Amendment’s Due Process Clause.\(^{37}\) By this time, only Connecticut, Maryland, Massachusetts, North Carolina, and Vermont did not have some state constitutional or statutory provisions prohibiting placing a defendant twice in jeopardy.\(^{38}\) As Tables One and Three in Part IV evince, none of these states have since provided constitutional protections, although Connecticut, Massachusetts and Vermont have since enacted statutory protections.\(^{39}\) The courts in all five states have held

\(^{32}\) *Abbate* v. United States, 359 U.S. 187, 195 (1959). The Supreme Court decided *Abbate* the same day as *Bartkus* v. Illinois.

\(^{33}\) Id. at 201.

\(^{34}\) *Bartkus* v. Illinois, 359 U.S. 121, 155 (1959); see infra note 52 and accompanying text.

\(^{35}\) *Abbate*, 359 U.S. at 203 (Black, J., dissenting).


\(^{37}\) *Id.* at 794-96. Justice Thurgood Marshall wrote in the majority opinion:

In an increasing number of cases, the Court ‘has rejected the notion that the *Fourteenth Amendment* applies to the States only a ‘watered-down, subjective version of the individual guarantees of the *Bill of Rights* . . . .’ . . . . For the same reasons, we today find that the double jeopardy prohibition of the *Fifth Amendment* represents a fundamental ideal in our constitutional heritage, and that it should apply to the States through the *Fourteenth Amendment*. Insofar as it is inconsistent with this holding, *Palko* v. *Connecticut* is overruled.


\(^{39}\) CONN. GEN. STAT. § 54-200 (2009) addresses double jeopardy by implication when it states that “[n]o acquittal or conviction for any criminal offense, had upon any complaint issued by the procurement or at the solicitation of the person committing it, shall be a bar to another complaint or information for the same offense.” MASS. GEN. LAWS ch. 263, § 7 (2010) provides: “A person shall not be held to answer on a second indictment or complaint for a crime of which he has been acquitted upon the facts and merits . . . .” VT. STAT. ANN. tit. 13, § 6556 (LexisNexis 2010) provides double jeopardy protection for offenses if the accused “was acquitted by a jury upon the merits on a former trial.” VT. STAT. ANN. tit. 13, § 6557 (LexisNexis 2010) further limits the protection to instances in which a former acquittal is not based on (a) a variance between the
double jeopardy protections to exist under these statutes, other provisions of their state constitutions, or at common law.40

Prior to the Court’s application of double jeopardy protections to the states in Benton, the Supreme Court had occasion in 1959 to determine whether a state prosecution following an unsuccessful federal prosecution violated due process in Bartkus v. Illinois.41 The Supreme Court ruled in a 5-4 decision authored by Justice Felix Frankfurter that it did not.42 After being acquitted of bank robbery under federal law, the defendant was convicted for the same conduct under applicable state law and given a life sentence as a habitual felony offender.43 While the Court acknowledged that (1) the Federal Bureau of Investigation shared the fruits of its investigation with state prosecutors and (2) the sentencing of Bartkus’ co-defendants in federal court was delayed until after they testified against Bartkus in the state court proceedings, the Court concluded that each prosecution was, for all intents and purposes, separately conducted.44 The majority opinion observed that of the twenty-eight state courts that had addressed the issue of state prosecution following federal prosecution, twenty-seven concluded that a prior federal prosecution was not a bar to subsequent proceedings for the same act under state law.45 The Court agreed with the overwhelming majority of state courts and held that subsequent state prosecutions were not generally barred under the Due Process Clause of the Fourteenth Amendment.46 In doing so, the Court recognized the unfairness that could conceivably be visited on states and their interests in enforcing their own

40. See State v. Rawls, 502 A.2d 374, 375 n.3 (Conn. 1985) (“Although the Connecticut constitution contains no specific double jeopardy provision, the due process guarantees of article first, § 8, [of the Connecticut Constitution] have been held to include such a protection.”); Apostledes v. State, 575 A.2d 792, 794 (Md. Ct. Spec. App. 1990) (“The Fifth Amendment to the United States Constitution and the Maryland common law provide that no person should be put in jeopardy twice for the same offense.”); Commonwealth v. Vanetzian, 215 N.E.2d 658, 660 (Mass. 1966), overruled in part on other grounds by Commonwealth v. Lewis, 409 N.E.2d 771, 774 (Mass. 1980) (“Both the common law and our statutes provide that a person may not be twice put in jeopardy for the same offence.”); State v. Crocker, 80 S.E.2d 243, 245 (N.C. 1954) (“It is a fundamental and sacred principle of the common law, deeply imbedded in our criminal jurisprudence, that no person can be twice put in jeopardy of life or limb for the same offense. While the principle is not stated in express terms in the North Carolina Constitution, it has been regarded as an integral part of the ‘law of the land’ within the meaning of Art. I, sec. 17 [of the North Carolina Constitution].” (citations omitted)); State v. Deso, 1 A.2d 710, 715 (Vt. 1938) (“Where one offense is a necessary element in, and constitutes a part of, another, and both are in fact but one transaction, an acquittal or conviction of one is a bar to a prosecution for the other.”).
42. Id. at 121, 139.
43. Id. at 121-22.
44. Id. at 122-24.
45. Id. at 135-36 (noting Florida as the only state not permitting state prosecution following a federal prosecution but, even here, the Court noted that this restriction might be limited to the Florida court’s interpretation of the Eighteenth Amendment instituting Prohibition).
46. Id. at 136.
criminal laws if a former prosecution for a comparatively minor federal offense precluded prosecution for a major state offense.\(^{47}\)

Justice Black, joined by Chief Justice Warren and Justice Douglas (even then, the more liberal bastions of the Court), strongly dissented, arguing that since there was nothing superficially wrong with the initial federal prosecution,\(^{48}\) the Due Process Clause of the Fourteenth Amendment should preclude a subsequent state prosecution, just as the Fifth Amendment would prevent a second federal prosecution.\(^{49}\) Black contended that to hold otherwise “limits our already weakened constitutional guarantees against double prosecutions.”\(^{50}\) Black concluded his dissent with an argument that has resonated throughout the ensuing debate on the constitutional viability of the dual sovereignty doctrine:

> The Court apparently takes the position that a second trial for the same act is somehow less offensive if one of the trials is conducted by the Federal Government and the other by a State. Looked at from the standpoint of the individual who is being prosecuted, this notion is too subtle for me to grasp. If double punishment is what is feared, it hurts no less for two “Sovereigns” to inflict it than for one. If danger to the innocent is emphasized, that danger is surely no less when the power of State and Federal Governments is brought to bear on one man in two trials, than when one of these “Sovereigns” proceeds alone. In each case, inescapably, a man is forced to face danger twice for the same conduct.\(^{51}\)

Justice Black’s words have been quoted frequently in criticisms of the dual sovereignty doctrine over the ensuing fifty years.\(^{52}\)

\(A.\) \textit{Other Sovereigns Included Under the Doctrine}

The Supreme Court has expanded the dual sovereignty doctrine to encompass sovereign relations other than those existing between the federal and state governments. In 1985, a divided Court (7-2) in \textit{Heath v. Alabama} found the application of the dual sovereignty doctrine to successive state prosecutions “inescapable”\(^{53}\) where a husband hired contract killers to abduct his wife from their Alabama home along the Alabama-Georgia border, kill her, and discard the

\(^{47}\) \textit{Bartkus}, 359 U.S. at 137. The Court previously inverted this argument for the protection of federal interests in \textit{Lanza} and \textit{Abbate}. \textit{See supra} note 30 and accompanying text.

\(^{48}\) \textit{See Bartkus}, 359 U.S. at 150 (Black, J., dissenting) (“So far as appears the trial was conducted fairly by an able and conscientious judge.”).

\(^{49}\) \textit{Id.}

\(^{50}\) \textit{Id.}

\(^{51}\) \textit{Id. at} 155.


body in Georgia. The Court found that Heath’s guilty plea and life sentence under Georgia law did not preclude the State of Alabama from prosecuting Heath for capital murder because the states are no less sovereign with respect to each other than they are with respect to the federal government. Using a similar rationale, in 1978 the Court held in United States v. Wheeler that prosecution of a Native American for statutory rape under federal law was not barred by Wheeler’s earlier conviction for a lesser included offense by a Navajo tribal court. The Court reasoned that because American Indian tribes have not given up full sovereignty and retain the attributes of sovereignty as to their people and their territory, Native American tribes are separate sovereigns until Congress acts to remove their sovereignty.

Similar sovereignty issues also arise when jurisdiction lies in multiple countries. Although many European countries regard prosecutions by one nation as prohibitive of a subsequent prosecution by another for the same conduct under the doctrine of non bis in idem, the United States and many of the states have invoked the doctrine of dual sovereignty in subsequent prosecutions to preclude the assertion of double jeopardy protections following prosecutions in foreign jurisdictions. The importance of the doctrine of dual sovereignty to American jurisprudence is underscored by the United States’ successful preservation of the doctrine as a specifically recognized exception to double jeopardy in the International Covenant on Civil and Political Rights, a United Nations treaty authored in 1966 pursuant to the Universal Declaration of Human Rights.

54. Id. at 83-84.
55. Id. at 84-85, 94.
56. Id. at 88-91.
58. Id. at 332.
59. See also William Bradford, Tribal Sovereignty and United States v. Lara: “Another Such Victory and We Are Undone:” A Call to an American Indian Declaration of Independence, 40 TULSA L. REV. 71, 84 (2004) (characterizing the Court’s decision as recognizing the “primeval sovereignty” of the Navajo nation “at least with regard to their own memberships”). But see United States v. Enas, 255 F.3d 662, 664, 667 (9th Cir. 2001), cert. denied, 534 U.S. 1115 (2002) (excepting delegated power to Indian tribes by Congress from dual sovereignty exception because of same source of power).
60. The international law equivalent of double jeopardy. See BLACK’S LAW DICTIONARY (9th ed. 2009) (translated as “[n]ot twice for the same thing” which “essentially refers to the double-jeopardy bar”).
61. See Dax Eric Lopez, Note, Not Twice for the Same: How the Dual Sovereignty Doctrine is Used to Circumvent Non Bis In Idem, 33 VAND. J. TRANSNAT’L L. 1263, 1271-74 (2000) (discussing the evolution and treatment of the concept “non bis in idem”). See also Chua Han Mow v. United States, 730 F.2d 1308, 1309, 1313 (9th Cir. 1983) (prosecuting and convicting Malaysian citizen for importation and distribution of a controlled substance under federal law not barred by his conviction in Malaysia for the same offense).
B. Exceptions to the Doctrine

The Supreme Court has recognized extremely narrow exceptions to the dual sovereignty doctrine: one involving governments collusively acting to maintain multiple prosecutions alluded to in the Court’s opinion in *Bartkus*, and two others involving a state boundary line dispute and territories of the United States. The first of these exceptions is seen in *Puerto Rico v. Shell Co.*, where the Court held that the doctrine had no application where the federal government and a territory of the United States engaged in dual prosecutions because territories, unlike states, are not separate sovereigns with respect to the U.S. government. Using the same rationale, the D.C. Circuit Court of Appeals prohibited prosecutions by the District of Columbia following federal prosecutions because the district is a federal territory.

The second exception is displayed in *Nielsen v. Oregon*, a peculiar case in which the alleged crime was committed in a river that comprised the boundary between two states. The Court held that both states had concurrent jurisdiction over the river, and one state could not prosecute a citizen of the second for doing something that the other condones. Specifically, Oregon could not prosecute Nielsen for operating a “purse net” when Nielsen held a valid license to use the net from the State of Washington. Because of the unusual fact situation, the Supreme Court limited its holding in *Nielsen* to the specific facts of that case when later deciding *Heath v. Alabama*.

A third limitation to the dual sovereignty doctrine exists where state and federal prosecutors conspire to reprosecute a defendant following an unsuccessful prosecution, the possibility of which was alluded to by the Supreme Court in dicta in *Bartkus*. While state and federal officials are free to...
cooperate, they may not manipulate the system to get a second bite at the apple.\textsuperscript{74} To date, the Supreme Court has not reversed a conviction on this basis.\textsuperscript{75} In \textit{Bartkus}, the Court did not address the level of intergovernmental collusion or cooperation necessary to sustain such a claim, nor has it done so since.\textsuperscript{76} \textit{United States v. Belcher}\textsuperscript{77} is the only lower court finding of a double jeopardy bar to a subsequent prosecution based on intergovernmental collusion not reversed on appeal.\textsuperscript{78} Belcher is unique for several reasons. First, the state’s attorney who unsuccessfully prosecuted the defendants in state court was appointed as a special prosecutor in the federal case,\textsuperscript{79} and second, the ruling in the defendants’ favor was never appealed.\textsuperscript{80}

\textbf{C. Recent Applications of the Dual Sovereignty Doctrine I—The Rodney King Case}

One of the most prominent recent examples of the effects of the dual sovereignty doctrine surrounded Rodney King’s beating at the hands of law enforcement in Los Angeles in 1991. Even with a videotape of the officers beating Mr. King in evidence, a state trial jury acquitted the officers of any criminal wrongdoing under California law.\textsuperscript{81} The United States then prosecuted the officers for the same conduct, alleging that the police violated Mr. King’s federally protected civil rights.\textsuperscript{82} The officers’ conviction in federal court renewed interest in the dual sovereignty doctrine, especially as applied to civil rights statutes.\textsuperscript{83}

\begin{itemize}
  \item Id.
  \item Dawson, supra note 75, at 296 n.113.
  \item Belcher, 762 F. Supp. at 668.
  \item See Dawson, supra note 75, at 296 n.113 (noting that the Belcher decision was never appealed in the context of delineating a number of other citations rejecting sham prosecution claims).
  \item See United States v. Koon, 34 F.3d 1416, 1424-25 (9th Cir. 1994) [hereinafter \textit{Koon II}] (reciting the factual and legal background of the state case leading to the filing of civil rights charges in federal court). Technically, one police officer was not acquitted on one count, but there was a hung jury on that count. \textit{Id.} at 1425. A much more detailed recitation of the videotape depiction of the injuries inflicted on Mr. King is contained in the district court’s Sentencing Memorandum, \textit{United States v. Koon}, 833 F. Supp. 769, 774-80 (C.D. Cal. 1993).
  \item \textit{Koon II}, 34 F.3d at 1425.
  \item The initial acquittal of the officers involved in the Rodney King beating and their subsequent conviction in federal court spawned a veritable cottage industry of law review and academic journal articles on the dual sovereignty doctrine. See generally Akhil Reed Amar & Jonathan L. Marcus, \textit{Double Jeopardy Law After Rodney King}, 95 COLUM. L. REV. 1 (1995); Susan N. Herman, \textit{Double Jeopardy All Over Again: Dual Sovereignty, Rodney King, and the ACLU}, 41 UCLA L. REV. 609 (1994); Hoffman, supra note 62; Merkl, supra note 10; Christina Galye
\end{itemize}
Remarkably, the Ninth Circuit Court of Appeals gave short shrift to the defendants' argument that they had been subjected to Bartkus-like collusion by state and federal prosecutors, as did the U.S. Supreme Court, which refused to grant certiorari on the issue. The only issue accepted by the Supreme Court for certiorari review regarded the appropriate standard of review governing appeals from a district court’s departure under the federal sentencing guidelines in assessing punishment against the convicted officers.

D. Recent Applications of the Dual Sovereignty Doctrine II—United States vs. Robert Angleton

A high profile Texas case that was denied certiorari by the U.S. Supreme Court provides a much better example of the current debate. Robert Angleton, a Houston bookie and police informant, was indicted for capital murder in 1997 for allegedly paying his brother to kill Angleton’s wife. When Angleton’s brother committed suicide in jail in 1998, he left a note claiming sole responsibility for the murder. At the state trial, Angleton was acquitted by a jury of all charges. At the time, Angleton was the only person tried for capital murder in state court in Harris County to be acquitted since the reinstatement of capital punishment in Texas in 1976.

Six months after his acquittal, federal law enforcement agents began investigating Angleton for potential Racketeer Influenced and Corrupt Organizations Act (RICO) charges related to his gambling activities and income tax evasion, but did not instigate an investigation into the murder of Angleton’s wife until the Harris County District Attorney’s Office sought assistance from
In 2000, the U.S. Attorney’s Office requested that the Federal Bureau of Investigation (FBI) expand its investigation to include the murder. The request resulted in the creation of an “ad hoc task force” of FBI agents and the Houston Police Department (HPD) officers who originally investigated the homicide. The HPD officers were sworn in as U.S. Deputy Marshals to allow them access to FBI files. The joint task force was (1) allowed access to all of the evidence used to prosecute Angleton in the state trial, (2) provided assistance by the state assistant district attorneys who prosecuted the capital murder case, and (3) allowed to interview the state jurors who served on the capital murder jury to determine why they acquitted Angleton.

Although the investigation resulted in no RICO or tax evasion charges, in 2002 a grand jury indicted Angleton under a federal murder-for-hire statute that required the jurisdictional element of some involvement in interstate commerce. At the hearing on Angleton’s motion to dismiss the federal indictment on double jeopardy grounds, both the government and the defense agreed that the federal prosecution would rely upon much of the same evidence utilized in the state prosecution. The trial court denied Angleton’s motion, but acknowledged that the state and federal crimes were identical for purposes of determining double jeopardy under United States v. Blockburger, the U.S. Supreme Court opinion that established the test for determining whether offenses are the same for Fifth Amendment double jeopardy purposes when prosecuted by a single sovereign.

Angleton’s attorney, distinguished Houston criminal defense lawyer Mike Ramsey, contended that there was Bartkus-like collusion between the State of

92. Id.
93. Id.
94. Angleton, 221 F. Supp. 2d at 700.
95. Id.
96. Id.
97. Id. at 701. Angleton was indicted under 18 U.S.C. § 1958 (1996), “[u]se of interstate commerce facilities in the commission of murder-for-hire,” which carries a punishment of death or life imprisonment, or a fine not to exceed $250,000.00, or both. The interstate commerce allegation contended Angleton caused his brother to travel in interstate commerce from California to Texas to murder Angleton’s wife. Id. at 701.
98. See Angleton, 221 F. Supp. 2d at 700 (noting that the parties agreed that the federal government would use much of the same evidence as used in the state court).
99. Id. at 740-41.
100. Id. at 706.
101. United States v. Blockburger, 284 U.S. 299 (1932). Since the Double Jeopardy Clause protects against multiple prosecutions by the same sovereign for the same offense, the Court in Blockburger faced the task of defining what constitutes the same offense since the Constitution does not. The Court responded:

The applicable rule is that where the same act or transaction constitutes a violation of two distinct statutory provisions, the test to be applied to determine whether there are two offenses or only one, is whether each provision requires proof of a fact which the other does not.

Id. at 304. The test has become popularly known as “the Blockburger Test.”
Texas and the federal government following the state’s unsuccessful prosecution. In support, Ramsey cited the extensive involvement of state officials in the federal prosecutorial efforts. The trial court responded that the facts of Bartkus and many other subsequent prosecution cases considered by the federal courts indicated there could be extensive involvement between governments without such cooperation rising to the level of a collusive prosecution. Among the examples cited by the trial court were prosecutions in which (1) state sovereigns requested federal prosecution, (2) the initial sovereign turned over all of its evidence to the second, and (3) cross designations of state officials as federal officials were made. The court concluded it could find no collusion on the record before it.

The trial court also ruled against Angleton’s argument that the findings by the state jury barred any inconsistent findings in a retrial under the doctrine of collateral estoppel for two reasons. First, the federal government was not a party to the initial prosecution, and, second, collateral estoppel normally does not apply to successive prosecutions by state and federal governments under controlling authority in jurisdictions under the domain of the Fifth Circuit Court

102. Angleton, 221 F. Supp. 2d at 713.
103. Id.
104. Id. at 713-15.
105. Id. at 713, 715, 720.
106. Id. at 722.
107. In Ashe v. Swenson, 397 U.S. 436, 443-44 (1970), the Court described the concept and application of “collateral estoppel” as follows:

‘Collateral estoppel’ is an awkward phrase, but it stands for an extremely important principle in our adversary system of justice. It means simply that when an issue of ultimate fact has once been determined by a valid and final judgment, that issue cannot again be litigated between the same parties in any future lawsuit. Although first developed in civil litigation, collateral estoppel has been an established rule of federal criminal law at least since this Court’s decision more than 50 years ago in United States v. Oppenheimer, 242 U.S. 85.

The federal decisions have made clear that the rule of collateral estoppel in criminal cases is not to be applied with the hypertechnical and archaic approach of a 19th century pleading book, but with realism and rationality. Where a previous judgment of acquittal was based upon a general verdict, as is usually the case, this approach requires a court to ‘examine the record of a prior proceeding, taking into account the pleadings, evidence, charge, and other relevant matter, and conclude whether a rational jury could have grounded its verdict upon an issue other than that which the defendant seeks to foreclose from consideration.’ (footnote omitted). The inquiry ‘must be set in a practical frame and viewed with an eye to all the circumstances of the proceedings.’ Sealfon v. United States, 332 U.S. 575, 579. Any test more technically restrictive would, of course, simply amount to a rejection of the rule of collateral estoppel in criminal proceedings, at least in every case where the first judgment was based upon a general verdict of acquittal. (footnote omitted).
of Appeals, which encompasses Texas. Ramsey responded to the local press: “The federal system is arrogant to the extent it believes only it can do justice. If a state jury is not to be believed when it acquits, how can it be believed when it convicts?” The question was rendered moot when Angleton subsequently fled to the Netherlands using a fraudulent passport.

Angleton was detained by Dutch authorities, who refused to extradite him to the United States to be retried for the murder-for-hire offense or the related failure to appear charge. Dutch authorities did, however, agree to extradite Angleton on three charges related to passport fraud. Under the extradition treaty between the United States and the Netherlands, a person extradited on some charges, but not others, can only be prosecuted for those charges for which he or she is extradited. Upon his return, Angleton pleaded guilty to the federal passport charges and was assessed a sixty month prison sentence that was affirmed by the Fifth Circuit in an unpublished opinion.

II. PROBLEMS OCCASIONED BY THE FEDERALIZATION OF CRIMINAL LAW

Although seemingly at odds with the Fifth Amendment protections against double jeopardy, the topic of dual sovereignty was not particularly contentious until the latter part of the twentieth century when federal criminal statutes increasingly addressed conduct already criminalized under state law. In the American conception of federalism, matters pertaining to health and safety have traditionally been considered matters of state primacy. Consequently, states

111. Id.
112. Id.
114. Angelton, 2006 U.S. App. LEXIS 24449, at *2-4. See 5TH CIR. R. 47.5 (limiting publication of “opinions that merely decide particular cases on the basis of well-settled principles of law” of little interest to persons other than the parties).
115. See Merkl, supra note 10, at 175 (“By permitting subsequent prosecution by a State or the Nation for the same underlying criminal acts, the Supreme Court vitiates the Fifth Amendment’s [double jeopardy protection].”).
116. See generally Merkl, supra note 10 (discussing the expansion of federal criminal law).
117. See James A. Strazzella, The Federalization of Criminal Law: Task Force on Federalization of Criminal Law, 1998 A.B.A. SEC. CRIM. JUST.: Constitutional law recognizes that ‘preventing and dealing with crime is much more the business of the States than it is of the Federal Government . . .’ (quoting Patterson v. New York, 432 U.S. 197, 201 (1977)). In practice, most criminal conduct in America has always been, and still is, defined by state legislatures, investigated by state agents, prosecuted by state prosecutors, tried in state courts, and punished in state prisons. This accords with the historical American principle that the general police power lies with the states and not
historically have determined what conduct should be criminalized, as well as the appropriate sanctions for each crime.\textsuperscript{118} In 1903, the Supreme Court ruled in \textit{Champion v. Ames} that the Commerce Clause was an appropriate vehicle for Congress to use to supplement state criminal law.\textsuperscript{119} Following \textit{Champion v. Ames}, Congress passed an increasing number of criminal laws dealing with conduct customarily falling within the purview of state law. In the latter half of the twentieth century, particularly the last quarter, a marked expansion of federal criminal law occurred.\textsuperscript{120}

Given the increasing overlap between state and federal criminal laws, many scholars argue that, even though dual sovereignty may have once served a useful purpose when federal criminal law was limited to conduct pertaining distinctly to the national government, this is no longer the case with the increasing federalization of criminal law.\textsuperscript{121} Instances abound in which federal criminal statutes overlap existing state law.\textsuperscript{122} For example, the federal government now has statutes addressing failure to pay child support (The Child Support Recovery Act),\textsuperscript{123} murder and robbery (The Hobbs Act),\textsuperscript{124} and auto theft/carjacking (The Anti-Car Theft Act).\textsuperscript{125}

\begin{flushright}
with the federal government, although there clearly is an appropriate sphere for federal criminal legislation.
\end{flushright}

\textit{Id.} at 24-25.

\begin{flushleft}
118. \textit{Id.}
\end{flushleft}

\begin{flushleft}
119. In \textit{Champion v. Ames}, 188 U.S. 321 (1903), the Supreme Court upheld Charles Champion’s conviction for transporting Paraguayan lottery tickets from Texas to California via the Wells Fargo Company under a federal statute that prohibited sending lottery tickets in interstate or foreign commerce. \textit{Id.} at 363-64. The Court held:

We decide nothing more in the present case than that lottery tickets are subjects of traffic among those who choose to sell or buy them; that the carriage of such tickets by independent carriers from one State to another is therefore interstate commerce; that under its power to regulate commerce among the several States Congress—subject to the limitations imposed by the Constitution upon the exercise of the powers granted—has plenary authority over such commerce, and may prohibit the carriage of such tickets from State to State; and that legislation to that end, and of that character, is not inconsistent with any limitation or restriction imposed upon the exercise of the powers granted to Congress.

\textit{Id.}
\end{flushleft}

\begin{flushleft}
\end{flushleft}

\begin{flushleft}
121. \textit{See, e.g.}, Owlsley, \textit{supra} note 52, at 779-80 nn.83-85 (citing a variety of articles critical of the dual sovereignty doctrine).
\end{flushleft}

\begin{flushleft}
122. \textit{See Beale, \textit{supra} note 120, at 42-44 (listing laws enacted by Congress which overlap with existing state law).
\end{flushleft}

\begin{flushleft}
The American Bar Association (ABA) has bemoaned the increase in federal criminal laws.126 A task force created by the ABA in 1998, chaired by former U.S. Attorney General Edwin Meese III,127 revealed that federal criminal law has increased in astonishing proportions in recent decades, blurring the line of what constitutes a federal crime.128 The task force found that over forty percent of the federal criminal law enacted since the American Civil War was enacted after 1970, with an estimated 1000 bills addressing criminal acts being introduced in the 105th Congress through July 31, 1998 alone.129 The Task Force characterized the trend toward federalization as troubling, and expressed deep concern that local values would be lost:

Local crimes involve local values and should be handled by state law. . . . Community views also differ from state to state on related issues . . . In the participatory democracy of our large nation, with varying local values, citizen views about such matters are more likely to be felt and acted upon through representatives at the local level, rather than at the federal level where most of those in power are more

\[124. \text{18 U.S.C. § 1951 (2006) makes a federal offense of conduct interfering with interstate commerce by threat of violence. See James S. Gorlick & Harry Litman, Prosecutorial Discretion and the Federalization Debate, 46 Hastings L.J. 967, 973 (1995) (claiming that the statute “potentially federalizes any convenience store hold-up”). See also Litman & Greenberg, supra note 123, at 83 (recognizing that without federal prosecutorial discretion, Hobbs Act prosecutions could “result in federal prosecution of most convenience store holdups, duplicating and or supplanting state prosecutions and expanding the business of the federal criminal justice system exponentially”).}

125. 18 U.S.C. § 2119 (2006) makes carjacking a federal offense if a person, while possessing a firearm, "takes a motor vehicle that has been transported, shipped, or received in interstate or foreign commerce from the person or presence of another by force and violence or by intimidation, or attempts to do so." The statute has been described as “making almost any theft of an automobile a federal offense, intrud[ing] on a traditionally local area of law enforcement without lowering the crime rate.” 126 Stephen Chippendale, Note, More Harm Than Good: Assessing Federalization of Criminal Law, 79 Minn. L. Rev. 455, 479 (1994).

126. The Task Force on the Federalization of Criminal Law remonstrated:

\[\text{The expanding coverage of federal criminal law, much of it enacted in the absence of a demonstrated and distinctive federal justification, is moving the nation rapidly toward two broadly overlapping, parallel, and essentially redundant sets of criminal prohibitions, each filled with differing consequences for the same conduct. Such a system has little to commend it and much to condemn it.} \]

Strazzella, supra note 117.

127. Id. at 1.
128. Id. at 2.

129. \text{Id. at 7, 10-11 & n.15. See also Laura G. Dooley, The Dilution Effect: Federalization, Fair Cross-Sections and the Concept of Community, 54 DePaul L. Rev. 79, 102 (2004); Merkl, supra note 10, at 178, n. 11 & 13 (providing the findings of the task force).}
removed from the affected local values and more preoccupied with issues of national and international concern. 130

The First Congress initially established only seventeen federal crimes when it passed the Crimes Act of 1790. 131 All seventeen offenses directly implicated federal interests. 132 Federal criminal law increased incrementally thereafter, until a significant federalization of criminal law began with the passage of the Thirteenth, Fourteenth, and Fifteenth Amendments following the Civil War. 133 The expansion continued into the twentieth century with the passage of the National Prohibition (Volstead) Act. 134 More recent federal laws have addressed such problems as organized crime and drug and illegal alien trafficking. 135 Today, there are in excess of 3000 federal criminal laws throughout the United States Code.136

The Supreme Court has only recently shown any inclination to invalidate federal statutes due to a lack of federal jurisdiction under the Commerce Clause, such as the Gun Free School Zones Act and the Violence Against Women Act, considered in United States v. Lopez 137 and United States v. Morrison, 138

130. Strazzella, supra note 117, at 44.
132. See Beale, supra note 120, at 40 (describing the pre-Civil War crimes as involving treason, bribery of federal officials, perjury in federal court, theft of federal government property, and revenue fraud); Kathleen F. Brickey, Criminal Mischief: The Federalization of American Criminal Law, 46 HASTINGS L.J. 1135, 1137 (1995).
133. See Brickey supra note 132, at 1139-40 (explaining how the federal criminal law expanded with the enactment of federal civil rights acts). Commonly referred to as “the Civil War Amendments,” the last sections of the Thirteenth, Fourteenth, and Fifteenth Amendments specifically authorize Congress to enforce the amendments by appropriate legislation.
136. See Beale, supra note 120, at 44; Dooley, supra note 129, at 101 n. 132 (both approximating the number at more than 3000). The ABA Task Force on Federalization of Criminal Law was of the opinion, based on its findings, that the actual number of federal crimes was “unquestionably larger” than the frequently cited 3000 figure and growing. Strazzella, supra note 117, at 9-11, n. 11.
respectively. With the expansion of federal criminal law and the rise of cooperative federalism in law enforcement, many have argued that the concept of dual sovereignty is at odds with the original intent of the Founders. 139

The increase in the number of overlapping federal laws also comes at a time when jurisdictional disputes concerning the province of the national and state governments have increased. 140 While some cite cooperative federalism in law enforcement as a reason the dual sovereignty doctrine is outmoded, 141 a festering of tensions also exists in other areas where the federal government has entered realms traditionally left to the states. 142 Often these intrusions, particularly in the area of civil and regulatory law, preempt rather than coexist with state law addressing the same issue. 143 Some scholars have gone so far as to question

561-68. The Court concluded that the government’s arguments in its brief on appeal (Congress made no explicit findings regarding how gun possession in public schools affected interstate commerce) that possession of firearms in a school zone affects the national economy because of (1) the costs of increasing violent crime are passed on to the entire population through insurance, (2) violent crime in certain areas decreases the willingness of citizens to travel to these areas, and (3) possession of guns in schools threatens the learning environment in ways that will result in a less productive citizenry, were too remotely connected to interstate commerce. Id. at 563-64. The Court observed that, were it to accept the government’s reasoning:

[It] is difficult to perceive any limitation on federal power, even in areas such as criminal law enforcement or education where States historically have been sovereign. Thus, if we were to accept the Government's arguments, we are hard pressed to posit any activity by an individual that Congress is without power to regulate.

Id. at 564.

138. United States v. Morrison, 529 U.S. 598, 617 (2000). One of Morrison’s victims, Christy Brzonkala, appealed a ruling by the Fourth Circuit Court of Appeals holding unconstitutional the Violence Against Women Act, 42 U.S.C. § 13981 (1994), which created a federal civil remedy for victims of gender-based violence. Id. at 605. Writing for the majority, Chief Justice Rehnquist stated Congress had overstepped its authority under the Commerce Clause, id. at 613-19, and the legislation could not be premised upon the enforcement provisions of the Fourteenth Amendment because the Fourteenth Amendment protected only against actions by state rather than private actors. Id. at 621-27.

139. See, e.g., Paul G. Cassell, The Rodney King Trials and the Double Jeopardy Clause: Some Observations on Original Meaning and the ACLU’s Schizophrenic Views of the Dual Sovereignty Doctrine, 41 UCLA L. REV. 693, 709-15 (1994) (asserting that, although not free from doubt, a cogent argument can be made that the Framers intended to adopt double jeopardy as it existed under English Common Law which allowed an accused to raise the defense of autrefois acquit, loosely translated as “formerly acquitted,” if he or she were acquitted in a court of competent jurisdiction, but also providing possible objections to such an interpretation).

140. See Litman & Greenberg, supra note 123, at 73-74 (recognizing how the expansion of federal criminal law has been criticized for “the potential dire consequences for federalism” due to the concurrent state and federal jurisdiction of many crimes).

141. Daniel A. Braun, Praying to False Sovereigns: The Rule Permitting Successive Prosecutions in the Age of Cooperative Federalism, 20 AM. J. CRIM. LAW 1, 7-9 (1992) (stating that in the current age of cooperative federalism, that double jeopardy protections should be re-examined).

142. See infra note 143 (listing federal preemption in products liability as an example of federal intrusions in other area of the law).

143. See David G. Owen, Federal Preemption of Products Liability Claims, 55 S.C. L. REV. 411, 441 (2003); Catherine M. Sharkey, Products Liability Preemption: An Institutional Approach,
whether increased federal involvement in traditional areas of state primacy threatens the structure of American federalism.\textsuperscript{144}

In order for Congress to criminalize conduct, there must be a constitutional basis.\textsuperscript{145} Notwithstanding the Supreme Court’s opinions in \textit{Lopez} and \textit{Morrison}, the Court’s increasing willingness over the past thirty years to find federal jurisdiction under the Commerce and Necessary & Proper Clauses and the enforcement provisions of the Fourteenth Amendment has resulted in a massive increase in the number of topics now deemed appropriate subjects of federal criminal legislation and regulation.\textsuperscript{146} As a result, many have been critical of the dual sovereignty doctrine,\textsuperscript{147} and especially critical of its continuation in mundane cases where it serves no useful purpose.\textsuperscript{148} Scholars have expressed concern that the increasing federalization of criminal law and the resulting overlap with state laws addressing the same subject matter, coupled with the Supreme Court’s continuing recognition of the dual sovereignty doctrine, vitiates

\textsuperscript{76} GEO. WASH. L. REV. 449, 482-84 (2008) (discussing arguments for and against federalization of products liability law).

\textsuperscript{144} See Bradley Scott Shannon, \textit{The Jurisdictional Limits of Federal Criminal Child Pornography Law}, 21 U. HAW. L. REV. 73, 125 n.185 (1999) (listing a substantial number of articles and scholarly works critical of the overfederalization of criminal law and its intrusion upon areas traditionally regarded as matters of state concern); Note, \textit{Mens Rea in Federal Criminal Law}, 111 HARV. L. REV. 2402, 2416-19 (1998) (expressing concerns that the federalization of crime is (1) contrary to the presumption that states are in a better position to address criminal matters, (2) opens the door to selective federal prosecution, (3) overcrowds the dockets of the federal courts, and (4) threatens the balance of power between the states and the federal government). In the 1998 Year-End Report on the Federal Judiciary, Chief Justice Rehnquist expressed the following concerns regarding the increased federalization of crime:

\begin{quote}
The trend to federalize crimes that traditionally have been handled in state courts not only is taxing the Judiciary’s resources and affecting its budget needs, but it also threatens to change entirely the nature of our federal system . . . Federal courts were not created to adjudicate local crimes, no matter how sensational or heinous the crimes may be. State courts do, can, and should handle such problems.
\end{quote}


\textsuperscript{145} See Screws v. United States, 325 U.S. 91, 109 (1945) ("[T]he administration of criminal justice rests with the States except as Congress, acting within the scope of those delegated powers, has created offenses against the United States.").

\textsuperscript{146} See Merkl, supra note 10, at 77-80 (listing recent federal criminal statutes which are arguably examples of Congress’ overreaching, but stating “regardless of how one assesses the merits of the new laws, one thing is certain: Congress continues to enact federal crimes, creating a greater overlap with state law").

\textsuperscript{147} See generally Herman, supra note 82, at 618 n.32 (1994) (listing over ten scholarly works generally critical of the dual sovereignty doctrine). A note by Yale law student Michael Dawson maintained that the dual sovereignty doctrine undermined the concept of popular sovereignty by taking the final word away from juries impinging upon the rights of criminal defendants. Dawson, supra note 75, at 282, 299-302.

much of the double jeopardy protection an accused would otherwise be afforded under the Fifth Amendment. 149 Some claim that the dual sovereignty doctrine constitutes little more than a form of prosecutorial appeal from an adverse result not permitted at law. 150

In his 1998 Year-End Report on the Federal Judiciary, Chief Justice Rehnquist abjured what he described as Congress’s increasing tendency to federalize crimes. 151 Noting that federal criminal cases had tripled since 1997, 152 Rehnquist saw Congress’s actions as taxing the limited resources of the federal judiciary and related agencies, such as pre-trial services and the prosecutorial arm of the Department of Justice (DOJ). 153 Rehnquist also perceived this tendency as threatening to alter the nature of the entire federal system in the United States. 154 Invoking Abraham Lincoln and Dwight D. Eisenhower, Rehnquist argued that matters able to be “adequately handled” by the state courts should be left to those courts. 155

In the report, Rehnquist points to a 1995 recommendation by the Judicial Conference that advocated making the federal courts a “distinctive judicial forum of limited jurisdiction” exercising criminal jurisdiction “only to further clearly defined and justified national interests . . . .” 156 Rehnquist endorsed the Conference’s recommendation that federal criminal jurisdiction should only extend to five types of cases:

1. offenses against the federal government or its inherent interests;
2. criminal activity with substantial multi-state or international aspects;
3. criminal activity involving complex commercial or institutional enterprises most effectively prosecuted using federal resources and expertise;
4. serious high level or widespread state or local government corruption; and
5. criminal cases raising highly sensitive local issues. 157

The ABA Special Task Force on the Federalization of Criminal Law echoed Rehnquist’s concerns. 158 Others have made similar suggestions, but often argue

---

149. See e.g. Merkl, supra note 10, at 175 (criticizing the dual sovereignty exception to the double jeopardy protections).
150. Id. at 175-76, 193 (stating that the dual sovereignty exception to double jeopardy allows “one government effectively to appeal the acquittal of another [government]” even though the same government could not do so).
152. Id. at 2.
153. Id. at 2-3.
154. Id. at 2.
155. Id. (attributing to both the principle that “matters that can be handled adequately by the states should be left to them; matters that cannot be so handled should be undertaken by the federal government”).
156. Id. at 3.
157. Rehnquist, supra note 144, at 3.
for the inclusion in federal criminal statutes of areas where there has been “demonstrated state failure,” such as enforcement of civil rights.159

The federalization of criminal law may also affect a defendant’s ability to obtain a fair trial. At least one law professor argues that the federalization of crime results in the dilution of minority representation in jury pools in urban communities.160 Laura Dooley maintains that when jurisdiction is moved from state to federal court, the eligible jurors are usually selected from a much larger geographic area than would be the case with a state venire that would be limited to the confines of the city or, more typically, the county or parish.161 The resulting federal jury pool also includes potential jurors from the surrounding “collar counties” and bedroom communities.162 Dooley maintains that these

158. Strazzella, supra note 117, at 49-50. The ABA Task Force also concluded, among other things, that inappropriate federalization of criminal law: (1) causes “serious problems to the administration of justice . . . [by] threaten[ing] fundamental allocations of responsibility between state and federal authorities;” (2) “disrupts the important constitutional balance of federal and state systems;” (3) has a detrimental impact on state criminal justice actors who bear the brunt of criminal law enforcement; (4) potentially “relegates the less glamorous prosecutions to the state system” which may have adverse effects on citizen perception, power, and confidence in state and local mechanisms; (5) “creates an unhealthy concentration of policing power at the federal level;” (6) “increases unreviewable federal prosecutorial discretion;” (7) may cause an adverse impact on the federal prison system; (8) leads to “inappropriately disparate results for similarly situated defendants” depending on the system in which they are prosecuted; (9) “accumulates a large body of law” that Congress must continually update and address while “divert[ing] Congressional attention” from matters that would be more appropriate for federal prosecutorial action; and (10) “represents an unwise allocation of scarce resources needed to address the genuine issues of crime.” Id. at 50.

159. See George D. Brown, Constitutionalizing the Federal Criminal Law Debate: Morrison, Jones, and the ABA, 2001 U. ILL. L. REV. 983, 989-90 (2001) (noting that federal prosecutorial efforts may sometimes be superior to state prosecutions and that the federal government may be able to address areas of “demonstrated state failure”); Rory K. Little, Myths and Principles of Federalization, 46 HASTINGS L.J. 1029, 1077-81 (1995) (regarding superior federal prosecutorial resources, and utilizing the term “demonstrated state failure” to refer to both failures occasioned by inadequate resources and those related to unwillingness to pursue prosecutions, as was experienced in some areas of the South during the Civil Rights Era); Geraldine Szott Moohr, The Federal Interest in Criminal Laws, 47 SYRACUSE L. REV. 1127, 1142-43 (1997) (concerning areas of “demonstrated state failure,” and proposing the creation of rebuttable presumptions against expanding federal jurisdiction conditioned upon a showing state prosecutions are “demonstrably inadequate”). Like Little’s article, Moohr recognizes areas of failure occasioned by both inability and unwillingness to prosecute. Id. See also Herman, supra note 83, at 611-12 (explaining that, within the ACLU a few years before the Rodney King beating, “there was little disagreement among the members of the National Board of Directors that the ACLU should oppose the dual sovereignty doctrine as a harmful, misconceived, unsustainable legal fiction,” but also noting a reluctance of the Board to abandon dual sovereignty in the area of civil rights violations). Herman notes that in 1989, the ACLU’s Due Process Committee recommended that the ACLU oppose dual sovereignty without exception. Id. at 615 & nn.23-24.

160. See generally Dooley, supra note 129.

161. Id. at 81, 87-88. (describing how federal juries are selected and noting that they must be drawn from the entire district, which often includes multiple counties).

162. Id. at 81.
suburban areas dilute minority representation because they are likely to have fewer minorities in their populations.\textsuperscript{163}

III. THE PETITE POLICY—THE UNITED STATES RESPONSE TO DUAL SOVEREIGNTY AND PRIORITIZATION OF PROSECUTIONS

The dual sovereignty doctrine and the growing tendency to federalize much of the criminal law have been much maligned.\textsuperscript{164} British Common Law and the law in most other common-law heritage countries, as well as the law of those American states identified in Tables Two and Three in Part IV, regard a previous prosecution in which double jeopardy attaches as a limitation or bar to a subsequent prosecution.\textsuperscript{165} In response to concerns expressed by the U.S. Supreme Court regarding the dual sovereignty doctrine in \textit{Bartkus v. United States} and \textit{Abbate v. United States},\textsuperscript{166} the federal government took steps beginning in the late 1950s to curtail potential abuses of the doctrine.\textsuperscript{167}

\textit{A. The Petite Policy—\textemdash\textemdash}

The DOJ first promulgated guidelines concerning subsequent prosecutions during the Eisenhower Administration.\textsuperscript{168} The guidelines addressed when the federal government should prosecute after a state prosecution, and which sovereign should have priority in conducting prosecutions.\textsuperscript{169} The multiple prosecution aspect of the guidelines came to be known as the “Petite Policy” after the 1960 Supreme Court per curiam opinion that acknowledged its existence.\textsuperscript{170} In \textit{Petite}, the Supreme Court granted the government’s motion to vacate Petite’s second conviction in federal court following a conviction on

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{163} Id. at 88-90.
\item \textsuperscript{164} See e.g., Beale, supra note 148, at 771 n.117; Herman, supra note 83, at 618 n.32. (listing scholarly articles generally critical of the dual sovereignty doctrine as a basis to circumvent double jeopardy protections).
\item \textsuperscript{165} See Merkl, supra note 10, at 186 & nn.52-53 (quoting William Blackstone for the English Common Law principle that, if “a man is once fairly found not guilty upon an indictment, or other prosecution, before any court having competent jurisdiction of the offence, he may plead such acquittal at bar of any subsequent accusation for the same crime” and noting that the English practice of using writs of \textit{autrefois acquit} and \textit{autrefois convict} continues today in Australia).
\item \textsuperscript{166} \textit{Bartkus v Illinois}, 359 U.S. 121 (1959) is most commonly cited as the impetus for the \textit{Petite Policy}. The language of the majority opinion in \textit{Bartkus} expressly cited the need for restraint, stating “[t]he greatest self-restraint is necessary when the federal system yields results with which a court is in little sympathy.” \textit{Id.} at 138-39. In \textit{Abbate}, the Supreme Court refused to overrule the dual sovereignty doctrine because there was a possibility federal law enforcement efforts would be hindered if they were barred by state prosecutions when a defendant’s conduct impinged more heavily on federal interests. \textit{Abbate v. United States}, 359 U.S. 187, 195 (1959).
\item \textsuperscript{168} Id.
\item \textsuperscript{169} Id. at 178 n. 69.
\item \textsuperscript{170} Petite v. United States, 361 U.S. 529, 530-31 (1960).
\end{enumerate}
\end{footnotesize}
related charges in the Eastern District of Pennsylvania.\textsuperscript{171} The Court’s action rendered it unnecessary for the Court to reach the double jeopardy issue raised by the defendant.\textsuperscript{172} However, Justice William Brennan, joined by Justices Black and Douglas, concurred to opine that the case should have been reversed outright on double jeopardy grounds.\textsuperscript{173}

The DOJ policy regarding successive prosecutions was actually initiated the year before the \textit{Petite} decision by U.S. Attorney General William Rogers in a memorandum addressed to U.S. Attorneys’ Offices nationwide (hereinafter Rogers Memorandum).\textsuperscript{174} The Supreme Court acknowledged in \textit{Rinaldi v. United States}\textsuperscript{175} that the Policy was specifically formulated to address its concerns expressed in \textit{Bartkus} and \textit{Abbate} regarding multiple prosecutions.\textsuperscript{176} \textit{Bartkus}, in particular, admonished both state and federal governments to utilize

---

\textsuperscript{171} Id. Although the \textit{Petite} case involved two federal charges, the government urged the Supreme Court to vacate the second conviction under the general policy “that several offenses arising out of a single transaction should be alleged and tried together,” rather than constitute the basis for multiple prosecutions under the dictates of fairness to the accused and of official and orderly law enforcement. \textit{Id.} at 530. See Jean F. Rydstrom, Annotation, \textit{Effect on Federal Criminal Prosecution or Conviction of Prosecutor’s Noncompliance with Petite Policy Requiring Prior Authorization of Attorney General for Federal Trial Where Accused has been Previously Prosecuted for Same Acts in State Court}, 51 A.L.R. \textit{Fed.} 852 (1981) § 2(a) n.4 (explaining that \textit{Petite v. United States} involved two federal charges and its relation to federal-state successive prosecutions).

\textsuperscript{172} \textit{Petite}, 361 U.S. at 531.

\textsuperscript{173} \textit{Id.} at 533 (Brennan, J., concurring).

\textsuperscript{174} See Podgor, supra note 167, at 178-79 (discussing the history of the \textit{Petite} Policy).


\textsuperscript{176} \textit{Id.} at 27-28. The Court’s per curiam opinion included the following language:

What has come to be known as the \textit{Petite} policy was formulated by the Justice Department in direct response to this Court’s opinions in \textit{Bartkus v. Illinois}, 359 U.S. 121 (1959), and \textit{Abbate v. United States}, 359 U.S. 187 (1959), holding that the Constitution does not deny the State and Federal Governments the power to prosecute for the same act. As these decisions recognize, in our federal system the State and Federal Governments have legitimate, but not necessarily identical, interests in the prosecution of a person for acts made criminal under the laws of both. These cases reflect the concern that if the \textit{Double Jeopardy Clause} were applied when the sovereign with the greater interest is not the first to proceed, the administration of criminal justice may suffer. \textit{Bartkus v. Illinois}, supra, at 137; \textit{Abbate v. United States}, supra, at 195. Yet mindful of the potential for abuse in a rule permitting duplicate prosecutions, the Court noted that “[t]he greatest self-restraint is necessary when that federal system yields results with which a court is in little sympathy.” \textit{Bartkus v. Illinois}, supra, at 138.

In response to the Court’s continuing sensitivity to the fairness implications of the multiple prosecution power, the Justice Department adopted the policy of refusing to bring a federal prosecution following a state prosecution except when necessary to advance compelling interests of federal law enforcement (footnote reference to Rogers Memorandum omitted). The \textit{Petite} policy was designed to limit the exercise of the power to bring successive prosecutions for the same offense to situations comporting with the rationale for the existence of that power.

\textit{Id.} at 28-29.
self-restraint in pursuing multiple prosecutions. The Rogers Memorandum recognized that while federal prosecutors might constitutionally pursue

178. In United States v. Mechanic, the opinion quotes the following parts of the memorandum:

MEMORANDUM TO THE UNITED STATES ATTORNEYS

In two decisions on March 30, 1959, the Supreme Court of the United States reaffirmed the existence of a power to prosecute a defendant under both federal and state law for the same act or acts. That power, which the Court held is inherent in our federal system, has been used sparingly by the Department of Justice in the past. The purpose of this memorandum is to insure that in the future we continue that policy. After a state prosecution there should be no federal trial for the same act or acts unless the reasons are compelling.

In Abbate v. United States and Bartkus v. Illinois (citations omitted) the Supreme Court held that there is no violation of the double jeopardy prohibition or of the due process clause of our federal Constitution where there are prosecutions of the defendant, both in the state and in the federal court, based upon the same act or acts.

This ruling reaffirmed the holding in United States v. Lanza, 260 U.S. 377 (citation omitted) decided by the Supreme Court in 1922. * * *

* * *

But the mere existence of a power, of course, does not mean that it should necessarily be exercised. * * *

* * *

The Court held then that precedent, experience and reason supported the conclusion of separate federal and state offenses.

It is our duty to observe not only the rulings of the Court but the spirit of the rulings as well. In effect, the Court said that although the rule of the Lanza case is sound law, enforcement officers should use care in applying it.

Applied indiscriminately and with bad judgment it, like most rules of law, could cause considerable hardship. Applied wisely it is a rule that is in the public interest. Consequently -- as the Court clearly indicated -- those of us charged with law enforcement responsibilities have a particular duty to act wisely and with self-restraint in this area.

Cooperation between federal and state prosecutive officers is essential if the gears of the federal and state systems are to mesh properly. We should continue to make every effort to cooperate with state and local authorities to the end that the trial occur in the jurisdiction, whether it be state or federal, where the public interest is best served. If this be determined accurately, and is followed by efficient and intelligent cooperation of state and federal law enforcement authorities, then consideration of a second prosecution very seldom should arise.
successive prosecutions in federal court under Supreme Court decisions such as *Bartkus* and *Abbate*, the practice might prove unwise in many instances, and should be informally controlled by the DOJ.\(^{179}\)

Today, the *Petite* Policy has been formalized as Section 9-2.031 of the United States Attorneys’ Manual (hereinafter Manual), a DOJ publication composed of departmental policy statements that can be found online.\(^{180}\) The Policy acts in concert with the more general “Principles of Federal Prosecution” (hereinafter Principles) found in Chapter 9-27 of the Manual, but specifically addresses the issue of subsequent prosecutions.\(^{181}\) The policy statements in the Manual are intended to serve as guidelines that promote consistency in prosecutions, education of neophyte federal prosecutors, and acting as some constraint on prosecutorial discretion.\(^{182}\) The preface of the Principles states that the principles “should promote the reasoned exercise of prosecutorial authority and contribute to the fair, evenhanded administration of the Federal criminal laws.”\(^{183}\)

The *Petite* Policy applies to any situation where a state or federal prosecution has resulted in a decision on the merits, whether conviction or otherwise, after double jeopardy has attached.\(^{184}\) The *Petite* Policy applies only if a trial has actually commenced against a defendant in state court, and does not

---

In such event I doubt that it is wise or practical to attempt to formulate detailed rules to deal with the complex situation which might develop, particularly because a series of related acts are often involved. However, no federal case should be tried when there has already been a state prosecution for substantially the same act or acts without the United States Attorney first submitting a recommendation to the appropriate Assistant Attorney General in the Department. No such recommendation should be approved by the Assistant Attorney General in charge of the Division without having it first brought to my attention.

/s/ William P. Rogers
Attorney General.

United States v. Mechanic, 454 F.2d 849, 855 n.5 (8th Cir. 1971).

179. *Id.* See Podgor, *supra* note 167, at 179 (discussing the relationship between the Court’s decision in *Abbate* and the Rogers Memorandum).


181. *Id.*

182. See Podgor, *supra* note 167, at 169-75 (discussing the purposes and limitations of DOJ internal guidelines).


184. U.S. ATT’YS’ MANUAL, *supra* note 180, § 9-2.031(C) (stating that the *Petite* Policy “applies whenever there has been a prior state or federal prosecution resulting in acquittal, a conviction, including one resulting from a plea agreement, or a dismissal or other termination of the case on the merits after jeopardy has attached”).
apply to cases where state charges have been dismissed.\textsuperscript{185} It also does not apply to circumstances where double jeopardy would normally not attach, such as a hung jury, the declaration of a mistrial, or a reversal of a conviction on appeal.\textsuperscript{186}

Under the Policy, prosecutors presume that any prior trial vindicated federal interests.\textsuperscript{187} However, this presumption can be overcome if there are factors suggesting an un-vindicated federal interest, including:

1. a failure to convict resulting from incompetence, corruption, intimidation or undue influence;
2. court or jury nullification in clear disregard of the evidence or the law;
3. the unavailability of significant evidence not timely discovered or known by the prosecution, or because it was kept from the trier of fact’s consideration because of an erroneous interpretation of the law;
4. the failure in a prior state prosecution to prove an element of a state offense that is not an element of the contemplated federal offense; and
5. the exclusion of charges in a prior federal prosecution out of concern for fairness to other defendants, or for significant resource considerations that favored separate federal prosecutions.\textsuperscript{188}

The presumption may also be overcome where the violation of federal law (1) involves a “compelling federal interest,” (2) the offense involves “egregious conduct,” or (3) where any prior prosecution is regarded as “manifestly inadequate” in light of the federal interest at issue.\textsuperscript{189}

The underlying rationale for subsequent federal prosecutions under the \textit{Petite} Policy is expressed in the Manual as:

\textsuperscript{185} See United States v. Gomez, 776 F.2d 542, 550 (5th Cir. 1985) (in response to the defendants’ argument that a federal prosecution after state charges were dismissed violated the \textit{Petite} Policy, the court replied “the government assures us that this policy is not implicated in this case, and this assurance suffices for our purposes”). \textit{See also} GEO. L. J. ANN. REV. CRIM. PROC., supra note 74, at 452 n.1465 (“The \textit{Petite} policy applies only if the defendant is actually brought to trial in state court.”) (citing United States v. Gomez, 776 F.2d 542 (5th Cir. 1985) as an example).


\textsuperscript{187} U.S. ATT’YS’ Manual, \textit{supra} note 180, § 9-2.031(D).

\textsuperscript{188} Id.

\textsuperscript{189} Id. Whether a substantial federal interest exists is determined under the general principles of federal prosecution found in the U.S. Attorneys’ Manual, which lists the following non-exclusive considerations:

1. federal law enforcement priorities;
2. the nature and seriousness of the offense;
3. the deterrent effect of prosecution;
4. the [defendant]’s culpability in connection with the offense;
5. the [defendant]’s history with respect to criminal activity;
6. the [defendant]’s willingness to cooperate in the investigation or prosecution of others; and
7. the probable sentence or other consequences if the [defendant] is convicted.

(1) vindicating “substantial federal interests” with “appropriate federal prosecutions;”
(2) protecting accuseds from multiple prosecutions for substantially the same offense;
(3) promoting effective use of DOJ resources; and
(4) encouraging state and federal prosecutorial cooperation.\(^{190}\)

As the Court observed in \textit{Rinaldi}, the Policy seeks to limit subsequent federal prosecutions to situations in which a significant federal interest truly exists.\(^{191}\)

Even if the Policy criteria are satisfied, the \textit{Petite} Policy requires one additional hurdle before a subsequent federal prosecution may be initiated: approval by an appropriate higher level U.S. Assistant Attorney General,\(^{192}\) usually the Assistant Attorney General for the Criminal Division,\(^{193}\) unless the state charge is only a minor part of the federal charges under consideration, as in a RICO prosecution.\(^{194}\) In a 1996 article, two attorneys with the DOJ’s Office of Policy Development claimed that statistics indicated that the DOJ typically pursued fewer than 150 dual prosecutions each year, representing only a tiny fraction of the total number of cases with overlapping state and federal jurisdiction, and an even smaller percentage of the approximately 65,000 annual federal prosecutions.\(^{195}\)

---

\(^{190}\) U.S. ATT’YS’ MANUAL, supra note 180, § 9-2.031(A) (emphasis added).


\(^{192}\) U.S. ATT’YS’ MANUAL, supra note 180, § 9-2.031(E).

\(^{193}\) \textit{See Litman & Greenberg, supra note 123, at 77 n.15} (Litman and Greenberg note that the Civil Rights, Antitrust, and Environmental Divisions may have cases involving the \textit{Petite} Policy that would require the approval of a different assistant attorney general, but the number of such cases would be much smaller in volume). A recent article by an attorney formerly with the DOJ Witness Immunity Unit (WIU) of the Office of Enforcement Operations in the Criminal Division reveals that the process has changed somewhat, with the WIU reviewing initial requests for \textit{Petite} Policy waiver and sending its recommendation to the Assistant Attorney General, who makes the ultimate decision on waiver. Kathleen N. Coleman, \textit{The Witness Immunity Office of Enforcement Operations}, 55 U.S. ATT’YS BULL (EOUSA) 10, 12-13 (January 2007). An email exchange between the author and the current attorney in charge of handling waiver requests for the WIU disclosed that the tax division may have different criteria. In fact, Directive No. 115 of the Tax Division of the DOJ expressly states that \textit{Petite} Policy waiver decisions are for the Assistant Attorney General for the Tax Division and may not be delegated. U.S. DEP’T OF JUST. CRIM. TAX MANUAL § 3.00, Directive No. 115 1(d) (2008), available at \url{http://www.justice.gov/tax/readingroom/2008ctm/CTM%20Chapter%203.htm}. A reference to \textit{Petite} policy concerns is also contained on page III-122 of the Fourth Edition of the \textit{Antitrust Division Manual} (available at \url{www.justice.gov/atr/public/divisionmanual/atrdivman.pdf}), although the decision process is not elaborated.

\(^{194}\) U.S. ATT’YS’ MANUAL, supra note 180, § 9-2.031(B) (making specific reference to RICO cases, in which acts previously prosecuted under state law are exempted from the policy to the extent the acts constitute “only a minor part of the contemplated federal charges” and “as long as those acts or offenses do not represent substantially the whole of the contemplated federal charge, and, in a RICO prosecution, as long as there are a sufficient number of predicate offenses to sustain the RICO charge if the previously prosecuted offenses were excluded”).

\(^{195}\) \textit{Litman & Greenberg, supra note 123, at 77.}
Although considerations of state prosecution are addressed elsewhere in the Manual, Section 9-2.031 requires federal prosecutors to consult with state prosecutors to determine the most appropriate single forum in which to bring charges against a defendant when the two jurisdictions overlap. One of the major areas of concern involving “substantial federal interests” is civil rights violations. During the Carter Administration, Attorney General Griffin Bell required federal prosecutors to more strictly adhere to the Petite Policy before initiating civil rights cases. Nonetheless, civil rights violations are still regarded as one of the more appropriate usages of the dual sovereignty doctrine because of some states’ inadequate past treatment of such abuses.

B. Problems with the Petite Policy—

Because the Petite Policy is a creation of the DOJ rather than constitutionally or statutorily mandated, the Policy vests no rights in an accused subject to multiple prosecutions, other than those which the federal government elects to provide. In essence, the Policy is no more than an instance of prosecutorial discretion that incorporates written guidelines. The preface of the Principles states as much when it clarifies that the principles “have been cast in general terms with a view to providing guidance rather than to mandating results. The intent is to assure regularity without regimentation and to prevent unwarranted disparity without sacrificing necessary flexibility.” When prosecutors ignore the DOJ’s Petite Policy, courts have routinely treated the guidelines as strictly internal and not enforceable by outside parties, including

197. U.S. ATT’YS’ MANUAL, supra note 180, § 9-2.031(A) provides:
In order to insure the most efficient use of law enforcement resources, whenever a matter involves overlapping federal and state jurisdiction, federal prosecutors should, as soon as possible, consult with their state counterparts to determine the most appropriate single forum in which to proceed to satisfy the substantial federal and state interests involved, and, if possible, to resolve all criminal liability for the acts in question.
198. See e.g., Herman, supra note 83, at 632-39 (discussing the ACLU’s struggle in deciding whether there should be a double jeopardy exception for civil rights violations).
200. See e.g., Little, supra note 159, at 1077-81 (citing the civil rights prosecutions as an example of a time when federal intervention was necessary despite state objection).
201. Hagburg, supra note 2, at 589 (noting that the Petite Policy is not “constitutionally mandated” resulting in criminal defendants having no right to protections under the policy “unless the government chooses to invoke it”). See, e.g., United States v. Jackson, 327 F.3d 273, 294-95 (4th Cir. 2003); United States v. Kriens, 270 F.3d 597, 603 (8th Cir. 2001) (holding that the Petite Policy does not create any personal rights for criminal defendants).
criminal defendants. In fact, the Principles expressly provide that none of the expressions of policy are intended to be relied upon to create a right or benefit, whether substantive or procedural, to any party in litigation with the United States. The Petite Policy itself contains a similar recitation.

A 2004 study of federal criminal cases performed by Ellen S. Podgor reveals that, unless other constitutional violations are included or the federal government seeks reversal, a defendant’s argument that a prosecution violated the Petite Policy or any other DOJ policy falls on deaf ears. Not surprisingly, the Petite Policy has been described as little more than a “‘wait and see’ approach [that] is significantly flawed” due to its susceptibility to manipulation and political pressure. Several jurists have argued that the right to policy adherence under DOJ policies such as the Petite Policy should be enforceable as a matter of due process.

203. See, e.g., Jackson, 327 F.3d at 294-95; Kriens, 270 F.3d at 603. See also cases cited in Podgor, supra note 167, nn. 73-76. (refusing to provide a judicial remedy for a violation of the Petite policy).

204. U.S. ATT’YS’ MANUAL, supra note 180, § 9-27.150.

205. The U.S. ATT’YS’ MANUAL provides that the Petite Policy:

[H]as been promulgated solely for the purpose of internal Department of Justice guidance. It is not intended, does not, and may not be relied upon to create any rights, substantive or procedural, that are enforceable at law by any party in any matter, civil or criminal, nor does it place any limitations on otherwise lawful litigative prerogatives of the Department of Justice. U.S. ATT’YS’ MANUAL, supra note 180, § 9-2.031(F).

206. Podgor, supra note 167, at 177-81.

207. Merkl, supra note 10, at 200.

208. See Judge Oliver Seth’s dissent in United States v. Thompson, in which Thompson’s federal prosecution subsequent to state prosecution was not approved by an appropriate higher level attorney general until after Thompson’s prosecution and sentencing because the Assistant U.S. Attorney prosecuting Thompson was not aware of the requirement. 579 F.2d 1184, 1188-89 (10th Cir. 1978). After surveying numerous other agency regulations that are enforced by the courts, Judge Seth dissented:

From the variety of situations in which the courts have held that agencies must follow their statements of policy and directives, it is apparent that if the statements are applicable, are made known, and are of guidance for the execution of the agency policies, it makes no difference what they are called nor how they are adopted. The public has a right to rely on them as against some inconsistent case by case subjective determination by a public official. The cases further demonstrate that it makes no difference whether the agency has express statutory authority to adopt the policy. The agency has to run its business, it has to function. This is done by regulations and policy statements, and it makes no difference whether Congress has expressly directed them or not if they are within the general objectives of the agency, and especially if they represent some self-restraint on its authority or discretion as in the case before us.

The Petite policy against dual prosecutions contains an exception to permit dual prosecution when necessary to advance ‘compelling interests of law enforcement.’
The challenges faced by the government when attempting to avail itself of the Petite Policy after a federal prosecution has commenced are a somewhat surprising problem that has arisen in recent years. *Rinaldi* indicates that defendants should receive the benefits of the Policy when the government urges it, and lower courts have made statements advocating the rights of the federal government under the Policy.

Thus I would hold that the postconviction attempt to authorize the prosecution of defendant Thompson in the federal court came too late to be effective as he had been convicted and sentenced some nine months before. United States v. Thompson, 579 F.2d 1184, 1191-92 (10th Cir. 1978) (Seth, J., dissenting).

Sentiments against the Petite Policy were also expressed by Judge Ruggero J. Aldisert in his dissent in *United States v. Wilson*, in which Wilson was prosecuted for a drug offense in federal court after a motion to suppress was granted in state court. 413 F.3d 382, 385 (3d Cir. 2005). In response to the government’s position at oral argument that it could refile in federal court when the government felt an injustice was done in state court, Judge Aldisert referred to the government’s conduct “as a blatant exercise of judge shopping” and further continued:

> I believe this policy generates serious problems. It increases the caseload in federal courts, runs counter to modern concepts of federalism, denigrates the quality of the state-court system, trial and appellate, demeans the professionalism of state-court judges who have more experience, indeed much more experience, in deciding federal constitutional questions in criminal proceedings than federal judges . . .
>
> The very admission in open court that the federal government will initiate a new prosecution in cases where state courts suppress evidence has a pernicious effect on the rights of state-court defendants seeking to vindicate Fourth Amendment rights. The federal government’s message to state judges is clear: ‘Do not suppress evidence. If you do, we’ll institute a new federal prosecution on the same facts even though the investigation and arrest were made by state authorities and the state conducted the prosecution.’ This policy allows the United States, in effect, to use federal courts to review any state judge’s federal constitutionally-based decision on a motion to dismiss.
>
> To me, this is appalling.

United States v. Wilson, 413 F.3d 382, 391-92 (3d Cir. 2005) (Aldisert, J., dissenting). The majority was sympathetic to Judge Aldisert’s concerns and expressed its own concerning the Petite Policy and dual sovereignty, but felt it was not in a position to overrule Supreme Court precedent. *Wilson*, 413 F.3d at 388-90.

209. In *Rinaldi*, the Court noted:

> The Petite policy was designed to limit the exercise of the power to bring successive prosecutions for the same offense to situations comporting with the rationale for the existence of that power. Although not constitutionally mandated, this Executive policy serves to protect interests which, but for the ‘dual sovereignty’ principle inherent in our federal system, would be embraced by the Double Jeopardy Clause. In light of the parallel purposes of the Government’s Petite policy and the fundamental constitutional guarantee against double jeopardy, the federal courts should be receptive, not circumspect, when the Government seeks leave to implement that policy.


210. See, e.g., United States v. Booth, 673 F.2d 27, 30 (1st Cir. 1982), cert. denied, 456 U.S. 978 (1982) (“The Petite policy and cases construing it stand only for the proposition that the government’s motion to dismiss should be granted when the government decides it is conducting separate prosecution for the same offense.”); United States v. Snell, 592 F.2d 1083, 1087-88 (9th Cir. 1979), cert. denied, 442 U.S. 944 (1979) (noting that the Supreme Court remands cases based on the Petite Policy only at the request of the DOJ).
Despite this, some Supreme Court justices and other courts have expressed reluctance to vacate or dismiss prosecutions already concluded, or well under way, solely on the government’s motion, indicating that the government’s right to secure a dismissal under the Petite Policy may be more limited than one might think.211

In Watts v. United States, Chief Justice Warren Burger dissented, joined by Justices Byron White and William Rehnquist, in a memorandum opinion that granted certiorari and vacated a conviction at the government’s request.212 Chief Justice Burger questioned the Court’s complicity in reversing a conviction under the Policy following a substantial commitment of prosecutorial and court resources to secure a conviction.213

Lower court opinions also show this judicial reluctance. In United States v. Cockrell, a federal judge in Texas denied a prosecutor’s request, made on the eve of trial, for dismissal under the Policy and issued what can only be described as a scathing rebuke of the prosecution’s conduct.214 The court chided the prosecution for its lack of candor in informing the court of the circumstances of the case, including a lack of certainty concerning the overlap of state and federal charges,215 before making specific findings as to why dismissal under the Policy was inappropriate.216 The court concluded that “[t]he contention of the government that the Petite Policy applies is frivolous, irrational, and devoid of any basis in fact.”217 In denying the motion to dismiss and ordering the case to trial, the court took the position that, even in the case of a Petite Policy request, it is incumbent upon the government to demonstrate that dismissal is in the public interest, and that a court has an obligation to deny a motion to dismiss “when the prosecutor’s actions clearly implicate a betrayal of the public interest.”218

Subsequently in Cockrell, after voir dire questioning had commenced, the federal government petitioned the Fifth Circuit for a writ of mandamus,

212. Id.
213. Id. at 1034-36. In his dissent, the Chief Justice noted that failure to adhere to DOJ policies is a matter for the Department and not for the courts, particularly after a defendant has been convicted and there is no error in the trial. Id. at 1036-37. The Chief Justice wrote:

The federal courts have no role in prosecutorial decisions, but, once the judicial power has been invoked, it is decidedly the role of federal courts to interpret the decisions of this Court and to assess the validity of judgments duly entered.

Surely it is not our function either to approve or disapprove internal prosecutorial policies and even less so their implementation.

Id. (Burger, C.J., dissenting).
215. Id. at 766-75.
216. Id. at 774-75.
217. Id. at 775.
218. Id. at 768-69 (quoting United States v. Hamm, 659 F.2d 624, 629 (5th Cir. 1981)).
requesting it to have the trial court dismiss the indictment.\textsuperscript{219} The Fifth Circuit responded in the government’s favor, giving the trial court three days to dismiss the indictment.\textsuperscript{220} After the trial court’s refusal to do so, the Fifth Circuit Court granted the requested writ of mandamus.\textsuperscript{221} However, the Fifth Circuit’s failure to elaborate on its reasoning, other than a generic recitation that the trial court abused its discretion,\textsuperscript{222} leaves unclear whether the appellate court regarded the entry of an order of dismissal, when requested by the government under the Policy, as a ministerial act with which the trial court could not refuse to comply.

A final problem with the Policy is its continued viability as an internal DOJ policy. In May 2008, the Criminal Division of the DOJ circulated a memorandum soliciting input for possible changes to the Policy.\textsuperscript{223} The memo notes that “the policy was adopted . . . during a time when there was little, if any, official coordination on criminal matters between U.S. Attorney’s Offices and their state and local prosecuting counterparts” and that increased coordination “has triggered Petite issues that may not have been contemplated when the policy was adopted.”\textsuperscript{224} The memo cites concerns expressed by DOJ lawyers “about the need for Petite waivers in cases where there has been close federal/state coordination of prosecutive responses” and observes “it has been suggested that the policy is over-inclusive and that prior convictions for minor, unrelated state charges should not trigger the waiver requirement.”\textsuperscript{225} A January 29, 2010, exchange of emails between the author and the designated contact at the Office of Enforcement Operations revealed that, while the solicitation of input generated many suggestions, the DOJ felt there was no pressing need to make substantive changes to the Policy.\textsuperscript{226}

\textsuperscript{219} Cockrell, 353 F. Supp. 2d at 777-78. \\
\textsuperscript{220} Id. at 777. \\
\textsuperscript{221} Id. at 779. The amended version of the Memorandum Opinion and Order sent by the district court to the Fifth Circuit panel provides its own version of what transpired, including the trial court’s refusal to dismiss the case because the court felt it would violate its responsibility under FED. R. CRIM. P. 48(a), and a statement that “the court cannot overlook the possibility that the motion to dismiss . . . was a ruse designed to help [the assistant U.S. attorney handling the case] and the United States Attorneys Office clean up his docket so that he could be moved to another place.” Id. at 778 n. 1 The Fifth Circuit’s initial order found that the district judge had “exceeded the bounds of his discretion under Rule 48(a) in denying the government’s motion.” Id. at 777 (citing In re United States, No. 05-10023, slip op. at 1-2 (5th Cir. Jan. 10, 2005)). \\
\textsuperscript{222} See id. at 779 (stating that the order granting the writ of mandamus did not provide any explanation). \\
\textsuperscript{223} Memorandum from Maureen H. Killion, Director, U.S. Department of Justice, Criminal Division, Office of Enforcement Operations to Kenneth Melson, Director, U.S. Department of Justice, Executive Office of the United States Attorney (May 22, 2008). \\
\textsuperscript{224} Id. \\
\textsuperscript{225} Id. \\
\textsuperscript{226} Email from Thomas C. Taylor, Associate Director for Policy, U.S. Department of Justice, Criminal Division, Office of Enforcement Operations, to author (January 28, 2009) (on file with author) (indicating that while the survey had generated suggestions, it demonstrated no pressing need for change, but did result in the creation of a standardized electronic application for DOJ waiver of the policy that is available from the FOIA reading room).
C. Determining Priority in Prosecution—

Even where the dual sovereignty doctrine permits a subsequent prosecution, many states limit or prohibit subsequent state prosecutions when a defendant has previously been prosecuted by the federal government or, in some instances, any government.227 It is a fundamental tenet of constitutional law that, while states cannot fall below the baseline of constitutional protections applied to the states under the Fourteenth Amendment, states are free to provide greater protections as a matter of state law.228 This maxim is commonly referred to as the “new judicial federalism.”229 For states that limit or bar subsequent prosecutions, determining which sovereign gets the first opportunity to prosecute a defendant in cases involving overlapping state and federal jurisdiction may often mean the difference between a state being able to vindicate its interests under the state’s criminal law and no state prosecution of any kind.230 Even in states that do not limit the dual sovereignty doctrine, primacy of prosecution may affect state interests in ways that impact the availability and usage of state resources.231

The Principles set forth the DOJ criteria for initiating or declining prosecution when the possibility for prosecution in another jurisdiction exists.232 The section requires federal prosecutors to weigh all relevant considerations, including three which are expressly set forth:

1. the strength of the other jurisdiction’s interest in prosecution;
2. the other jurisdiction’s ability and willingness to prosecute effectively; and
3. the probable sentence or other consequences if the person is convicted in the other jurisdiction.233

227. See infra Part IV.
228. See KANOVITZ & KANOVITZ, supra note 12, at § 1.13.
229. See Staci L. Beavers & Jeffrey S. Walz, Modeling Judicial Federalism: Predictors of Defendants’ Rights Under State Constitutions, 1969-1989, 28 PUBlius 43, 43-44 (1998) (examining the increasing role state courts play in maintaining or expanding constitutional rights with the decreasing amount of constitutional protections afforded by the Burger and Rehnquist Courts, and noting that “[t]his development of state civil liberties law has been often referred to as a new era of ‘judicial federalism’”).
230. See Bartkus v. Illinois 359 U.S. 121, 136-37 (U.S. 1959) (discussing the implications of holding that due process prohibits a successive state prosecution after a federal prosecution, mainly that it would deprive the states of their “historic right and obligation to maintain peace and order”).
231. See id. Possible collateral effects could entail the time in which a case goes to trial, availability of witnesses, stale memories or evidence, etc.
233. Id. A search of all federal and state cases in the WestLaw ALLFEDS and ALLSTATES data bases revealed that no court has referred to this provision by title or section number. The section is entitled “Initiating and Declining Charges—Prosecution in Another Jurisdiction.” Five searches were run in the federal and state case data bases constructed as follows:
Search 1—9-27.240
Search 2—“Initiating and Declining Charges—Prosecution in Another Jurisdiction”
Search 3—“Initiating and Declining Charges” and “Prosecution in Another Jurisdiction”
As is the case with all policies in the Manual, the Manual creates no vested rights to the accused’s benefit for violation of the Policy regarding declining prosecution when the criteria specified are satisfied by another jurisdiction.\textsuperscript{234} For the same reasons, one would surmise that states also have no rights to initial prosecution if the Policy is violated, even if all of the requisite conditions are satisfied.

\textbf{IV. \textsc{State Responses to the Issue of Subsequent Prosecutions}}

Many states limit subsequent prosecutions under the dual sovereignty doctrine. As previously noted, where the U.S. Supreme Court has made provisions of the Bill of Rights applicable to the states through the doctrine of selective incorporation, states are bound to provide no less protection than that afforded by the U.S. Constitution. However, the states are free to provide greater protections, including double jeopardy protections.\textsuperscript{235} Conversely, states are free not to limit subsequent prosecutions as long as they comply with Fifth Amendment double jeopardy protections and the decisions by the U.S. Supreme Court interpreting those protections.\textsuperscript{236}

States address the issues of double jeopardy and dual sovereignty in a variety of ways.\textsuperscript{237} As Table One indicates, most state constitutions expressly provide general protections against double jeopardy.\textsuperscript{238} While some state constitutions are explicit in what does and does not constitute double jeopardy,\textsuperscript{239} most approximate the guarantees contained in the Fifth Amendment.\textsuperscript{240} The Declaration of Rights in the Constitution of Maryland merely adopts the rights guaranteed by the U.S. Constitution.\textsuperscript{241} Montana actually incorporates protections against dual sovereignty into its constitution.\textsuperscript{242}

\begin{verbatim}
Search 4—initiating /3 declining /3 charges /3 prosecution /5 jurisdiction
Search 5—initiating /3 declining /3 charges and prosecution /5 jurisdiction
All five searches indicated there were no cases fitting the search criteria.
234. \textit{See supra} Part III.B.
236. \textit{See} Benton \textit{v. Maryland}, 395 U.S. 784, 793-96 (U.S. 1969) (declaring that the double jeopardy protections in the Fifth Amendment are incorporated to the states so that the same constitutional standards apply to both state and the federal government).
237. For a general discussion of state prosecutions following federal prosecutions, \textit{see} Carolyn Kelly MacWilliam, Annotation, \textit{Conviction or Acquittal in Federal Court as Bar to Prosecution in State Court for State Offense Based on Same Facts—Modern View}, 97 A.L.R. 5th 201 (2002).
238. \textit{See infra}, Table One.
239. \textit{See, e.g.}, GA. \textit{CONST.} art. I, § I, ¶ XVIII; MO. \textit{CONST.} art. I, § 19 (excluding matters such as mistrials, new trials and reversals on appeal from double jeopardy protection).
240. Most state constitutions employ some variant of a barebones guarantee against being twice placed in jeopardy. \textit{See, e.g.}, DEL. \textit{CONST.} art. I, § 8 (“no person shall be for the same offense twice put in jeopardy of life or limb.”); ILL. \textit{CONST.} art. I, § 10 (“[n]o person shall . . . . be twice put in jeopardy for the same offense.”).
241. MD. \textit{CONST.} art. 2.
242. MONT. \textit{CONST.} art. II, § 25 (providing that “[n]o person . . . . shall be again put in jeopardy for the same offense \textit{previously tried in any jurisdiction.”}) (emphasis supplied).
\end{verbatim}
The constitutions of some states, such as North Carolina, provide no explicit double jeopardy protection. In response, North Carolina’s Supreme Court, along with many other state supreme courts, has held that double jeopardy protections exist under the state’s due process clause, other constitutional language, or under the state’s common law. Regardless of constitutional, statutory, or court opinion language, all states are required to afford protections against double jeopardy since the U.S. Supreme Court’s 1969 holding in *Benton v. Maryland*, which incorporated and applied the Fifth Amendment’s double jeopardy protections to the states through the Due Process Clause of the Fourteenth Amendment.

Table One—State Constitution Double Jeopardy Provisions—

<table>
<thead>
<tr>
<th>State</th>
<th>Constitution</th>
<th>Manner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>Art. 1, Sec. 9</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Alaska</td>
<td>Art. I, Sec. 9</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Arkansas</td>
<td>Art. 2, Sec. 8</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Arizona</td>
<td>Art. II, Sec. 10</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>California</td>
<td>Art. I, Sec. 15</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Colorado</td>
<td>Art. II, Sec. 18</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Connecticut</td>
<td>Art. I, Sec. 8</td>
<td>No Specific Double Jeopardy Provisions,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Encompassed in Due Process and by Statute</td>
</tr>
<tr>
<td>Delaware</td>
<td>Art. I, Sec. 8</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Florida</td>
<td>Art I, Sec. 9</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Georgia</td>
<td>Art. I, Sec. 1</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Hawaii</td>
<td>Art. I, Sec. 10</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Idaho</td>
<td>Art. I, Sec. 13</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Illinois</td>
<td>Art. I, Sec. 10</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Indiana</td>
<td>Art. I, Sec. 14</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Iowa</td>
<td>Art. I, Sec. 12</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Kansas</td>
<td>BOR, Sec. 10</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Kentucky</td>
<td>BOR, Sec. 13</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Louisiana</td>
<td>Art. I, Sec. 13</td>
<td>No Specific Double Jeopardy Provisions,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

243. See e.g. N.C. CONST. art. I §§ 22-23 (providing for modes of prosecution and rights of accused, but not mentioning protection from double jeopardy).

244. See, e.g., State v. Nugent, 89 S.E.2d 781, 783 (N.C. 1955) (discussing a defendant’s right to sufficient language in a charging instrument, the court stated in dicta that the right to be informed of the charges against a defendant was “a substantial redeclaration of the common law rule requiring the charge against the defendant to be set out in the warrant or indictment with such exactness that the defendant can have a fair and reasonable opportunity to prepare his defense, can avail himself of his conviction or acquittal as a bar to subsequent prosecution for the same offense, and can enable the court, on conviction, to pronounce sentence according to law”) (emphasis added).

The states in Table Two limit, or have in the past limited, the dual sovereignty doctrine by court decision. The Supreme Courts of Michigan and

246. Commonwealth v. Burke, 172 N.E.2d 605, 606-07 (Mass. 1961) (recognizing double jeopardy protections in the Fifth Amendment of the U.S. Constitution and noting that there is no similar provision in the state constitution, but that it is recognized in state statute).

247. State v. O’Brien, 170 A. 98, 103-04 (Vt. 1934) (finding that double jeopardy protections are established in common law).


<table>
<thead>
<tr>
<th>State</th>
<th>Section/Article/Para</th>
<th>Protection Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maine</td>
<td>Art. I, Sec. 8</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Maryland</td>
<td>DOR, Art. 2</td>
<td>Adoption of U.S. Constitution</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>None</td>
<td>Statutory and Common Law Protections,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commonwealth v. Burke (1961)</td>
</tr>
<tr>
<td>Michigan</td>
<td>Art. I, Sec. 15</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Minnesota</td>
<td>Art. I, Sec. 7</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Mississippi</td>
<td>Art. I, Sec. 22</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Missouri</td>
<td>Art. I, Sec. 19</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Montana</td>
<td>Art. II, Sec. 25</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Nebraska</td>
<td>Art. I, Sec. 12</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Nevada</td>
<td>Art. I, Sec. 8</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>Pt. I, Art. 16</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>New Jersey</td>
<td>Art. I, Para. 11</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>New Mexico</td>
<td>Art. II, Sec. 15</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>New York</td>
<td>Art. I, Sec. 6</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>North Carolina</td>
<td>Art. I, Secs. 22, 23</td>
<td>No Specific Double Jeopardy Provisions,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other Sections Afford Protection</td>
</tr>
<tr>
<td>North Dakota</td>
<td>Art. I, Sec. 12</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Ohio</td>
<td>Art. I, Sec. 10</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>Art. II, Sec. 21</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Oregon</td>
<td>Art. I, Sec. 12</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>Art. I, Sec. 10</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>Art. I, Sec. 7</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>South Carolina</td>
<td>Art. I, Sec. 12</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>South Dakota</td>
<td>Art. VI, Sec. 9</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Tennessee</td>
<td>Art. I, Sec. 10</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Texas</td>
<td>Art. I, Sec. 14</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Utah</td>
<td>Art. I, Sec. 12</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Vermont</td>
<td>None</td>
<td>Statutory and Common Law Protections,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State v. O’Brien (1934)</td>
</tr>
<tr>
<td>Virginia</td>
<td>Art. I, Sec. 8</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Washington</td>
<td>Art. I, Sec. 9</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>West Virginia</td>
<td>Art. III, Sec. 5</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>Art. I, Sec. 8</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
<tr>
<td>Wyoming</td>
<td>Art. I, Sec. 11</td>
<td>Standard Double Jeopardy Protections</td>
</tr>
</tbody>
</table>
New Hampshire\textsuperscript{249} have found unstated protections against the dual sovereignty doctrine under their state constitutions. Before the decisions by the Michigan and New Hampshire Supreme Courts, the Pennsylvania Supreme Court decided in \textit{Commonwealth v. Mills} that the U.S. Supreme Court had erroneously interpreted the double jeopardy protections afforded by the Fifth Amendment in \textit{Bartkus}.\textsuperscript{250} The court’s opinion chastised the U.S. Supreme Court for being overly deferential to the prosecutorial rights of the federal and state governments at the expense of individuals’ rights protected under the Bill of Rights.\textsuperscript{251} The Pennsylvania Supreme Court, without any reference to state law, declared that Pennsylvanians did, in fact, possess protections against dual sovereignty unless the interests of the two sovereigns were “substantially different.”\textsuperscript{252} As Table

\begin{table}
\centering
\begin{tabular}{|l|l|}
\hline
| Michigan | New Hampshire | Pennsylvania |
\hline
An acquittal | Acquittal | Acquittal |
\hline
A second prosecution will be permitted unless the interests of the two sovereigns are substantially different.
\hline
\end{tabular}
\caption{Comparison of State Laws on Dual Sovereignty}
\end{table}

\textsuperscript{249}. In \textit{State v. Hogg}, 385 A.2d 844 (N.H. 1978), the New Hampshire Supreme Court was much more emphatic about the wrongheadedness of the dual sovereignty doctrine, holding “[i]t is fundamentally and morally wrong to try a man for a crime of which he has already been tried and found not guilty” even if an initial acquittal is ill-founded. \textit{Id.} at 845. The court also reasoned that the New Hampshire Constitution existed before New Hampshire’s ratification of the United States Constitution and, therefore, the drafters could not possibly have contemplated the doctrine of dual sovereignty in granting double jeopardy protections. \textit{Id.} at 845-46.

\textsuperscript{250}. \textit{Commonwealth v. Mills}, 286 A.2d 638, 640-41 (Pa. 1971) (arguing that the majority in \textit{Bartkus} failed to recognize that the interests of both sovereigns might be the same).

\textsuperscript{251}. \textit{Id.} at 640-42. The Pennsylvania Supreme Court railed against the majority in \textit{Bartkus}, arguing:

\begin{quote}
More important, [the \textit{Bartkus} majority opinion] failed to really examine the interest of the individual.

\ldots

It is wrong to retry a man for a crime of which he previously has been found innocent, wrong to harass him with vexatious prosecution, and wrong to punish him twice for the same offense.

\ldots

The striking feature of the aforementioned rules and policies is that the focus is always on the individual, on a person’s basic and fundamental rights. This feature is the common thread that runs across all of the provisions of the [United States] Bill of Rights, and we believe this is the element the Supreme Court failed to adequately consider in \textit{Bartkus}.

Mr. Justice Frankfurter, not once in his majority opinion in \textit{Bartkus} refers to the interest of the individual, he consistently focused on the interests of the Federal and State governments, and, for the purpose of double jeopardy, transformed one act into two. \ldots

We are talking about the two governments protecting their interests, when we really should be talking about the individual, since by focusing on the individual we see that it matters little where he is confined—in a federal or state prison—the fact is that his liberty is taken away twice for the same offense.

\textit{Id.} at 641.
\end{quote}

\textsuperscript{252}. \textit{Id.} at 642. Having only discussed the United States Supreme Court opinion in \textit{Bartkus}, the Pennsylvania Supreme Court, without discussing any Pennsylvania authority for the proposition, held “[h]enceforth in Pennsylvania, a second prosecution and imposition of punishment for the same offense will not be permitted unless it appears from the record that the interests of the Commonwealth of Pennsylvania and the jurisdiction which initially prosecuted and imposed punishment are substantially different.” (footnote omitted).
Three indicates, the Pennsylvania legislature subsequently adopted statutory limitations on the dual sovereignty doctrine as well. Following the U.S. Supreme Court’s decision in Benton v. Maryland, an Ohio intermediate appellate court decreed that dual sovereignty protections exist in Ohio. The decision was subsequently reversed when the Ohio Supreme Court determined the lower court had misinterpreted Benton to require such protection. Ohio presently provides no dual sovereignty limitations.

<table>
<thead>
<tr>
<th>State</th>
<th>Opinion</th>
<th>Basis for Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michigan</td>
<td>People v. Cooper</td>
<td>State Constitution</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>State v. Hogg</td>
<td>State Constitution</td>
</tr>
<tr>
<td>Ohio</td>
<td>State v. Fletcher</td>
<td>Erroneous Interpretation of Benton v. Maryland (1969) by trial and intermediate</td>
</tr>
<tr>
<td></td>
<td>(1970)</td>
<td>court of appeals (reversed)</td>
</tr>
</tbody>
</table>

Both the Pennsylvania and Michigan decisions limited protections against dual sovereignty to situations in which the state’s interest was vindicated by the prior prosecution.

In 1996, three judges of the Michigan Supreme Court voted to overrule Cooper as wrongly decided because it failed to adhere to a Michigan tenet of statutory construction that Michigan law will provide no greater protection than federal law absent a compelling reason. Because three judges also voted against overruling Cooper, the case is still law in Michigan, however weakened it may be. In 1998, the Michigan Supreme Court refused to address the continuing viability of Cooper in an en banc opinion in People v. Childers on

253. 18 PA. CONS. STAT. ANN. § 111 (West 2010).
256. See infra Table Four.
257. See supra notes 245-47 and 251-52.
260. See People v. Childers, 587 N.W.2d 17 n.3 (1998) (“In People v. Mezy (citation omitted), a majority of this Court held that there was no need to address the merits of Cooper, while three justices voted to overrule Cooper.”).
grounds that it was unnecessary to review Childers’ state conviction because the military prosecution of Childers failed to vindicate the state’s interests.261

Like Pennsylvania, many states adopted statutory limitations, and sometimes even bars, to the application of the dual sovereignty doctrine.262 Some bar subsequent state prosecutions for the same “offense,”263 while others prohibit prosecutions for the same “act,” “behavior,” or “conduct,”264 often with the state’s courts later equating these terms with “same offense.”265

A common requirement for the prohibition against successive prosecution in states with a statutory limitation is the standard double jeopardy jurisprudence requirement that double jeopardy attach in the prior prosecution either by conviction, acquittal, or improper termination.266 Other common statutory exceptions to state protections from the application of the dual sovereignty doctrine include the following:

261. Id. at 219-23.
262. See infra Table Three.
263. See e.g., 720 ILL. COMP. STAT. ANN. 5/3-4(c) (LexisNexis 2010); KY. REV. STAT. ANN. § 505.050 (West 2010); MISS. CODE. ANN. § 99-11-27 (2010); UTAH CODE ANN. § 76-1-404 (2010). See also GA. CODE ANN. § 16-1-8(c) (2010)(barring subsequent prosecution for the same “crime”).
266. See ARK. CODE ANN. § 5-1-114 (2010) (conviction, acquittal, or improper termination); CAL. PENAL CODE § 793 (West 2010) (conviction or acquittal); COLO. REV. STAT. ANN. § 18-1-303(1)(a) (West 2010) (conviction, acquittal, or improper termination); DEL. CODE ANN. tit. 11, § 209(1) (2010) (conviction, acquittal, or improper termination); HAW. REV. STAT. ANN. § 701-112 (LexisNexis 2010) (conviction, acquittal, or improper termination); IDAHO CODE ANN. § 19-315 (2010) (conviction or acquittal); 720 ILL. COMP. STAT. ANN. 5/3-4 (LexisNexis 2010) (conviction, acquittal, or improper termination) (the 2010 amendments made one change not affecting the dual sovereignty protections and other changes described as “non-substantive); IND. CODE ANN. § 35-41-4-5 (LexisNexis 2010) (conviction, acquittal, or improper termination); KAN. STAT. ANN. § 21-3108(3) (2010) (acquittal, conviction, or improper termination); KY. REV. STAT. ANN. § 505.050 (West 2010) (acquittal, conviction, or improper termination); MINS. STAT. ANN. § 609.045 (West 2010) (acquittal or conviction); MISS. CODE. ANN. § 99-11-27 (2010) (acquittal or conviction); MONT. CODE ANN. § 46-11-504 (2010) (acquittal conviction, or improper termination); MONT. CODE ANN. § 171.070 (LexisNexis 2010) (acquittal or conviction); N.J. STAT. ANN. § 2C:1-11 (West 2011) (acquittal, conviction, or improper termination); OKLA. STAT. ANN. tit. 22, § 130 (West 2010) (acquittal or conviction); 18 PA. CONS. STAT. ANN. § 111(1) (West 2010) (acquittal, conviction, or improper termination); UTAH CODE ANN. § 76-1-404 (2010) (acquittal conviction or improper termination); WASH. REV. CODE ANN. § 10.43.040 (LexisNexis 2010) (acquittal or conviction).
(1) where the initial prosecution requires proof of different facts or elements not required in the state prosecution;\(^{267}\)

(2) where the state statute addresses different harms or evils than the offense originally prosecuted;\(^{268}\) and

(3) where the state offense was not consummated at the time of the first prosecution.\(^{269}\)

The type of separate sovereigns to which the protection extends differs among the states. Many statutes provide expansive protections incorporating prior prosecutions in another “jurisdiction.”\(^{270}\) Others extend protection following prosecutions in another state, territory, or country,\(^{271}\) or just another state or country.\(^{272}\) Still others limit the initial jurisdiction to only state and federal courts,\(^{273}\) or only an initial prosecution by the U.S. government.\(^{274}\) The Mississippi Supreme Court held in 1997 that the statutory language “another state, territory or country” in its dual sovereignty statute did not encompass the


\(^{272}\) See Wash. Rev. Code Ann. § 10.43.040 (LexisNexis 2010).


U.S. government, thereby allowing the state government to pursue state charges for capital murder following a federal conviction for kidnapping.\textsuperscript{275} Most states addressing the issue include Native American tribal nations within their dual sovereignty bars,\textsuperscript{276} but the Supreme Court of Washington recently held that Native American tribes do not fall within the state’s statutory dual sovereignty bar that specifies prosecution in “another state or country” because such tribes are not expressly included in the statute.\textsuperscript{277}

<table>
<thead>
<tr>
<th>Table Three—State Statutory Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>State</strong></td>
</tr>
<tr>
<td>Arkansas</td>
</tr>
<tr>
<td>California</td>
</tr>
<tr>
<td>Colorado</td>
</tr>
<tr>
<td>Delaware</td>
</tr>
<tr>
<td>Georgia</td>
</tr>
<tr>
<td>Hawaii</td>
</tr>
<tr>
<td>Indiana</td>
</tr>
<tr>
<td>Kansas</td>
</tr>
<tr>
<td>Kentucky</td>
</tr>
<tr>
<td>Minnesota</td>
</tr>
<tr>
<td>Nevada</td>
</tr>
<tr>
<td>New Jersey</td>
</tr>
<tr>
<td>New York</td>
</tr>
<tr>
<td>North Dakota</td>
</tr>
<tr>
<td>Oklahoma</td>
</tr>
<tr>
<td>Pennsylvania</td>
</tr>
<tr>
<td>Utah</td>
</tr>
<tr>
<td>Washington</td>
</tr>
<tr>
<td>Wisconsin</td>
</tr>
</tbody>
</table>

\textsuperscript{275} Evans v. State, 725 So. 2d 613, 658-59 (Miss. 1997) (emphasis supplied), cert. denied, 525 U.S. 1133 (1999). The court held that under Mississippi law, resort to statutory construction is unnecessary where there is no ambiguity in the statute. \textit{Id.} Without explaining why the statute was unambiguous, and even while acknowledging the defendant’s argument that other states had reached different results, the court declared that the statute unambiguously excluded prosecutions by the United States. \textit{Id.} at 659.

\textsuperscript{276} See John E. Theuman, Annotation, \textit{Validity of State Prosecution Subsequent to Tribal Court Prosecution}, 116 A.L.R. 5th 313, §§ 2, 4[a].

\textsuperscript{277} State v. Moses, 37 P.3d 1216, 1218-20 (Wash. 2002).
The remaining states listed in Table Four have not adopted a limitation on subsequent state prosecutions. Many have done so explicitly in the legislation or state court holdings cited in Table Four. Others have not had occasion to address the issue as indicated in Table Four. Alaska and South Dakota at one time had dual sovereignty bars, but both have since repealed the protections.

<table>
<thead>
<tr>
<th>State</th>
<th>Authority</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td><em>Ex Parte Heath</em> (1984)</td>
</tr>
<tr>
<td>Alaska</td>
<td>Dual sovereignty protections repealed in 2008</td>
</tr>
<tr>
<td>Arizona</td>
<td><em>State v. Poland</em> (1982)</td>
</tr>
<tr>
<td>Connecticut</td>
<td><em>State v. Moeller</em> (1979)</td>
</tr>
<tr>
<td>Louisiana</td>
<td><em>State v. de la Beckwith</em> (1977)</td>
</tr>
</tbody>
</table>

278. Only state court decisions affirming the dual sovereignty doctrine after the Supreme Court’s application of double jeopardy protections to the states in *Benton v. Maryland* are included. Otherwise the state is considered not to have addressed the issue, even though prior state decisions may have approved the notion of subsequent state prosecutions. See, e.g., *State v. Mosely*, 114 S.E. 866, 868-69 (S.C. 1922); *State v. O’Brien*, 170 A. 98, 100-01 (Vt. 1934); *State v. Henson*, 114 S.E. 273, 274-75 (W. Va. 1923) (involving state alcohol related prosecutions following federal convictions under the Volstead Act during Prohibition, with all three courts finding no double jeopardy, since alcohol could be regulated under both state and federal law after the ratification of the Eighteenth Amendment).


280. [S.D. CODIFIED LAWS §§ 22-5-8 and 23-2-12](https://www.slo.legislature.gov/Legislation/Legislation.cfm) were repealed in 1978 after the state supreme court held in *State v. West* that the owner of a multi-million dollar meat packing enterprise who had fallen on hard times could not be prosecuted in state court after being convicted of several counts in federal court and acquitted of others. *State v. West*, 260 N.W.2d 215, 219-20 (S.D. 1977). Although the statutes had been law since South Dakota was a territory, the *West* decision was the first time a court had interpreted them. *Id.* at 220. The statutes were repealed the following year. See *State v. Chavez*, 668 N.W.2d 89, 99-100 (S.D. 2003).


282. See *supra* note 279.


Drug cases comprise a special category for many states. The states listed in Table Five have specific statutes that afford limitations on the dual sovereignty doctrine in drug cases. Two states that provide no comprehensive dual sovereignty limitations, Alabama and Arizona, at one time afforded specific statutory protection for drug offenses, but both statutes have since been repealed.

297. See supra note 280.  
301. See infra Table Five.  
302. See infra Table Five.
Table Five—State Statutory Limitations for Drug Cases

<table>
<thead>
<tr>
<th>State</th>
<th>Statute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>ALA. CODE § 20-2-77 (Repealed)</td>
</tr>
<tr>
<td>Arizona</td>
<td>ARIZ. REV. STAT. ANN. § 36-2532, et seq. (Repealed)</td>
</tr>
<tr>
<td>Arkansas</td>
<td>ARK. CODE. ANN. § 5-64-418 (2010)</td>
</tr>
<tr>
<td>Michigan</td>
<td>MICH. COMP. LAWS ANN. § 333.7409 (West 2010)</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>OKLA. STAT. ANN. tit. 63, § 2-413 (West 2010)</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>WIS. STAT. ANN. § 961.45 (West 2010)</td>
</tr>
</tbody>
</table>

V. A MORE JUST AND REASONABLE APPROACH

While dual sovereignty remains a constitutionally viable doctrine, it clearly has its detractors. The U.S. Supreme Court has emphasized the need for restraint, even while affirming the constitutionality of the doctrine. In response, the DOJ has adopted internal policies, albeit unenforceable ones, intended to circumscribe the use of dual sovereignty to only instances in which demonstrable federal interests are at stake.

At least two state supreme courts, New Hampshire and Pennsylvania, have not been nearly as kind in expressing their opinions concerning the doctrine. Additionally, nearly half the states have limited the doctrine by constitution or statute. All tolled, half the states and the federal government affirmatively

304. ARIZ. REV. STAT. ANN. § 36-2532 et. seq., repealed by Laws 1981, Ch. 264, § 38, effective September 1, 1981.
305. ARK. CODE. ANN. § 5-64-418 (2010) (barring prosecution for drug offenses if the defendant was already convicted or acquitted for the “same act” under federal law or another state’s law).
306. MICH. COMP. LAWS ANN. § 333.7409 (West 2010) (barring state prosecution for drug offenses if the defendant was already convicted or acquitted for the “same act” under federal law or another state’s law).
307. OKLA. STAT. ANN. tit. 63, § 2-413 (West 2010) (barring state prosecution for drug offenses if the defendant was already convicted or acquitted for the “same act” under federal law or another state’s law).
308. WIS. STAT. ANN. § 961.45 (West 2010) (barring state prosecution for drug offenses if the defendant was already convicted or acquitted for the “same act” under federal law or another state’s law).
309. See supra note 166 and accompanying text.
310. See supra Part III.
311. See supra note 249 (New Hampshire) and supra notes 250-52 (Pennsylvania) and accompanying text.
312. See supra Tables 2 & 3.
313. Possibly more, if the position of Michigan could be ascertained in light of recent court holdings. See supra notes 248, 258-60 and accompanying text.
limit subsequent prosecutions in some manner if double jeopardy attaches in an initial prosecution by a different sovereign.\textsuperscript{314}

Many judges, lawyers, and scholars have condemned the dual sovereignty doctrine as having outlived its usefulness, particularly in light of the increasing federalization of criminal law.\textsuperscript{315} Some maintain that the dual sovereignty doctrine may even discourage cooperation among sovereigns in pursuing criminal charges, and engender distrust among them, particularly between the states and the federal government.\textsuperscript{316} Other critics advocate the application of the doctrine only in areas that states fail to address issues adequately, such as civil rights violations.\textsuperscript{317} Concerns expressed by judges and scholars do little more than beg the questions of whether the dual sovereignty doctrine, in its current manifestation, is truly compatible with double jeopardy protections, and, if not, whether it can be reformulated so as to address recurrent questions about the relationship.

Given that there must be a constitutional basis for the exercise of federal criminal jurisdiction,\textsuperscript{318} the problem might best be approached on federal preemption grounds. In areas of federal jurisdiction where the United States believes the states may be more lax than federal interests warrant, Congress could preempt state law and make such issues solely matters of federal concern under the Supremacy Clause of the U.S. Constitution.\textsuperscript{319} This would bar states from initiating such prosecutions, thereby obviating the need for the dual sovereignty doctrine. Such an approach would undoubtedly strain an already overtaxed federal criminal justice system, but it would also force the federal government to identify those crimes it truly regards as matters of national concern that the Supreme Court envisioned when it discussed the Petite Policy in Rinaldi.\textsuperscript{320}

Another approach might be to make the requirements of the Petite Policy, both substantive and procedural, inure to the benefit of criminal defendants and also put the onus on defendants to raise the issue pre-trial, much as double

\begin{itemize}
\item \textsuperscript{314} See supra Tables 2 & 3.
\item \textsuperscript{315} See supra Part II.
\item \textsuperscript{316} Merkl, supra note 10, at 196.
\item \textsuperscript{317} See id. at 199 (suggesting that Congress should have a “more restrained approach to enacting criminal laws and limiting them to certain areas as appropriate for federal criminal jurisdiction). See supra note 133 and accompanying text.
\item \textsuperscript{318} Screws v. United States, 325 U.S. 91, 108-09 (1945).
\item \textsuperscript{319} The power to preempt in Article VI might be particularly useful in areas such as Civil Rights violations. Article VI, Clause 2 of the U.S. Constitution provides:

This Constitution, and the Laws of the United States which shall be made in Pursuance thereof; and all Treaties made, or which shall be made, under the Authority of the United States, shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, any Thing in the Constitution or Laws of any State to the Contrary notwithstanding.

U.S. CONST. art. VI, cl. 2.
\item \textsuperscript{320} Rinaldi v. United States, 434 U.S. 22, 29 (1977).
\end{itemize}
jeopardy objections are. This would ensure that the objectives of the Policy would be achieved before resources are invested in a trial. It would also have the benefit of calling technical violations, such as failure to obtain a waiver of the Policy from the appropriate DOJ authority, to the attention of prosecutors so they can be remedied, if possible. Substantive issues addressed by the Policy, such as vindication of federal interests or an inadequate or incompetent prior prosecution, could be addressed in a pre-trial hearing using a preponderance of the evidence standard, with the government having the burden of proof.

Yet another approach might be to make only an acquittal an absolute bar to reprosecution by the federal government. However, while this might solve the problem of defendants who plead guilty to another sovereign’s offense to escape harsher punishment under a federal statute, it would not address the issue of jury nullification of unpopular laws, as seen in the South during the Civil Rights Era.321 In the twenty-first century, one would hope such problems do not exist to the extent they did before the Civil Rights Movement. However, racial prejudice continues to exist, as do biases based on sect, ethnicity, and sexual orientation.

This approach also fails to address the equally dogmatic issue of subsequent prosecutions by state governments. While a ban or enforceable limitation on reprosecution by the federal government would require only an appropriate statute or regulation, prohibiting subsequent prosecutions by the states under longstanding dual sovereignty jurisprudence would require either a constitutional amendment or an abrupt, wholesale change of heart by the U.S. Supreme Court regarding the constitutionality of the dual sovereignty doctrine, neither of which is likely.

There does, however, remain the possibility that, in an era that has witnessed an enormous increase in federal criminal law that often duplicates or overlaps state and local law, the Supreme Court may come to the realization that its prior pronouncements relating to the need for vindication of federal and state interests are overstated in light of intervening circumstances. The Court might well conclude that the cure of dual sovereignty may have become worse than the disease of unvindicated sovereign interests.

SPOT DELIVERY: WHY IT IS ILLEGAL AND WHY CAR DEALERS HAVE BEEN GETTING AWAY WITH THEFT FOR SO LONG

Gregory R. Ingalsbe* and Raymond G. Ingalsbe**

I. INTRODUCTION: HOW “SPOT DELIVERY” WORKS

Jane Jones is in the market to buy a new automobile. She goes to Dave’s Dealership to search the lot for the perfect car. A salesperson at Dave’s shows Jane a new, fuel-efficient economy car and she instantly falls in love. She takes it for a test drive and returns to Dave’s to complete the purchase. The price of the new car is $15,000. Jane trades in her old car for $1,500 and finances the remaining $13,500. Jane signs several forms prepared by Dave’s Dealership. One of the forms is a Retail Installment Sales Contract (“RISC”). The RISC discloses terms, including the identity of the creditor, the amount financed, the Annual Percentage Rate (“APR”), and the total of payments. The Federal Truth in Lending Act (“TILA”) requires disclosure of these as well as other terms. The RISC discloses the dealership, not a third-party bank or finance company, as the creditor. After receiving Jane’s credit application and obtaining her credit report, Dave’s offers financing at 6% interest. Another form (or provision within the RISC itself) the dealership has Jane sign purports to disclose that if Dave’s chooses not to assign the contract at the 6% APR, the

* Gregory R. Ingalsbe, lead author on this article, is a 2011 J.D. candidate at Salmon P. Chase College of Law, Northern Kentucky University. He earned a B.A. in Sociology from Boston University in 2004.
** Raymond G. Ingalsbe, who co-authored this article, is an attorney in private practice in Palm Beach Gardens, Florida. He earned his B.A. in Psychology from the University of Cincinnati, and his J.D. from the University of Miami (Florida). Mr. Ingalsbe received Florida Bar certification as a civil trial lawyer in 1985. He was a 30-year member of the Florida Justice Association, f/k/a Academy of Florida Trial Lawyers and is involved in attempting to draft changes to Florida’s Deceptive and Unfair Trade Practices Act with the assistance of the Florida Justice Association. Mr. Ingalsbe was the first and only chairman of the Palm Beach County Bar Association Consumer Law Committee. He is an occasional lecturer on sales fraud and deceptive and unfair trade practices, has received two pro bono awards from the Palm Beach County Bar Association for his work in consumer law, and received an award for his work on the Academy of Florida Trial Lawyers Journal. Mr. Ingalsbe is a charter member of the National Association of Consumer Advocates, and has been a consumer plaintiffs’ attorney for the past 20 years. Mr. Ingalsbe specializes in auto-dealer spot delivery fraud/TILA, FCRA, ECOA, and UDAP cases, both individual and class actions. He currently has cases pending in Florida, Ohio, Kentucky, and Indiana.

3. Id. §§ 1601 – 1667f.
dealership can unilaterally revoke the contract. After delivering the car to Jane, Dave’s transmits her credit information to nine different banks in an attempt to assign the contract at face value.

Two weeks later, Dave’s contacts Jane claiming that there were problems with the financing and asks her to return to the dealership at once to get the financing squared away. When Jane arrives, the dealer tells her the bank rejected credit at 6% APR but would finance the deal at 12% APR. In reality, because Dave’s Dealership had no intention of honoring its RISC obligations, and in fact did not do so, Dave’s Dealership alone rejected credit, which is an adverse action requiring written notice. The dealer then gives Jane the “Hobson’s choice” to either sign a new RISC at the 12% APR or surrender the car that she has fallen in love with and now depends upon for transportation. Having no other transportation and having been told by Dave’s that it already sold her trade-in, Jane is left with no choice but to sign the new RISC, resulting in a substantial increase in the total sale price of the car. This happens every day in almost every state. It is a violation of several federal statutes, and it is criminal theft.

The scenario described above is known euphemistically as “spot delivery” by car dealers, and “yo-yo” by consumer lawyers. In other words, the dealership delivers the car to the purchaser on the spot (the same day the buyer first steps on the lot), with no intention of itself extending the credit as offered, and with no intention of assigning the RISC unless the prospective assignee agrees to buy it as written. Dealers often do this intending to call the consumer back to the lot to increase profit at the consumer’s expense. They “spot deliver” only to take the consumer out of the car-buying market, let him or her fall in love with the car, and call him or her back to the dealership (as if on a yo-yo string) knowing full-well that the consumer will, and often must, agree to the new, more expensive terms—classic bait and switch. Unfortunately, to date, most courts that have confronted this issue have sided with the dealerships in contravention of consumer protection statutes and without any analysis of TILA or Regulation Z, TILA’s implementing regulation.

4. Id. § 1691(d)(2); 12 C.F.R. § 202.9(a) (2011).
6. See id.
7. See Elwin Griffith, Lenders and Consumers Continue the Search for the Truth in Lending Under the Truth in Lending Act and Regulation Z, 44 San Diego L. Rev. 611, 614-15 (2007) (describing the bait-and-switch tactic as harmful in these cases because a consumer would not expect the lender to make certain disclosures without advising the consumer that the disclosed terms would not last very long).
The problem exists because most courts and, indeed, most attorneys that have confronted the issue are victims of “groupthink”: “The act or practice of reasoning or decision-making by a group, esp[ecially] when characterized by uncritical acceptance or conformity to prevailing points of view.” As it pertains to spot delivery, groupthink results from a belief that some third party is the creditor who rejected the buyer’s credit; that the dealer is not a creditor. Often, judges and consumers’ attorneys alike have not read or do not understand the controlling statutes and regulations; they uncritically accept the dealer’s argument on spot delivery which, when analyzed, amounts to bait and switch, and assume that it is legal. They point to the body of case law erroneously holding that spot delivery is legal based upon the unsupportable assumption that the dealer itself is not the creditor. Not only is spot delivery illegal, it is a violation of TILA, the Fair Credit Reporting Act (“FCRA”), the Equal Credit Opportunity Act (“ECOA”), most state Unfair Deceptive Acts and Practices (“UDAP”) statutes, and even state criminal codes.

II. BACKGROUND

A. The Truth in Lending Act and Regulation Z

Congress enacted the Truth in Lending Act in 1968 to protect consumers. TILA’s stated purpose is “to assure a meaningful disclosure of credit terms so that the consumer will be able to compare more readily the various credit terms available to him . . . and to protect the consumer against inaccurate and unfair credit . . . practices.” As a federal statute, TILA trumps any inconsistent state law. TILA itself provides that it does not “annul, alter, or affect the laws of any states relating to the disclosure of information in connection with credit transactions, except to the extent that those laws are inconsistent with the provisions of this subchapter and then only to the extent of the inconsistency.” Congress went on to clarify that certain classes of credit transactions within a state are exempt from TILA requirements if state law subjects those transactions

16. Id. § 1601(a).
17. U.S. CONST. art. VI, cl. 2.
to requirements that are “substantially similar” to TILA’s, “or that such law gives greater protection to the consumer, and that there is adequate provision for enforcement.” In short, Congress enacted TILA to provide consumers more protection from unscrupulous credit sellers than inconsistent state laws, as some state laws allow spot delivery by restricting its application, believing that it benefits consumers.

TILA requires several disclosures to be made in every consumer credit transaction, including:

1. The creditor’s identity;
2. The “amount financed”;
3. The “finance charge”;
4. The “annual percentage rate”;
5. The “total of payments”;
6. The payment schedule; and
7. The “total sale price.”

Regulation Z, the Truth in Lending Act’s implementing regulation, defines a creditor as “[a] person who regularly extends consumer credit that is subject to a finance charge or is payable by written agreement in more than four installments . . . and to whom the obligation is initially payable, either on the face of the note or contract, or by agreement . . . .” An assignee is not a creditor under TILA.

Indeed, the commentary to Regulation Z discusses the difference between assignees and creditors:

Assignees. If an obligation is initially payable to one person, that person is the creditor even if the obligation by its terms is simultaneously assigned to another person. For example:

- An auto dealer and a bank have a business relationship in which the bank supplies the dealer with credit sale contracts

19. Id. § 1666j(b).
20. See Hickman v. Cliff Peck Chevrolet, Inc., 566 F.2d 44, 47 (8th Cir. 1977) (concluding “that state law should not be applied where to do so would require a course of action in contravention of a specific direction in a federal statute”). See also Williams v. First Gov’t. Mortg. & Investors Corp., 176 F.3d 497, 500 (D.C. Cir. 1999) (explaining that “[n]othing in TILA or its legislative history suggests that Congress intended the Act’s disclosure regime to provide the maximum protection to which borrowers are entitled nationwide; states remain free to impose greater protections for borrowers”).
22. 15 U.S.C § 1638(a); 12 C.F.R. § 226.18 (2011).
24. Id. § 226.1.
25. Id. § 226.2(a)(17)(i).
that are initially made payable to the dealer and provide for the immediate assignment of the obligation to the bank. The dealer and purchaser execute the contract only after the bank approves the creditworthiness of the purchaser. Because the obligation is initially payable on its face to the dealer, the dealer is the only creditor in the transaction.  

Thus, the dealership is the only creditor in spot delivery transactions, and is therefore subject to TILA requirements.

Regulation Z defines consummation as “the time that a consumer becomes contractually obligated on a credit transaction.” The point at which consummation occurs is important because sellers must make TILA disclosures before consummation of the transaction. The American Heritage Dictionary defines “to consummate” as “to bring to completion or fruition; conclude.” Unfortunately for consumers, some courts have misconstrued when consummation occurs in motor vehicle financing transactions.


In 1998, Diane Janikowski purchased a new Ford Escort from Lynch Ford. The salesperson prepared a purchase order and RISC, disclosing the APR at 5.9%. The salesperson informed Janikowski that he was not sure a bank would accept the contract on those terms, but Janikowski signed the buyer’s order and RISC anyway. The purchase order included what is often referred to as a “Seller’s Right to Cancel,” stating that if the dealership could not obtain the agreed-upon financing within five business days, either party to the transaction could cancel the contract. The day after the original sale, Lynch Ford called Janikowski to inform her that her financing had been approved at 11.9% APR—a
full six percentage points higher than the original agreement.\textsuperscript{38} She returned to
the dealership later that day and signed a new RISC at the 11.9\% interest rate.\textsuperscript{39}

Janikowski later sued Lynch Ford for violations of TILA, as well as an
Illinois consumer protection statute, claiming that the dealership violated TILA
by first offering a 5.9\% APR but later requiring the 11.9\% RISC.\textsuperscript{40} She also
argued that the dealership violated TILA by failing to disclose that the 5.9\%
interest rate was an estimate, as she believed was required by the Act.\textsuperscript{41} The
United States District Court for the Northern District of Illinois granted summary
judgment for Lynch Ford, and Janikowski appealed.\textsuperscript{42}

On appeal the Seventh Circuit Court of Appeals affirmed.\textsuperscript{43} First, the court
acknowledged that TILA requires RISCs to reflect in writing “‘the terms of the
legal obligation of the parties,’ . . . before the ‘consumer becomes contractually
obligated on a credit transaction.’”\textsuperscript{44} However, the court rejected Janikowski’s
claim that Lynch Ford’s initial disclosure of 5.9\% and subsequent increase to
11.9\% violated TILA.\textsuperscript{45} Instead, the court held that Janikowski entered into two
different valid contracts, one at 5.9\% and another at 11.9\%.\textsuperscript{46} When Janikowski
signed the first contract, “Lynch Ford disclosed a contractual rate of 5.9\%. That
disclosure reflected the terms of her legal obligations, as required by
regulation.”\textsuperscript{47} However, Lynch Ford cancelled that contract the following day
after the 5.9\% APR was denied, and Janikowski signed a new contract at 11.9\% APR.\textsuperscript{48}
Thus, according to the Seventh Circuit, “even though Janikowski did not
eventually obtain financing at 5.9\%, Lynch Ford did not violate TILA because it
accurately disclosed her legal obligations under the two contracts.”\textsuperscript{49}

The Seventh Circuit likewise rejected Janikowski’s claim that Lynch Ford’s
failure to disclose the 5.9\% APR as an estimate was a violation of TILA.\textsuperscript{50} The
court cited Regulation Z: “If any information necessary for an accurate
disclosure is unknown to the creditor, the creditor shall make the disclosure
based on the best information reasonably available at the time the disclosure is
provided to the consumer, and shall state clearly that the disclosure is an
estimate.”\textsuperscript{51} Nevertheless, the court reiterated its contention that 5.9\% was not
an estimate; it was the contractual rate. 52 “If the financing condition had been satisfied, Janikowski would be able and obligated to purchase the car at 5.9%. However, when Janikowski did not receive approval for financing at 5.9%, she could have canceled the contract and refused to purchase the Escort.” 53 Lynch Ford, on the other hand, exercised its “right” to cancel the contract. 54

Thus, the Seventh Circuit found that Lynch Ford disclosed the actual terms of the sale in both RISCs, conditioned on the dealership’s ability to assign the contract to a third party that would accept the terms as written. 55 Most courts confronted with this spot delivery issue have adopted a similar line of reasoning. 56 However, some courts view it differently and hold that these sales practices are indeed a violation of TILA. 57

2. Patton v. Jeff Wyler Eastgate, Inc. 58

In 2005, Brian and Jennifer Patton entered into a contract to purchase a used Ford Windstar from Jeff Wyler Eastgate (“Wyler Eastgate”). 59 The Pattons signed a RISC that identified them as the “Buyer (and Co-Buyer)” and Wyler Eastgate as the “Creditor-Seller.” 60 The RISC contained a clause captioned “No Cooling Off Period,” which disclosed that “[s]tate law does not provide for a ‘cooling off’ or cancellation period for this sale. After you sign this contract, you may only cancel it if the seller agrees or for legal cause. You cannot cancel this contract simply because you change your mind.” 61 The RISC also contained a merger clause providing that “[t]his contract is the entire agreement between you and us relating to this contract. Any change in this contract must be in writing and we must sign it. No oral changes are binding.” 62

The Pattons also signed Wyler Eastgate’s “Purchase Spot Delivery Agreement.” 63 This form purported to disclose that Wyler Eastgate was allowing the Pattons to take possession of the Windstar even though financing had not been finalized, and that the parties intended the purchase to be final only upon

52. Janikowski, 210 F.3d at 768.
53. Id.
54. See id. at 768.
55. See id. at 768.
59. Id. at 908.
60. Id.
61. Id. at 909.
62. Id.
63. Id.
the RISC being duly assigned. 64 The form went on to provide that if Wyler Eastgate could not assign the RISC to a third party within forty-five days of the original transaction, the Pattons were required to immediately return the Windstar or pay the entire balance due. 65 The spot delivery form was incorporated into the Buyer’s Order (purchase order) by reference. 66 However, “[n]either the Purchase Spot Delivery Agreement nor the Installment Contract contain[ed] language which would incorporate the terms of one agreement into the other.” 67 The Pattons traded in their old car and took delivery of the Windstar. 68 Over a month later, Wyler Eastgate told the Pattons that it was unable to assign the RISC to a third party and demanded that they return the Windstar. 69 By this time, Wyler Eastgate had sold the Pattons’ trade-in, although the dealership eventually reclaimed it and returned it to them. 70

The Pattons filed suit in federal court claiming violations of TILA and Regulation Z. 71 Specifically, the Pattons claimed that the “Purchase Spot Delivery Agreement” (“Spot Delivery Agreement”) rendered the RISC’s financing terms meaningless by giving the dealership the unilateral right to cancel the contract if it could not assign it to a third party. 72 By purporting to retain this right, the Pattons claimed that the terms were illusory because at the time the deal was consummated the deal was subject to change or cancellation at the dealer’s discretion. 73 Wyler Eastgate contested the allegations, claiming that it could not be liable under TILA because it was not a creditor under that statute. 74

The district court first analyzed whether Wyler Eastgate was a creditor under TILA and Regulation Z. 75 It noted that the statutory definition of creditor has two requirements: “(1) that the creditor be someone who regularly extends consumer credit and (2) that the creditor is someone to whom the debt is initially payable based on the face of the obligation.” 76 Noting that Wyler Eastgate identified itself as the creditor on the face of the RISC, the court found that it had satisfied the second prong of the definition—that it was the “person to whom

64. Patton, 608 F. Supp. 2d at 909.
65. Id.
66. Id.
67. Id.
68. Id.
69. Id.
70. Patton, 608 F. Supp. 2d at 911.
71. Id.
73. Id. ¶ 4.
74. Patton, 608 F. Supp. 2d at 911. Wyler Eastgate also asserted that the Spot Delivery Agreement did not violate TILA “because other federal courts have upheld the practice of ‘spot delivery’ against TILA challenges.” Id.
75. Id.
76. Id. at 912.
the debt arising from the consumer credit transaction is initially payable on the face of the evidence of indebtedness.”

The court also concluded that Wyler Eastgate regularly extended credit to consumers, thus satisfying the first prong of the definition. Because the RISC listed Wyler Eastgate as the creditor to which the Pattons were to make their monthly payments, the dealership was extending credit to consumers. “Until that Installment Contract was assigned or cancelled in a contractual manner, Wyler Eastgate was the entity obligated to provide financing to the Pattons on the terms provided for their purchase of the Windstar.” Because between sixty and seventy percent of Wyler Eastgate’s 390 average monthly sales were financing transactions using RISCs similar to the one signed by the Pattons (which identified Wyler Eastgate as the creditor), the court had no trouble concluding that the dealership regularly extended credit to consumers. Therefore, Wyler Eastgate met the TILA definition of “creditor.”

The Patton court distinguished Janikowski on two bases. First, in Janikowski, the purchase order gave both the dealer and the consumer the right to cancel the contract within five days if the financing terms could not be assigned, whereas in Patton, the RISC bound both parties immediately. Second, the transaction in Janikowski was expressly conditioned on assignment to a third-party bank or financial institution, and was not immediately binding as in Patton. In any event, the Patton court found that Wyler Eastgate’s practice of using the Spot Delivery form violated TILA. While the RISC in Patton provided for its assignment, it did not make the transaction’s validity contingent upon that assignment. Rather, the Spot Delivery form purported to give Wyler Eastgate the unilateral ability to revoke the RISC if it chose not to assign it, rendering the TILA disclosures in the RISC illusory. “The representation in the [RISC] that Wyler Eastgate is providing credit on the financing terms stated in the TILA disclosures section for the Pattons’ purchase is rendered meaningless by the language in the Purchase Spot Delivery Agreement.”

Finally, the court concluded that state contract law permitting courts to read several documents related to a single transaction collectively did not apply in

77. Id. (citing 15 U.S.C. §1602(f) (2006)).
78. Id. at 913.
79. Id.
81. Id.
82. Id.
83. Id. at 914-15.
84. Id. at 914.
85. Id. at 914-15.
86. Patton, 608 F. Supp. 2d at 916.
87. Id. at 915.
88. Id.
89. Id.
First, the RISCs merger clause “purport[ed] to ‘contain the entire agreement between [the parties] relating to [the] contract’” without referencing or incorporating any other document, thus excluding the Spot Delivery Agreement as part of the contract. Second, the inconsistencies between the RISC and Spot Delivery Agreement “make it impossible to reconcile the two contractual documents and read them as a coherent single contract.” Third, and perhaps most importantly, the legislative purpose behind TILA is to protect consumers. Therefore, TILA should be “considered liberally in favor of the consumer. The purpose of TILA would be frustrated if automobile dealerships are permitted to rescind the terms of integrated automobile retail installment sales contracts by use of a second, contradictory form.” Accordingly, the court granted summary judgment in favor of the Pattons.

B. The Fair Credit Reporting Act

Congress enacted the Fair Credit Reporting Act (“FCRA”), another consumer protection statute, in 1970 to ensure “fair and accurate credit reporting.” Not only did Congress determine that the entire banking system relies upon a system of fair credit reporting, but also that consumer privacy depends upon it. The FCRA’s compliance procedures require a user of consumer reports (e.g., a car dealer) to certify its “permissible purpose” for accessing its customers’ credit reports. Likewise, the FCRA states that a person may not request a consumer report for any reason except for a certified permissible purpose. The FCRA borrows its definition of “creditor” from the Equal Credit Opportunity Act (“ECOA”), stating that “[a]ny person who regularly extends, renews or continues credit; any person who regularly arranges for the extension, renewal, or continuation of credit; or any assignee of an original creditor who participates in the decision to extend, renew, or continue credit.”

The FCRA provides an exclusive list of permissible purposes for which a credit-reporting agency may provide a consumer report to a person or entity,
such as a car dealer. The statute also prohibits using or obtaining a consumer report for any purpose unless the reporting agency is authorized to provide it (that is, unless the user is requesting it for a certified permissible purpose). In pertinent part, the “permissible purposes” provision allows a credit reporting agency to release a consumer report to a person who “intends to use the information in connection with a credit transaction involving the consumer on whom the information is to be furnished and involving the extension of credit to, or review or collection of an account of, the consumer . . . .” An example of a dealer’s FCRA permissible purpose certification is as follows:

You [subscriber-dealer] certify that you will only order Information Services that are consumer reports (“Basic Reports”) credit risk scores (“Scores”) and other enhancements to the Basic Reports only while operating as a motor vehicle dealership and only on written authorization from your customer in connection with your proposed extension of credit to that customer.

There are three elements to a claim for improper use or acquisition of a consumer’s credit report: “(i) that there was a ‘consumer report’ within the meaning of the statute; (ii) that the defendant used or obtained it; and (iii) that the defendant did so without a permissible statutory purpose.” The FCRA provides remedies for willful noncompliance as well as negligent noncompliance. Persons who are willfully noncompliant are liable for actual damages, punitive damages, court costs, and attorney’s fees. Persons who are negligently noncompliant are responsible only for actual damages and attorney’s fees.

C. The Equal Credit Opportunity Act and Regulation B

The ECOA is another consumer protection statute implicated in spot delivery transactions, requiring creditors to issue written adverse action notices to consumers against whom an adverse action was taken. The ECOA defines “adverse action” as “a denial or revocation of credit, a change in the existing credit arrangement, or a refusal to grant credit in substantially the amount or on

105. Id. § 1681b(a).
106. Id. § 1681b(f).
107. Id. § 1681(b)(3)(A).
111. Id. § 1681o.
112. Id. § 1681n; Regulation B, 12 C.F.R. § 202.9(a) (2011).
114. Id. § 1691(d).
substantially the terms requested.”\textsuperscript{115} The ECOA defines “creditor” (adopted by the FCRA) as “any person who regularly extends, renews or continues credit; any person who regularly arranges for the extension, renewal, or continuation of credit; or any assignee of an original creditor who participates in the decision to extend, renew, or continue credit.”\textsuperscript{116} Thus, if a creditor, as defined by the ECOA, denies or revokes credit to a consumer who has applied for credit, the creditor is required to notify the consumer in writing of the reasons such adverse action was taken.\textsuperscript{117} The creditor can satisfy its obligation by providing the consumer with written notification (1) of the reasons the adverse action was taken, or (2) that the adverse action was taken and disclosing that the consumer has a right to be apprised of the reasons, as well as from whom that information may be obtained.\textsuperscript{118}

The ECOA provides that consumers, either individually or as a class, are entitled to actual damages suffered as a result of a creditor failing to comply with any ECOA requirement.\textsuperscript{119} Punitive damages are also available to consumers, either individually or as a class.\textsuperscript{120} The court determines the amount of punitive damages to a certified class after considering “among other relevant factors, the amount of any actual damages awarded, the frequency and persistence of failures of compliance by the creditor, the resources of the creditor, the number of persons adversely affected, and the extent to which the creditor’s failure of compliance was intentional.”\textsuperscript{121} However, the ECOA caps punitive damage awards to individuals at $10,000, and punitive damage awards in a class action at $500,000 or one-percent of the creditor’s net worth, whichever amount is less.\textsuperscript{122} Additionally, equitable and declaratory relief are also available,\textsuperscript{123} as well as attorney’s fees.\textsuperscript{124} All of the ECOA provisions are implemented in Regulation B of the Code of Federal Regulations.\textsuperscript{125}

\textsuperscript{115} Id. § 1691(d)(6).
\textsuperscript{116} Id. § 1691a(e).
\textsuperscript{117} Id. § 1691(d)(2).
\textsuperscript{118} Id. § 1691(d)(2)(B).
\textsuperscript{120} Id. § 1691e(b).
\textsuperscript{121} Id.
\textsuperscript{122} Id.
\textsuperscript{123} Id. § 1691e(c).
\textsuperscript{124} Id. § 1691e(d).
D. State Law

1. Unfair and Deceptive Acts and Practices

Most states have Unfair and Deceptive Acts and Practices ("UDAP") statutes enacted to protect consumers.\footnote{See Carolyn L. Carter, Consumer Protection in the States: A 50-State Report on Unfair and Deceptive Acts and Practices Statutes 3 (2009), available at http://www.scribd.com/doc/14473414/50-State-Report-on-Unfair-and-Deceptive-Acts-and-Practices.} For example, the Ohio Consumer Sales Practices Act ("OCSPA") provides generally that "[n]o supplier shall commit an unconscionable act or practice in connection with a consumer transaction. Such an unconscionable act or practice by a supplier violates this section whether it occurs before, during, or after the transaction."\footnote{Ohio Consumer Sales Practices Act, OHIO REV. CODE ANN. § 1345.03(A) (West 2010).} When determining the unconscionability of an act or practice, courts are directed to consider:

1. Whether the supplier has knowingly taken advantage of the inability of the consumer reasonably to protect the consumer’s interests because of the consumer’s physical or mental infirmities, ignorance, illiteracy, or inability to understand the language of an agreement; . . .

2. Whether the supplier knew at the time the consumer transaction was entered into of the inability of the consumer to receive a substantial benefit from the subject of the consumer transaction; [and] . . .

3. Whether the supplier required the consumer to enter into a consumer transaction on terms the supplier knew were substantially one-sided in favor of the supplier . . . .\footnote{Id. § 1345.03(B).}

Similarly, the Florida Deceptive and Unfair Trade Practices Act ("FDUTPA") states generally that "[u]nfair methods of competition, unconscionable acts or practices, and unfair or deceptive acts or practices in the conduct of any trade or commerce are hereby declared unlawful."\footnote{Florida Deceptive and Unfair Trade Practices Act, FLA. STAT. ANN. § 501.204 (West 2010).} However, the FDUTPA does not provide specific instances of what in fact constitutes an unfair, deceptive, or unconscionable act or practice, leaving the determination to the trier of fact.\footnote{See id.} Nevertheless, a violation of a consumer protection statute such as TILA, the FCRA, or the ECOA would almost certainly constitute a violation of FDUTPA.\footnote{See also Kaumui v. Citibank, N.A., No. 09-000258 ACK-BMK, 2009 WL 3530373, at *1 (D. Haw. Oct. 28, 2009) (in which the court noted that the plaintiff claimed that the defendant’s alleged violation of the Truth In Lending Act constituted unfair and deceptive acts or practices in violation of Hawaii’s UDAP statutes).}
2. State Criminal Codes

Theft is a criminal offense in every state.132 For example, in Ohio, “[n]o person, with purpose to deprive the owner of property or services, shall knowingly obtain or exert control over either the property or services in any of the following ways: . . . (3) [b]y deception…”133 In Ohio, theft of property worth between $500 and $4,999.99 is a fifth-degree felony; theft of property worth between $5,000 and $99,999.99 is grand theft, as well as a fourth-degree felony.134 Similarly in Illinois, “[a] person commits theft when he knowingly: (2) [o]btains by deception control over property of the owner.”135 In Illinois, theft of property worth between $300 and $10,000 is a Class 3 felony.136

III. ANALYSIS

A. Truth In Lending Act—What Janikowski Got Wrong and What Patton Got Right

1. Janikowski—An Exercise in Judicial Sophistry

To put it generously, the Seventh Circuit in *Janikowski v. Lynch Ford* applied faulty reasoning to the facts in that case.137 More critically, the Seventh Circuit misunderstood and misconstrued TILA and Regulation Z, especially their consumer protection purpose, accepted without analysis Lynch Ford’s contentions, and turned both the plain language and the congressional purpose of the Truth in Lending Act on their head.138 The court’s opinion advances a nonsensical interpretation of TILA.139 As a published circuit appellate decision, courts around the country mistakenly follow *Janikowski* in the belief that a federal appellate decision always applies correct legal principles.140 Thus, not only is *Janikowski* wrong, but it has also created a snowball effect of case law that continues to work to the detriment of consumers.141

---

134. Id. § 2913.02(B)(2).
136. Id. at 5/16(b)(4).
137. See Janikowski v. Lynch Ford, Inc., 210 F.3d 765, 767-69 (7th Cir. 2000).
138. Id.
139. See id. at 767.
First, the Seventh Circuit concluded that Lynch Ford did not violate TILA by initially obtaining Janikowski’s signature on a 5.9% RISC, then changing it to an 11.9% RISC. The court reasoned that Janikowski had entered into two legally binding contracts, both of which disclosed the terms required by TILA. In other words, Janikowski was legally bound by 5.9% APR for one day, and when Lynch Ford called Janikowski back to the dealership she had every right to refuse to sign the new contract, which reflected financing terms that had more than doubled. Rather, she could have chosen to not buy the car and to cancel the contract. Perhaps the court reasoned that because the buyer’s order purported to give both buyer and seller the right to cancel the contract reflecting the 5.9% APR, the contract did not violate TILA.

The Seventh Circuit misconstrued TILA and the Lynch Ford RISC, and it accepted without question Lynch Ford’s contention that it could make its RISC contingent upon its desire to sell the RISC to a third party. The court did not stop to analyze Lynch Ford’s contention because, in fact, Lynch Ford claimed that there was no creditor and no consummation until and unless it chose to assign the RISC to some third party; the third party being the creditor, not Lynch Ford. In other words, Lynch Ford contended that it never had any TILA obligations other than to insert numbers on the RISCs in the TILA disclosure boxes, that TILA applied only if and when some third party accepted assignment, and that TILA then obligated that third party, not Lynch Ford. According to Lynch Ford, because it did not assign Janikowski’s RISC, there was no creditor and no consummation. This violates TILA and Regulation Z which mandate that consummation occurs when the buyer signs the RISC.

142. Janikowski, 210 F.3d at 767.
143. Id.
144. Id. at 766.
145. Id. at 768.
146. Although the opinion is silent in this respect, other courts have interpreted Janikowski this way. See, e.g., Patton v. Jeff Wyler Eastgate, Inc., 608 F. Supp. 2d, 907, 914 (S.D. Ohio 2007); Leguillou v. Lynch Ford, No. 99 C 3449, 2000 WL 198796, at *3 (N.D. Ill. Feb 14, 2000). Logically, however, any “right” to cancel is the dealership’s alone because it is the dealership who (falsely) advises the buyer that assignment failed. The dealer’s statement inevitably precipitates the dealer’s, not the buyer’s, cancellation (revocation) of the RISC.
147. Janikowski, 210 F.3d at 766.
148. See id. (citing Paragraph 9 of Janikowski’s Purchase Order, in which Lynch Ford asserts: “If financing cannot be obtained within 5 business days for purchaser according to the proposals in the retail installment contract executed between Seller and Purchaser…. Thus, by alleging that financing had not already been extended by the RISC, Lynch Ford is a fortiori claiming there was no creditor.).
149. See id.
150. See id.
151. See Bragg v. Bill Heard Chevrolet, Inc.-Plant City, 374 F.3d 1060, 1068 (11th Cir. 2004) (concluding that plaintiff’s “signature on these documents rendered him contractually obligated to the purchase of credit and thus constituted consummation for purposes of TILA disclosures. To give full effect to TILA’s goal of providing meaningful and timely disclosure of important credit terms, we hold that in a financing agreement containing a condition precedent where the condition
Consequently, when Janikowski signed the 5.9% RISC, both parties were legally bound.152

The Seventh Circuit did get one thing right: Lynch Ford’s TILA “disclosures” were not estimates.153 The 5.9% APR was the bait offered only to ensnare Janikowski into signing the RISC and taking delivery of the new car, and the next day to tell her financing could only be obtained at 11.9%—the switch.154 But who rejected the 5.9% financing? The court assumed without question that it must have been some third party and thereby ruled sub silentio that Lynch Ford was not the creditor in the transaction.155 The Seventh Circuit could not have come to this conclusion through reading TILA and Regulation Z, which mandate that the dealership is the creditor and is therefore bound by the TILA disclosures on the RISC.156

Regulation Z requires that the dealer be the creditor in a motor vehicle financing transaction because the dealer is the entity that extends the credit and to which the payments are initially payable.157 That is what the RISC says.158 To avoid any confusion, the Official Staff Interpretation of Regulation Z explicitly states that an assignee is not a creditor under TILA.159 Thus, in Janikowski, Lynch Ford was the creditor and was bound to the TILA disclosure requirements.160 As such, Lynch Ford was required to make the following disclosures and was bound by them:

(1) the creditor’s identity;
(2) the “amount financed”;
(3) the “finance charge”;
(4) the “annual percentage rate”;
(5) the “total of payments”;
(6) the payment schedule; and
(7) the “total sale price.”161

Indeed, Lynch Ford made these disclosures on the original RISC162 but apparently had no intention of being bound by them (as evidenced by its cancellation of the 5.9% contract).163

---

152. See id.
153. See Janikowski, 210 F.3d at 768.
154. See id.
155. See id.
162. See Janikowski v. Lynch Ford, Inc., 210 F.3d 765, 767 (7th Cir. 2000).
163. See id.
What makes Lynch Ford’s disclosures illegal as a violation of TILA is that Lynch Ford treated the terms as illusory.\footnote{164} It is irrelevant that Lynch Ford’s purchase order purported to give Janikowski the dubious right to cancel the contract if the 5.9% APR was rejected.\footnote{165} The fact that Lynch Ford claimed the right to cancel the contract rendered the disclosed terms absolutely meaningless.\footnote{166} TILA’s stated purpose is “to assure meaningful disclosure of credit terms. . . .”\footnote{167} Had the Seventh Circuit applied TILA’s purpose it surely would have arrived at the correct legal conclusion: that Lynch Ford’s practices violated TILA because the disclosures in the 5.9% RISC were truly meaningless to Janikowski.\footnote{168} It is irrelevant that Janikowski believed that some finance company was the creditor. TILA protections cannot be waived, knowingly or otherwise.\footnote{169}

Accordingly, the decision in Janikowski was precipitated by a lack of analysis of TILA, Regulation Z, and the RISC itself, as the original 5.9% RISC disclosed Lynch Ford as the creditor.\footnote{170} Once a creditor always a creditor, until assignment is legally perfected, and assignment of the 5.9% RISC was not legally perfected, which made Lynch Ford responsible under the RISC.\footnote{171} But, contrary to both the law and the evidence, the court in Janikowski erroneously believed that some third party was the creditor who rejected Janikowski’s credit.

\begin{footnotes}
\item[164] See id.
\item[165] See id. at 766.
\item[166] See Janikowski, 210 F.3d at 767-769. See also Patton v. Jeff Wyler Eastgate, Inc., 608 F. Supp. 2d 907, 915 (S.D. Ohio 2007) (explaining that a car dealership’s representation that it is extending credit on the terms disclosed within the TILA disclosure boxes is “rendered meaningless by the language in the [Spot Delivery Agreement] permitting the dealership, in essence, to cancel the purchase”).
\item[168] See Janikowski v. Lynch Ford, Inc., 210 F.3d 765 (7th Cir. 2000) (affirming summary judgment for lender). See also Parker v. DeKalb Chrysler Plymouth, 673 F.2d 1178, 1181 (11th Cir. 1982) (holding “[n]ot only does TILA contemplate a public interest in the enforcement of individual rights, but the public must rely largely on the efforts of individual consumers acting as ‘private attorneys general’ to achieve the disclosure system envisioned by the Act. If these private attorneys general are permitted to waive TILA claims in circumstances such as those presented in this case, the public interest in deterring inconsistent and undecipherable lending practices would be greatly hampered. The significance of this possibility becomes clearer when we consider that the confusion, and perhaps ignorance, that has led consumers in the past to accept draconian credit terms leaves them equally vulnerable when the lender gives them a general release of prospective claims to sign. The Truth in Lending Act was necessary exactly because borrowers could not force creditors to provide voluntarily a system of credit within which consumers could function intelligently.”) (citation omitted).
\item[169] Brooklyn Sav. Bank v. O’Neil, 324 U.S. 697, 704 (1945) (holding that “[w]here a private right is granted in the public interest to effectuate a legislative policy, waiver of a right so charged or colored with the public interest will not be allowed where it would thwart the legislative policy which it was designed to effectuate”). See also Parker, 673 F.2d at 1180-81 (applying Brooklyn Sav. Bank to TILA and spot delivery); Mills v. Home Equity Grp., 871 F. Supp. 1482, 1486 (D.D.C. 1994) (holding that rights under TILA “cannot be released or waived absent the narrowly drawn circumstances found in TILA for such waiver”).
\item[170] 210 F.3d at 766.
\end{footnotes}
TILA required the creditor (here, Lynch Ford) to disclose the actual terms of the deal, not terms that it claimed a right to cancel. TILA required Lynch Ford to disclose, for example, “[t]he ‘amount financed’, using that term, which shall be the amount of credit of which the consumer has actual use.”\(^\text{172}\) Therefore, whatever Lynch Ford disclosed as the “amount financed” on the 5.9% RISC was the only legally binding disclosure.\(^\text{173}\) It could not have retained the right to cancel the terms as such a claimed right would, and did, make all the disclosures illusory.\(^\text{174}\) Likewise, the “finance charge” disclosure in the 5.9% RISC was simply false and therefore entirely meaningless.\(^\text{175}\)

The disclosed terms were not estimates; they were the terms of the contract to which both parties should have been bound.\(^\text{176}\) By its terms, Lynch Ford was liable under the terms of the 5.9% RISC until and unless a legally valid assignment had been effected.\(^\text{177}\) Assignment never occurred, so Lynch Ford violated TILA when it revoked the contract and created a new contract at an APR six percentage points higher than the one to which it was legally bound.\(^\text{178}\) The Seventh Circuit took the only position that could have absolved Lynch Ford from liability under TILA by concluding that the two separate RISCs were both accurate and legally binding contracts, implicitly holding that the dealer’s purported right to cancel the contract was legal.\(^\text{179}\) However, a car dealership cannot retain the right to cancel a financing contract once it has been consummated, because such right would make the TILA disclosures illusory.\(^\text{180}\)

2. **Patton**—An Exercise in Statutory Analysis and Applying Congressional Intent

Analytically, the difference between *Janikowski* and *Patton* is that the district court judge in *Patton* analyzed TILA and Regulation Z and their purpose, thus arriving at the correct legal conclusion: that the dealer is the creditor and is bound by the TILA disclosures in the RISC.\(^\text{181}\) In doing so, the *Patton* court gave a reasoned, coherent explanation of why the practice of spot delivery is illegal as a violation of TILA, rather than following erroneous and unsupportable spot delivery case law.\(^\text{182}\)

\(^\text{173}\) See id.
\(^\text{174}\) See id.
\(^\text{175}\) See id.
\(^\text{176}\) Janikowski v. Lynch Ford, Inc., 210 F.3d 765, 768 (7th Cir. 2000).
\(^\text{178}\) Janikowski, 210 F.3d at 766-67.
\(^\text{179}\) Id. at 767.
\(^\text{180}\) Patton, 608 F. Supp. 2d at 913.
\(^\text{181}\) Id. at 915.
\(^\text{182}\) See id.
First, the *Patton* court correctly analyzed whether Wyler Eastgate was a creditor under TILA:183

The term “creditor” refers only to a person who both (1) regularly extends, whether in connection with loans, sales of property or services, or otherwise, consumer credit which is payable by agreement in more than four installments or for which the payment of a finance charge is or may be required, and (2) is the person to whom the debt arising from the consumer credit transaction is initially payable on the face of the evidence of indebtedness or, if there is no such evidence of indebtedness, by agreement.184

The court then noted that TILA and Regulation Z require creditors to make certain disclosures in every credit transaction with consumers.185 One of the required disclosures is the identity of the creditor.186 The court correctly found that Wyler Eastgate satisfied the second part of the definition because it held “itself out on the [RISC] as the Creditor-Seller to whom the Pattons as the Buyers were obligated to pay ‘the Amount Financed and the Finance Charge,’” thereby making the dealership “the person to whom the debt arising from the consumer credit transaction is initially payable on the face of the evidence of indebtedness.”187 Thus, the only question as to whether Wyler Eastgate was a creditor under TILA was whether it satisfied the first element of the statutory definition: whether Wyler Eastgate regularly extended credit “in more than four installments or for which the payment of a finance charge is or may be required.”188

Wyler Eastgate’s contention that it did not consider itself a creditor or that it did not itself ultimately honor terms in its RISC was undisputed.189 However, the fact that the dealership required all of its financing customers to sign a RISC that identified Wyler Eastgate as the creditor was significant to the court.190 “[The RISC] states that the Pattons, as buyers, were required to make 60 monthly payments to Wyler Eastgate, the creditor, in the sum of $493.93, which included a finance charge.”191 Despite the fact that Wyler Eastgate ultimately assigned all of its approximately 390 monthly financing deals to third parties, the language of the RISC unequivocally satisfied the second element of TILA’s definition of a “creditor.”192 Thus, “[u]ntil that [RISC] was assigned or cancelled in a

183. *Id.* at 911.
190. *Id.* at 912-13.
191. *Id.* at 913.
192. *Id.*
contractual manner, Wyler Eastgate was the entity obligated to provide financing to the Pattons on the terms provided for their purchase of the Windstar.\footnote{Id.} In other words, because the RISC made clear that the payments and finance charge were payable to the dealership, Wyler Eastgate extended credit; and because the dealership used the same or similar RISCs in most (if not all) of its 390 financing transactions per month, Wyler Eastgate extended credit regularly.\footnote{Id.} Accordingly, the dealership met TILA’s definition of a “creditor.”\footnote{Id.}

The court then correctly determined that Wyler Eastgate violated TILA by using the Spot Delivery Agreement which gave the dealership the right to cancel the contract if it decided not to assign it to a third party at the terms agreed upon.\footnote{Id. at 915.} “The representation in the [RISC] that Wyler Eastgate is providing credit on the financing terms stated in the TILA disclosures section for the Pattons’ purchase is rendered meaningless by the language in the Purchase Spot Delivery Agreement permitting the dealership, in essence, to cancel the purchase.”\footnote{Id. at 915.} Perhaps more important is the fact that the\footnote{Id. at 914-15.} Patton court rejected both Janikowski and\footnote{No. C2:05-CV-982, 2006 WL 2460872 (S.D. Ohio Aug. 12, 2006).} Gill v. Byers Chevrolet LLC,\footnote{Patton, 608 F. Supp. 2d at 915.} the latter having similar facts and coming from the same district as Patton, to the extent that those courts “found that the purchase order forms validly trumped the terms of the [RISCs] without violating TILA.”\footnote{Id. at 915-16.} Indeed, the court recognized that the purpose of TILA is to protect consumers by requiring creditors to provide meaningful disclosures, and by case law holding that car dealers could “rescind the terms of integrated automobile retail installment sales contracts by use of a second, contradictory form,”\footnote{Id. at 914.} and that a narrow interpretation of the statute would frustrate this purpose.

Patton, however, went too far in one very important respect: it attempted to distinguish Janikowski on its facts rather than acknowledge its legal shortcomings.\footnote{See supra Part III.A.1.} Patton distinguished Janikowski by noting that the purchase order in Janikowski gave both the seller and the buyer the right to cancel the contract, while Wyler Eastgate retained the sole right to cancel in Patton.\footnote{Id. at 915-16.} In fact, it is entirely irrelevant that Janikowski had the purported right to cancel the contract. As discussed above, Lynch Ford alone exercised its claimed right, which rendered TILA disclosures on the RISC illusory and thus violated TILA.\footnote{Id. at 914.} The Patton court’s attempt to distinguish Janikowski was unnecessary.
3. TILA Argument

The Truth in Lending Act makes the dealer in spot delivery financing transactions the creditor.204 There can only be one creditor in a dealer-initiated motor vehicle finance transaction.205 A potential assignee is exactly that, an “assignee.”206 The potential assignee in a spot delivery transaction is not the creditor, but is at best a buyer of the dealer’s interest in the RISC, that interest being the dealer’s right to receive the buyer’s monthly installment payments, which include a finance charge.207 The finance charge is the contract’s profit and is the inducement to the assignee to accept assignment and to pay the dealer-creditor for the RISC.

A motor vehicle RISC is a closed-end,208 secured credit contract.209 Appendix B to the ECOA’s Regulation B contains a model application form for closed-end, secured credit.210 Although creditors are permitted to change the format of the model form, a consumer credit application is to the creditor—one creditor, as there can only be one creditor.211 The dealer is that creditor and the consumer is applying for credit only with the dealer.212 The consumer is not applying for credit from any third party.213 Third parties are relevant only to the creditor for assignment purposes, not for consumer financing purposes, as the consumer has already been extended credit by way of the RISC.214

205. See, e.g., Parker v. Potter, No. 8:06-cv-183-T-26EAJ, 2008 WL 4539432, at *3 n.27 (M.D. Fla. Oct. 9, 2008), vacated on other grounds, 368 Fed. App’x 945 (11th Cir. 2010). Furthermore, TILA uses the term “creditor” in the singular in conjunction with its regulation of closed end credit, 15 U.S.C. § 1638(a)(1) (2006), and logic tells us that only one creditor can finance a car sale.
207. See id.
208. “Closed-end credit means consumer credit other than ‘open-end credit’...” 12 C.F.R. § 226.2(a)(10) (2011). “Open-end credit means consumer credit under a plan in which: (i) [t]he creditor reasonably contemplates repeated transactions; (ii) [t]he creditor may impose a finance charge from time to time on an outstanding unpaid balance; and (iii) [t]he amount of credit that may be extended to the consumer during the term of the plan (up to any limit set by the creditor) is generally made available to the extent that any outstanding balance is repaid.” Id. § 226.2(a)(20).
209. A “secured creditor” is “[a] creditor who has the right, on the debtor’s default, to proceed against collateral and apply it to the payment of the debt.” Black’s Law Dictionary 425 (9th ed. 2009).
210. 12 C.F.R. pt. 202, app. B (2011) (while Regulation B is the implementing regulation of the ECOA, Appendix B emphasizes the fact that the dealer alone is the creditor because the consumer is applying only to the dealer).
211. See, e.g., Parker v. Potter, No. 8:06-cv-183-T-26EAJ, 2008 WL 4539432, at *3 n.27 (M.D. Fla. Oct. 9, 2008), vacated on other grounds by Parker v. Potter, 368 Fed. App’x 945 (11th Cir. 2010).
213. See id.
Some of the cases that are adverse to consumers point to TILA at 15 U.S.C. § 1634 to support their decisions. This section provides that “[i]f information disclosed in accordance with this part is subsequently rendered inaccurate as the result of any act, occurrence, or agreement subsequent to the delivery of the required disclosures, the inaccuracy resulting therefrom does not constitute a violation of this part.” However, section 1634 applies only to “certain residential mortgage and variable rate transactions” refinancings, assumptions, and variable rate adjustments; and circumstances where early disclosures are rendered inaccurate prior to the date of consummation.

In closed-end motor vehicle credit contracts, the contract is consummated at buyer signing. In a residential mortgage transaction, the initial TILA disclosures are good faith estimates that may change after taxes and fees are fully determined, and it is only then that the parties consummate the mortgage contract. Thus, the plain language of TILA and Regulation Z provide that section 1634 applies only to certain mortgage transactions, and is not applicable in a spot delivery case wherein the RISC is consummated at signing.

Because only the dealer is the creditor in a spot delivery motor vehicle financing transaction, it is bound by the TILA disclosures and is required to honor those terms. When the RISC is signed, the deal is complete. The purchaser drives away with his or her car, and is bound to pay the dealership the installment payments disclosed on the RISC. If the dealership is unsuccessful in assigning the contract to a third party, it is stuck with the contract unless it opts

218. Id. § 226.20.
219. Id. § 226.17(f). See also Begala v. PNC Bank, Ohio, N.A., 163 F.3d 948, 951 (6th Cir. 1998).
222. Begala, 163 F.3d at 951.
225. Bragg, 374 F.3d at 1068.
for a recourse assignment. If the dealership calls the purchaser back to the lot to sign a new RISC, the dealer has violated TILA.

4. Groupthink

One likely reason many courts erroneously conclude that TILA permits spot delivery using a buyer-signed RISC is that they engage in “groupthink”: “a mode of thinking that people engage in when they are deeply involved in a cohesive in-group, when the members’ strivings for unanimity override their motivation to realistically appraise alternative courses of action.” Sociologists have tied groupthink to several major debacles in modern history, including the Bay of Pigs fiasco, the escalation of the Korean War, the failure to prevent the Pearl Harbor bombing, and the escalation of the Vietnam War. Groupthink has been popularly accepted since the 1972 publication of Irving Janis’s groundbreaking book, Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes. According to one report, on average social scientists cite the 1982 version of Janis’s book over 100 times per year. While researchers generally address groupthink in relation to political decision-making, the theory is equally applicable to judicial decision-making.

The symptoms of groupthink particularly relevant to judicial decisions include: (1) “the illusion of invulnerability,” (2) “an unquestioned belief in the group’s inherent morality,” (3) “the illusion of unanimity,” (4) “suppression

226. “Recourse” is “the right of a holder of a negotiable instrument to demand payment from the drawer or indorser if the instrument is dishonored.” BLACK’S LAW DICTIONARY 1389 (9th ed. 2009). See also “recourse” discussion, infra notes 313-17 and accompanying text.

227. See Patton, 608 F. Supp. 2d at 913.

228. In this section we intend to offer an explanation of why so many courts have come to erroneous conclusions on what is a rather straightforward violation of TILA. The “group” in groupthink often includes the consumers’ attorneys who also believe that dealers are not creditors.


230. See id. at 14.

231. See id. at 50.

232. See id. at 75.

233. Id. at 101.


236. See Oleson, supra note 9.

237. JANIS, supra note 229, at 36.

238. Id. at 198.

239. Id. at 38.
of personal doubts,”240 and (5) the rationalization of poor decisions.241 However, there are solutions that can mitigate the likelihood that groupthink will occur, such as encouraging group members to openly voice objections242 and examining the views of outside, independent experts.243 Unfortunately, many courts addressing the spot delivery issue have displayed symptoms of groupthink, while employing none of its solutions.244

a. The Illusion of Invulnerability

The “illusion of invulnerability . . . creates excessive optimism” and often leads to poor decision-making.245 Janis describes this symptom as the attitude that if “everyone else in our group decides that it is okay, the plan is bound to succeed.”246 As applied to courts and spot delivery, this illustration is better understood as “if other courts say it is okay, they must be correct.” This attitude leads to a virtual vacuum of critical legal analysis.247 As a result, those courts rarely, if ever, seriously consider the alternative position—that the other courts might have gotten it wrong.248 The illusion of invulnerability creates mysticism and leads courts to be uncritical of their counterparts’ decisions.249 Indeed, it often leads courts to reject the opinions of expert witnesses whose testimonies run counter to the judges’ preconceived beliefs.250 Presented with two opposing perspectives—one from an expert witness, and another a judicial opinion—too many judges will side with the judicial opinion.251 However, judges could

240. Id. at 40.
241. Id. at 198.
242. Id. at 180.
243. JANIS, supra note 229, at 214.
245. JANIS, supra note 229, at 197.
246. Id. at 36.
247. See Gill v. Byers Chevrolet LLC, No. C2:05-CV-982, 2006 WL 2460872, at *7-8 (S.D. Ohio 2006) (relying entirely on Janikowski v. Lynch Ford, 210 F.3d 765 (7th Cir. 2000) and Clark v. Troy and Nichols, Inc., 864 F.2d 1261 (5th Cir. 1989) to come to its conclusion that spot delivery does not violate TILA: “The consensus is in favor of holding that this type of practice does not constitute a violation of TILA.” See also Hunter, 2008 WL 1925265, at *5 (relying entirely on erroneous case law in four sentence discussion of plaintiffs’ FCRA claims without any analysis of or citation to the actual statute).
248. But see Gill, 2006 WL 2460872, at *7 (citing Rossman v. Fleet Bank (R.I.) Nat’l Assoc., 280 F.3d 384, 398 (3d Cir. 2002)) for a court that, in dicta, raised the possibility that the dissent in Clark, 864 F.2d at 1265, which argued that the bait and switch aspect of spot delivery was a violation of TILA, had the better argument.
250. See Janikowski, 210 F.3d at 769.
251. See, e.g., id.
largely avoid this “choice” if they were to read and understand TILA, Regulation Z, and the RISCs.\footnote{252} But many judges rely instead on the collective wisdom of sister courts. “Judges are susceptible to . . . the belief that, being pure, the judicial mission cannot fail.”\footnote{253}

b. Unquestioned Belief in the Group’s Morality

Related to the illusion of invulnerability is the “unquestioned belief in the group’s morality, inclining the members to ignore the ethical or moral consequences of their decisions . . . .”\footnote{254} This aspect of groupthink is particularly relevant to the spot delivery issue. When a court concludes that spot delivery does not violate TILA, it essentially tells car dealers that they are free to use unfair and deceptive acts and practices to defraud consumers.\footnote{255} One would search in vain for a judicial opinion both holding spot delivery to be permissible under TILA and also considering the consequences of that holding.\footnote{256} Rather, the opinions rely heavily, if not exclusively, on erroneous authority from other courts.\footnote{257} They incorrectly assume that the courts rendering the decisions considered the moral and ethical consequences, and follow them without analysis.\footnote{258}

c. The Illusion of Unanimity

Perhaps some courts ignore the moral and ethical consequences of their spot delivery decisions in an effort to achieve unanimity with other courts.\footnote{259} This oversight results from “a shared unanimity concerning judgments conforming to the majority view (partly resulting from self-censorship of deviations, augmented

\footnote{253. Oleson, supra note 9, at 689.}
\footnote{254. JANIS, supra note 229, at 198.}
\footnote{255. See, e.g., Janikowski, 210 F.3d at 767-68.}
\footnote{256. Compare id. at 769 (implicitly holding conditional sale spot delivery to be a legal sales practice without mentioning how that holding essentially means that car dealers can misrepresent the financing terms of the sale) with Bragg v. Bill Heard Chevrolet, Inc.-Plant City, 374 F.3d 1060, 1067-68 (11th Cir. 2004) (concluding that “[d]isclosures that come after the consumer executes a RISC . . . are likely to be of little or no value to that consumer”), and Daenzer v. Wayland Ford, Inc., No. 1:01CV133, 2002 WL 1050209, at *3 (E.D. Mich. 2002) (reasoning that the reason defendant car dealer refused to give a copy of the first signed RISC to the consumer was to give the dealer “the de facto ability to force the customer to sign a second contract changing the original terms of the bargain by depriving the customer of a RISC with the original terms.”).}
\footnote{257. See, e.g., Gill v. Byers Chevrolet LLC, No. C2:05-CV-982, 2006 WL 2460872, at *7-8 (S.D. Ohio 2006) (relying entirely on Janikowski, 210 F.3d 765 and Clark v. Troy and Nichols, Inc., 864 F.2d 1261 (5th Cir. 1989) to come to its conclusion that spot delivery does not violate TILA, rather than analyzing TILA’s stated purpose and the considering the consequences of such a holding: “The consensus is in favor of holding that this type of practice does not constitute a violation of TILA.”).}
\footnote{258. See Gill, 2006 WL 2460872, at *7-8.}
\footnote{259. See id. at *8.}
by the false assumption that silence means consent)... Courts are comprised of impartial judges who hear both sides of a dispute and resolve it, if possible, on legal grounds. Almost all spot delivery cases are decided at the motion to dismiss and the summary judgment stages because neither party is disputing the facts; they are instead disputing whether what they agree happened is legal or illegal. And because the spot delivery procedure is virtually the same at dealerships across the country, most cases have essentially identical facts. For this reason, courts presented with spot delivery issues tend to rely on erroneous case law, perhaps because when the same facts exist in previously decided cases, courts strive to achieve unanimity. It is this “concur rence-seeking” behavior that makes groupthink so dangerous. It encourages “judicial sophistry”: “a subtle, tricky, superficially plausible, but generally fallacious method of reasoning.”

d. Suppression of Personal Doubts

Similar to the illusion of unanimity is the “self-censorship of deviations from the apparent group consensus, reflecting each member’s inclination to minimize to himself the importance of his doubts and counterarguments.” Dissenting opinions are common in our judicial system, but it takes an independent judge to voice reservations about a practice that has been accepted as the norm for many years, such as spot delivery using a buyer-signed RISC. Janis discusses this phenomenon in relation to the Bay of Pigs invasion, describing how many of President Kennedy’s advisors, who were participants of the discussions leading

260. JANIS, supra note 229, at 198.
263. See Gill, 2006 WL 2460872, at *7-8 (relying entirely on Janikowski, 210 F.3d 765 and Clark v. Troy and Nichols, Inc., 864 F.2d 1261 (5th Cir. 1989) to come to its conclusion that spot delivery does not violate TILA: “The consensus is in favor of holding that this type of practice does not constitute a violation of TILA.”
264. JANIS, supra note 229, at 11.
266. JANIS, supra note 229, at 198.
267. See generally Stergiopoulos v. First Midwest Bancorp, 427 F.3d 1043, 1044 (7th Cir. 2005) (acknowledging the “common scenario” of spot delivery in FCRA case, and noting its “routine nature.”); Geller v. Onyx Acceptance Corp., No. 728614, 2001 WL 1711313, at *2 (Cal. Super. Ct. Nov. 13, 2001) (discussing at length, and relying in large part on, the widespread practice of third-party assignees secretly paying kickbacks to car dealers in spot delivery transactions to conclude that they are not unfair and deceptive: “Dealer participation payments have been an integral part of the indirect auto finance market since at least the 1960’s.”).
up to the fiasco, suppressed their own ideas because they “felt reluctant to raise questions that might cast doubt on a plan that they thought was accepted by the consensus of the group, for fear of evoking disapproval from their associates.”

It is arguably this tendency that leads to some erroneous judicial opinions. It is arguably this same tendency that led the Patton court to factually distinguish Janikowski rather than reject it completely. The Patton court should have expressed its doubts and voiced an objection to the Janikowski reasoning in addition to rejecting its conclusion.

e. Rationalizing Poor Decisions

Janis describes rationalizing poor decisions as the “collective efforts to rationalize in order to discount warnings which might lead the members to reconsider their assumptions before they recommit themselves to their past policy decisions.” In the spot delivery context, these warnings come from consumers’ attorneys and from reading TILA’s declaration of purpose. However, courts continue to recommit to the collective “wisdom” of erroneous case law. They rationalize their opinions in order to conform. For example, the court in Janikowski rationalized its decision by concluding that the plaintiff had signed two separate, legally binding contracts. This conclusion makes no sense; when Ms. Janikowski signed the 5.9% RISC it was a legally complete contract to which both parties were clearly bound. Lynch Ford then deceived Janikowski into signing a new RISC by misrepresenting to her that the bank rejected the 5.9% financing. But rather than “reconsider[ing] [its]

268. Janis, supra note 229, at 40-41.
270. Janis, supra note 229, at 198.
271. 15 U.S.C. § 1601 (2006) (“It is the purpose of this subchapter to assure a meaningful disclosure of credit terms so that the consumer will be able to compare more readily the various credit terms available to him and avoid the uninformed use of credit, and to protect the consumer against inaccurate and unfair billing and credit card practices.”).
273. 210 F.3d at 767.
274. See Bragg v. Bill Heard Chevrolet, Inc.-Plant City, 374 F.3d 1060, 1068 (11th Cir. 2004) (concluding that plaintiff’s “signature on these documents rendered him contractually obligated to the purchase of credit and thus constituted consummation for purposes of TILA disclosures. To give full effect to TILA’s goal of providing meaningful and timely disclosure of important credit terms, we hold that in a financing agreement containing a condition precedent where the condition of obtaining financing is within the exclusive control of the seller and third-party lender, consummation occurs when the consumer signs the contract.”).
275. Janikowski, 210 F.3d at 766.
assumption that spot delivery is legal, the Seventh Circuit Court of Appeals instead rationalized a rather esoteric (and erroneous) conclusion.

f. Groupthink Conclusion

Groupthink often results in bad case law. When judges succumb to the illusions of invulnerability and unanimity, they suppress their personal doubts and rationalize poor decisions.

The reification of the law as something distinct from judges allows them to maintain a sense of superior moral objectivity. By treating themselves as humble servants of the law (a law that is entirely separate from themselves), judges can indulge in a mythic illusion, telling themselves that they are but impartial functionaries of the law. They can reassure themselves that they are enforcing rules that have emerged from democratic compromises within a pluralistic society, thereby avoiding the voicing of support for one normative faction over another. So pretending, believing that they live in a world of legal abstractions (and not a world of human actors), judges can become blind to the consequences of their decisions.

Thus, judges addressing spot delivery issues need to thoroughly read and understand TILA, Regulation Z, and the RISCs themselves. They need to carefully consider the testimony of experts and always interpret TILA and Regulation Z with their consumer protection purpose in mind. Finally, judges should consider the ethical consequences of permitting unfair and deceptive sales practices to continue.

5. TILA Conclusion

Combining a buyer-signed RISC with a creditor disclaimer always violates the Truth in Lending Act. In every dealer-initiated, buyer-signed RISC, the dealer alone is required by TILA to disclose the TILA terms. One of these

---

276. JANIS, supra note 229, at 198.
277. Janikowski, 210 F.3d at 767.
278. See, e.g., Plessy v. Ferguson, 163 U.S. 537 (1896).
279. See id.
280. Oleson, supra note 9, at 689.
283. See JANIS, supra note 229, at 197.
284. Spot delivery would not be a violation of TILA if the purchaser did not sign a RISC, because it is the modifications of the mutually binding TILA disclosures on the RISC, or purporting to claim a right to modify them, that violates the statute. Rather, dealerships could legally spot deliver by having the purchaser sign only a buyer’s order explaining that the deal is contingent on third-party financing.
disclosures is the identity of the creditor, which is always the dealer itself. If the RISC form does not expressly identify the dealer as the creditor, it is a violation of TILA. Typically, the dealer will include a separate form (or provision within the RISC itself) that purports to condition the RISC on its ability to assign the contract. Whether the form or provision gives either party, or only the dealer, the right to cancel the contract, such claimed right renders the disclosures illusory and thus violates TILA, which mandates that retail installment contracts, when signed by the buyer, are the final written expression of the parties and are binding. Car dealers use conditional financing provisions to relieve themselves of their statutory and contractual obligations and to waive the buyers’ statutory and contractual rights, because such provisions either expressly or implicitly purport that some third party is the creditor, not the dealer.

Dealerships violate TILA and Regulation Z in their credit sales each time they purport to condition the contract terms on subsequent assignment. This is necessarily so because at the time of consummation as defined by TILA, the disclosures are subject to the dealer’s unilateral modification or revocation. The practice of spot delivery, therefore, is centered on disclosure of meaningless credit terms, while the intent of TILA is to provide meaningful credit terms. Spot delivery using a buyer-signed RISC cannot coexist with the Truth in Lending Act.

B. Fair Credit Reporting Act: An Analysis

Spot delivery practices also violate the Fair Credit Reporting Act every time a car dealer pulls a consumer’s credit report with no intention of itself extending the credit. The only permissible purpose for accessing customer credit reports is as a creditor extending credit to the consumer, which is never the case unless the dealer is an admitted “buy here, pay here” lot. However, the dealer could

289. See Patton, 608 F. Supp. 2d at 916. See also Bragg v. Bill Heard Chevrolet, Inc.-Plant City, 374 F.3d 1060, 1068 (11th Cir. 2004) (concluding that the contract was consummated, and therefore binding, when the consumer signed the RISC).
291. See Bragg, 374 F.3d at 1068; 12 C.F.R. § 226.2(a)(13).
294. Id. “Buy here, pay here” lots are automotive dealers that advertise that they finance their own transactions, and do not condition their sales on third-party financing. They actually service their own car loans as accounts receivable, i.e., the consumer pays, and the dealer accepts, the monthly installment payments disclosed in the RISC. Such dealers typically do not assign their RISCs.
cure this by entering into a recourse agreement\textsuperscript{295} with the assignee—an option that typically appears in the “Assignment” section of the RISC.\textsuperscript{296}

When requesting a consumer’s credit report from a credit reporting agency, the dealer must certify that it “intends to use the information in connection with a credit transaction involving the consumer on whom the information is to be furnished and involving the extension of credit to, or review or collection of an account of, the consumer . . . .”\textsuperscript{297} In other words, unless the dealership already has a credit account with the consumer, the only reason it can legally request the consumer’s report is with the intention of itself extending credit to the consumer.\textsuperscript{298} As discussed above, this is never the case with spot delivery.\textsuperscript{299} Indeed, the collateral documents or provisions purporting to condition financing on later assignment tell the consumer, and courts, that the dealership does not intend to honor the credit terms.\textsuperscript{300} But when the dealer prepares and the consumer signs a RISC, the dealer has extended credit.\textsuperscript{301} The RISC is an unambiguous, fully integrated, legally binding contract.\textsuperscript{302} When the dealer obtains the consumer’s credit report and then completes a RISC that it has no intention of honoring, it is a violation of the FCRA because its accessing of the customer’s credit report is not to extend credit; rather, it is to determine which potential assignee is most beneficial to the dealer in terms of profit.\textsuperscript{303}

Take the case of Jane Jones, the protagonist in our hypothetical.\textsuperscript{304} Dave’s Dealership conducted an initial credit check before extending Jane the 6\% APR RISC. The FCRA permits this only if Dave’s Dealership intended to honor the terms of the deal, i.e., only if Dave’s Dealership intended to provide the financing for Jane’s purchase at 6\%.\textsuperscript{305} However, because the Dealership had Jane sign a spot delivery disclaimer that purported to condition the deal on the Dealership’s desire to assign the contract as written, the dealership violated the FCRA because it did not intend to extend Jane the credit.\textsuperscript{306} Rather, Dave’s Dealership pulled Jane’s credit to determine which potential assignee would

\begin{footnotesize}
\begin{enumerate}
\item[295.] Recourse is “the right of a holder of a negotiable instrument to demand payment from the drawer and indorsers when the first liable party fails to pay.” Bryan A. Garner, A Dictionary of Modern Legal Usage 742 (2d. ed. 2001).
\item[298.] Id.
\item[299.] See, e.g., Patton v. Jeff Wyler Eastgate, Inc., 608 F. Supp. 2d 907, 912 (S.D. Ohio 2007) (reciting Dealership’s chief financial officer’s testimony that “Wyler Eastgate does not ‘as a rule . . . finance its own retail sales.’”).
\item[300.] See 12 C.F.R. pt. 226, app. A.
\item[301.] Cf. Bragg v. Bill Heard Chevrolet, Inc.-Plant City, 374 F.3d 1060, 1068 (11th Cir. 2004) (concluding that consummation of the transaction occurs at buyer signing, thus holding by implication that the terms in the TILA disclosure boxes are binding on both parties).
\item[302.] See 12 C.F.R. pt. 226, app. A.
\item[304.] See supra Part I.
\item[306.] Id.
\end{enumerate}
\end{footnotesize}
likely buy the RISC as written. This is not a permissible purpose under the FCRA, and thus Dave’s Dealership has violated that Act.  

So too have the potential assignees from whom the Dealership seeks to assign Jane’s RISC because they are assignees, not creditors, and Jane did not authorize any third party to access her credit report. In almost every spot delivery transaction, as in Jane’s case, the dealer contacts several potential assignees, who in turn access the consumer’s credit report for the impermissible purpose of determining whether or not to accept assignment of the dealer’s RISC, not to extend credit to the consumer. 

Accordingly, Jane can make out a prima facie case for an FCRA violation because: (1) there was a consumer report; (2) the dealer obtained it; and (3) the dealer did not have a permissible purpose. Noncompliance with the FCRA in Jane’s case was at best negligent, and more likely willful. If merely negligent, Dave’s Dealership would be liable to Jane for actual damages and attorney’s fees. 

The judge or jury finds willful noncompliance, then the Dealership is liable for actual damages, punitive damages, court costs, and attorney’s fees.

Spot delivering dealers are not “arranging” credit for consumers. They seek a purchaser of their interest in the RISC at face value. Even if an assignee refuses to buy the dealer’s RISC as written, the dealer can still assign it by agreeing to “limited recourse” or “full recourse” assignment. This would assure assignment of all RISCS which, in turn, would eliminate spot delivery/yo-yo transactions. In the event that the consumer misses an installment, the

307. See id.

308. See id. § 1681b(f) (stating that “[a] person shall not use or obtain a consumer report for any purpose unless (1) the consumer report is obtained for a purpose for which the consumer report is authorized to be furnished under this section; and (2) the purpose is certified in accordance with section 1681e of this title by a prospective user of the report through a general or specific certification.” Because determining whether or not to accept assignment is not a permissible purpose, the potential assignees may not lawfully access the purchasers’ credit.).


312. Id. § 1681n.

313. See id. § 1691a(e).

314. See generally ROGER D. BILLINGS, JR., HANDLING AUTOMOBILE WARRANTY AND REPOSESSION CASES §§ 12.3, 12.5 (2d ed. 2003). “Limited Recourse” is an agreement between the creditor-dealer and an assignee which provides that in the event that a borrower (e.g., a person financing a car) misses a payment, the car dealer will make that payment to the assignee, or will buy the loan back from the assignee if the missed payment occurs, say, within the first three months.

315. See generally id. “Full Recourse” is an agreement between the creditor-dealer and an assignee which provides that in the event the consumer defaults on the loan, the car dealer will repurchase the contract from the assignee.
assignee could look to the dealer for payment, or require the dealer to repurchase the loan. Thus, when a dealer tells a consumer that a bank, finance company, or credit union denied or rejected credit, it is a false statement made only because the dealer has chosen to defraud the consumer rather than agree to a recourse assignment.\footnote{316. The dealer always has the option of choosing “full recourse,” “limited recourse,” or “without recourse” on assignment, and virtually always chooses “without recourse.” See, e.g., 12 C.F.R. pt. 226, app. A (2011). “Without recourse” is defined as “a formula used to disclaim responsibility for future nonpayment, esp. of a negotiable financial instrument.” The New Oxford American Dictionary 1424 (2001). See also Black’s Law Dictionary 1740 (9th ed. 2009) (defining “without recourse” as “…without liability to subsequent holders. With this stipulation, one who indorses an instrument indicates that he or she has no further liability to any subsequent holder for payment.”).}

C. Equal Credit Opportunity Act—An Analysis

The Equal Credit Opportunity Act and Regulation B require a creditor to provide its financing customers with a written adverse action notice explaining the reasons for denying or revoking the extension of credit.\footnote{317. 15 U.S.C. § 1691(d) (2006); 12 C.F.R. § 202.9(a) (2011).} Regulation B defines “adverse action” in a credit transaction as “revocation, alteration, or termination” of credit.\footnote{318. 12 C.F.R. § 202.2(m). While section 202.2(c) is the “adverse action” subsection, section 202.2(m) states it more concisely.} Car dealers who prepare RISCs are creditors under the ECOA definition, i.e., “any person who regularly extends, renews or continues credit; any person who regularly arranges for the extension, renewal, or continuation of credit; or any assignee of an original creditor who participates in the decision to extend, renew, or continue credit.”\footnote{319. 15 U.S.C § 1691a(e).} As noted above in the TILA analysis, car dealers regularly extend credit.\footnote{320. See, e.g., Patton v. Jeff Wyler Eastgate, Inc., 608 F. Supp. 2d 907, 913 (S.D. Ohio 2007).} Spot delivering dealers do not “arrange for the extension” of credit.\footnote{321. See 15 U.S.C. § 1691a(e).} The RISC itself is the extension of credit.\footnote{322. See 15 U.S.C. § 1638(a) (2006); 12 C.F.R. § 226.18 (2011). See also 12 C.F.R. pt. 226, app. A (2011).} The ECOA requires dealers who revoke a buyer-signed RISC (e.g., Lynch Ford and Wyler Eastgate) to issue a written adverse action notice explaining the reasons for the adverse action.\footnote{323. See 15 U.S.C. § 1691(d); 12 C.F.R. § 202.9(a). See also Treadway v. Gateway Chevrolet Oldsmobile, Inc., 362 F.3d 971, 975-76 (7th Cir. 2004) (suffering the same infirmities as Janikowski v. Lynch Ford, Inc., 210 F.3d 765 (7th Cir. 2000), as the panel assumed that the dealer was not the creditor and that any third party in the circumstance was the creditor (the “lender”). “By unilaterally deciding not to send Treadway’s application to any lender, Gateway effectively denied credit to Treadway. Whether it is the lender or the dealership that makes the decision, both the action and the outcome are the same.” The lender is the dealership; there was no reason to differentiate).} Furthermore, car dealers often attempt to assign financing transactions to multiple third parties, presumably choosing the one assignee on each deal that
provides some particular benefit to the dealer.\textsuperscript{324} For example, in our Jane Jones hypothetical, Dave’s Dealership transmitted Jane’s credit application to nine banks and finance companies.\textsuperscript{325} The transmittals were not to obtain financing for Jane, but were to locate the most advantageous buyer for Dave’s RISC. When all nine potential assignees rejected Dave’s assignment requests, Dave’s revoked Jane’s RISC which was the only adverse action on Jane’s credit application.\textsuperscript{326} The ECOA required Dave’s Dealership to issue a written adverse action notice, but Dave’s did not do so.\textsuperscript{327} Thus, Dave’s Dealership has violated the ECOA.\textsuperscript{328}

D. Unfair Deceptive Acts and Practices—An Analysis

Spot delivery, because it violates TILA, the FCRA, and the ECOA, is an unfair and deceptive practice as a matter of law.\textsuperscript{329} A sales practice cannot be a violation of three federal consumer-protection statutes and still be fair to the consumer. Spot delivery is used to take the consumer out of the car-buying market.\textsuperscript{330} It is anti-competition, as the buyer believes he or she has a completed deal and will not shop further. Spot delivering dealers may well offer financing at terms they have no intention of assigning, just to put the buyer in a car that he

\begin{footnotesize}
\begin{enumerate}
\item See, e.g., Stergiopoulos v. First Midwest Bancorp, Inc., 427 F.3d 1043, 1044 (7th Cir. 2005).
\item See supra Part I.
\item See 12 C.F.R. § 202.2(m).
\item 15 U.S.C. § 1691(d).
\item 15 U.S.C. § 1691(d); 12 C.F.R. § 202.9(a).
\begin{enumerate}
\item Q: But it’s the dealership’s policy to spot deliver if at all possible in order to take the consumer out of the car buying market, correct?
\item A: Yes.
\item Q: And that’s really the philosophy of all the dealerships that you’ve just told me about, isn’t it?
\item A: I would say yes.
\end{enumerate}
\end{enumerate}
\end{footnotesize}
or she will fall in love with and proudly show to family, friends, and co-
workers. The reasoning is that when the buyer is called back to the dealership
days or even weeks later, the buyer is far more likely to acquiesce to the new,
more expensive financing terms. This is classic bait and switch.

Such bait and switch is by definition unfair and deceptive because it amounts
to theft. Without realizing the consequences of their decisions, many courts
have found that car transactions amounting to bait and switch are not violations
of state UDAP statutes, perhaps based on state statutes that permit spot
delivery. This logic is patently wrong. Any state law providing that spot
delivery is permitted is necessarily inconsistent with TILA because such laws
assume that dealers are not themselves creditors. Pursuant to the Supremacy
Clause in Article VI of the United States Constitution, federal law is the
“supreme Law of the Land; and the Judges in every State shall be bound thereby,
any Thing in the Constitution or Laws of any State to the Contrary
notwithstanding.” Thus, any state law that provides that spot delivery using a
buyer-signed RISC is legal is unconstitutional as a violation of TILA and the
Supremacy Clause.

Assume Dave’s Dealership from the hypothetical is in Ohio, and the series
of events in Jane Jones’s transaction all occurred in Ohio. Dave’s Dealership
would have violated the Ohio Consumer Sales Practice Act (“OCSPA”) as many
as three times. The OCSPA declares that “no supplier shall commit an
unconscionable act or practice in connection with a consumer transaction. Such
an unconscionable act or practice by a supplier violates this section whether it
occurs before, during or after the transaction.” One indicator of
unconscionability is “whether the supplier has knowingly taken advantage of
the inability of the consumer reasonably to protect the consumer’s interests

331. See Treadway v. Gateway Chevrolet Oldsmobile Inc., 362 F.3d 971, 973 (7th Cir. 2004).
332. See id.
333. See infra Part III.E.
334. See, e.g., Chastain v. N.S.S. Acquisition Corp., No. 08-81260-CIV, slip op., 2009 WL
1971621, at *8 (S.D. Fla. July 8, 2009); King v. King Motor Co. of Ft. Lauderdale, 900 So. 2d 619,
335. U.S. CONST. art. VI, cl. 2.
336. Id.
337. See, e.g., ARIZ. REV. STAT. ANN. § 44-1371 (2010); ARK. CODE ANN. § 23-112-316 (West
2010); COLO. REV. STAT. ANN. § 6-1-708 (West 2010); NEV. REV. STAT. ANN. § 482.554 (West
2010); N.H. REV. STAT. ANN. § 361-A: 10b (2010); UTAH CODE ANN. § 41-3-401 (West 2010); VA.
CODE ANN. § 46.2-1530 (West 2010); IDAHO ADMIN. CODE r.04.02.01.237 (2010). These laws
conditionally legalize spot delivery in the erroneous belief that they protect consumers, but are
preempted by TILA, the FCRA, and the ECOA, and are therefore unconstitutional. See also
2009) (holding that plaintiff’s state law UDAP claims were preempted by TILA).
338. U.S. CONST. art. VI, cl. 2.
339. See supra Part I.
340. See supra Part I.
341. Ohio Consumer Sales Practices Act, OHIO REV. CODE ANN. § 1345.03(A) (West 2010).
342. Id.
because of the consumer’s physical or mental infirmities, ignorance, illiteracy, or inability to understand the language of an agreement."345 Dave’s would have violated the OCSPA before the transaction when it represented to Jane in the RISC that it was the creditor, but purported to make her transaction contingent upon assignment. This takes advantage of Jane’s ignorance of her TILA right to meaningful credit terms.344 Dave’s would also have violated the OCSPA during the transaction when it obtained her signature on the “Seller’s Right to Cancel,” because no such “right” exists. Any such claim of that “right” is nothing other than attempted avoidance of the dealer’s TILA and RISC obligations.345 Finally, Dave’s would have violated the OCSPA after the transaction when it required her to sign a second RISC because it took advantage of Jane’s ignorance of the Dealership’s TILA requirements; she could not protect her own interests.346 Moreover, the fact that the Dealership violated the FCRA and the ECOA (in addition to TILA) would have rendered the transaction unconscionable.347

E. State Criminal Law—An Analysis

If a car dealer offers a consumer certain financing terms and the consumer accepts those terms and signs the contract, the transaction is complete.348 The RISC is legally binding on both parties.349 If the dealer subsequently revokes that contract and replaces it with a second, more expensive RISC and obtains the buyer’s signature on it, the dealer has committed theft—an offense codified in every state’s criminal code.350

For example, in Janikowski, Lynch Ford’s illegal profit from its six-point interest increase (5.9% to 11.9% APR), assuming an amount financed of $15,000 over forty-eight months, was $2,050.351 In Illinois, “[a] person commits theft when he knowingly: (2) [o]btains by deception control over property of the owner.”352 This is precisely what the “Seller’s Right to Cancel” provision is too often used for in financing deals: “If Seller elects to cancel, you may have the option of negotiating and signing a new contract with different financing terms,

343. Id. § 1345.03(B).
344. This is necessarily so because Dave’s deceived Jane into believing that a dealership has a legal right to condition the sale on third-party financing, which it does not because it would render the TILA disclosures illusory.
345. See Bragg v. Bill Heard Chevrolet, Inc.-Plant City, 374 F.3d 1060, 1068 (11th Cir. 2004).
348. See Bragg, 374 F.3d at 1068.
349. See id.
e.g., a larger down payment, a higher annual percentage rate . . . .”

Essentially, what this means is that if on attempted assignment the dealer cannot maintain the profit factored into the contract, it will revoke the RISC in favor of a second RISC to either maintain its profit margin or to increase it. The second RISC, resulting in either a larger down payment or a higher APR, is “obtain[ing] by deception control over property of the owner.”

The “property” is Janikowski’s money. When Lynch Ford represented to her that she could buy the Ford Escort at 5.9% APR, it induced her to sign the RISC notwithstanding its purported “right” to cancel. When Lynch Ford called her back to the dealership to tell her that credit was not available at 5.9% APR, it deceived her into believing that Lynch Ford was not the creditor and that she could buy the car only at an APR of 11.9%. This is entirely false for two reasons. First, Lynch Ford was the creditor at all times, and it was therefore bound by the first RISC and was required to honor the 5.9% APR. Second, Lynch Ford still could have assigned the RISC, either at a lower profit margin or by agreeing to full or limited recourse. In other words, it could have sold the 5.9% RISC by opting for recourse assignment. Instead, it deceived Janikowski into the new, more expensive terms which were more profitable to the dealership.

IV. CONCLUSION

The car dealer practice of spot delivery using a buyer-signed RISC is illegal as a violation of the Truth in Lending Act, the Fair Credit Reporting Act, the Equal Credit Opportunity Act, most state Unfair and Deceptive Act and Practices statutes, and often constitutes criminal theft. Astoundingly, most courts addressing the issue have concluded that this practice is legal. Given that federal and state consumer protection statutes, if properly construed, prohibit spot delivery using a buyer-signed RISC, it is not surprising that courts holding it to be a legal sales practice give reflexive lip service to TILA while offering faulty justification for their conclusion. State laws purporting to regulate spot

356. Id.
358. See supra notes 295, 314--15 and accompanying text.
delivery are unconstitutional as violations of TILA and the Supremacy Clause of Article VI of the United States Constitution.\textsuperscript{362} Courts misconstruing TILA have relied on the decisions of other courts to justify their opinions. Most notable among these are the progeny of \textit{Janikowski v. Lynch Ford}, which are decisions that offer no valid legal justification.\textsuperscript{363} This phenomenon is caused by “groupthink.”\textsuperscript{364} Like individuals, courts are often predisposed to concurrence-seeking behavior.\textsuperscript{365} They do not want to be considered non-conforming with the prevailing jurisprudence, even when that jurisprudence is legally erroneous and logically indefensible.\textsuperscript{366} Encouragingly, some courts have rejected the view that spot delivery is legal.\textsuperscript{367} Cases such as \textit{Patton v. Wyler Eastgate, Inc.}, with sound legal analysis and reasoning are bound to eventually prevail over the erroneous, but currently dominant view.\textsuperscript{368} Consumer protection depends on it.

\begin{footnotesize}
\footnotesize
\textsuperscript{362.} See, \textit{e.g.}, ARIZ. REV. STAT. ANN. § 44-1371 (2010); ARK. CODE ANN. § 23-112-316 (West 2010); COLO. REV. STAT. ANN. § 6-1-708 (West 2010); NEV. REV. STAT. ANN. § 482.554 (West 2010); N.H. REV. STAT. ANN. § 361-A:10b (2010); UTAH CODE ANN. § 41-3-401 (West 2010); VA. CODE ANN. § 46.2-1530 (West 2010); IDAHO ADMIN. CODE r.04.02.01.237 (2010). \textit{See, e.g.}, Kauinui v. Citibank, N.A., No. 09-000258 ACK-BMK, 2009 WL 3530373, at *1 (D. Haw. Oct. 28, 2009) (holding that plaintiff’s state law UDAP claims were preempted by TILA).

\textsuperscript{363.} See, \textit{e.g.}, \textit{Gill}, 2006 WL 2460872, at *8 (concluding that spot delivery must be legal because “[t]he consensus is in favor of holding that this type of practice does not constitute a violation of TILA.”).

\textsuperscript{364.} \textit{See Janis}, supra note 229; Oleson, supra note 9.

\textsuperscript{365.} \textit{See Oleson}, supra note 9, at 688-90.

\textsuperscript{366.} \textit{See id}.


\textsuperscript{368.} 608 F. Supp. 2d 907.
\end{footnotesize}
FRONT PAY UNDER THE FMLA

Michelle Eviston

I. INTRODUCTION

When an employer fails to comply with the Family and Medical Leave Act of 1993 ("FMLA"), the employee may be entitled to a number of remedies, including "front pay." Front pay is money awarded to compensate a plaintiff for future earnings that he or she would have earned if the employer had not violated the FMLA. Although all circuits agree that a judge must determine when a front pay award is available, the circuits disagree as to who should determine the amount of the award. Almost all of the circuit courts have determined that only a judge may calculate the amount of a front pay award; yet, the Sixth Circuit has said that a jury must determine the amount.

Under the Family and Medical Leave Act of 1993, covered employees may take unpaid leave so that they can care for themselves or immediate family without fearing job loss. If an employer violates the FMLA, the statute provides

---

* Michelle Eviston graduated from Northern Kentucky University in 2008 with a B.A. in Criminal Justice and a B.A. in Political Science. She received her J.D. from Salmon P. Chase College of Law in 2011.

2. See infra Part II. A.
4. Compare Traxler v. Multnomah County, 596 F.3d 1007, 1014 (9th Cir. 2010) (finding that the court determines the availability and amount of the front pay award under the FMLA) with Arban v. West Publ’g Corp., 345 F.3d 390, 406 (6th Cir. 2003) (finding that while the court determines the availability of the front pay award under the FMLA, the jury determines the amount of the award).
5. Traxler, 596 F.3d at 1014.
6. Arban, 345 F.3d at 406.
employees with certain remedies, including front pay. While front pay is not explicitly mentioned in the statute as an available remedy, courts have granted plaintiffs front pay awards under the FMLA as an equitable remedy in lieu of reinstatement.

When reinstatement is inappropriate or unavailable, courts may determine that an employee is eligible for a front pay award instead. Front pay allows an employee to receive the salary and benefits that would have been earned if the employee was not terminated. While the circuits agree that a judge must determine when a front pay award is available, they are split on who should determine the amount of the award—the judge or the jury. A majority of the circuits find that only a judge should calculate the amount of the front pay award. In contrast, the Sixth Circuit holds that a jury should calculate the amount.

In the recent decision of Traxler v. Multnomah County, in analyzing whether the judge or the jury should calculate the amount of a front pay award, the Ninth Circuit noted that this was an issue of first impression for the circuit. The court’s analysis focused on front pay being classified as an equitable remedy. Although the Seventh Amendment preserves the right to trial by jury for legal claims, there is no such right to a jury for equitable claims. Thus, the Ninth Circuit joined the majority of circuits in adopting the judicial determination approach.

This article argues all circuits should adopt the majority approach that only a judge should be able to determine the availability and amount of the front pay award, although judges should continue to use advisory juries to aid them in calculating the amount of the award. An advisory jury would give a judge the input of a jury who had heard all the facts, while allowing the judge to retain the ultimate power to decide the amount of the front pay award.

Part II of this article will examine the FMLA and review the statute’s legislative history. Additionally, it will examine the equitable remedies of

12. See infra Part II. B.
14. Compare Traxler, 596 F.3d at 1014 (9th Cir. 2010) (finding that the court determines the availability and amount of the front pay award under the FMLA) with Arban v. West Publ’g Corp., 345 F.3d 390, 406 (6th Cir. 2003) (finding that while the court determines the availability of the front pay award under the FMLA, the jury determines the amount of the award).
15. Traxler, 596 F.3d at 1011.
16. Arban, 345 F.3d at 405-06.
17. Traxler, 596 F.3d at 1009.
18. Id. at 1011-1012.
20. See id. at 1014.
reinstatement and front pay. It will also examine the purpose of a front pay award. Part III will examine the current circuit split existing between the judicial determination approach and the jury determination approach. Part IV will critique and analyze each approach to awarding front pay. Additionally, Part IV will advocate that the majority approach should be adopted, but suggest that judges should use advisory juries to aid them in calculating the amount of the award.

II. BACKGROUND AND FACTS

A. Family and Medical Leave Act

1. The Statute

In 1993 Congress passed the FMLA, which provides that “covered employer[s] must grant an eligible employee up to a total of [twelve] workweeks of unpaid leave during any [twelve]-month period”\(^{21}\) under certain circumstances.\(^{22}\) For example, an employee may take job-protected leave for the birth of a child or for the placement of an adopted or foster child.\(^{23}\) Additionally, an employee can take leave “to care for an immediate family member (spouse, child, or parent) with a serious health condition; or...when the employee is unable to work because of a serious health condition.”\(^{24}\)

An employee must meet at least three requirements to be eligible for leave benefits under the FMLA.\(^{25}\) First, an employee must have worked for the employer for at least twelve months.\(^{26}\) Second, an employee must have worked at least 1,250 hours in the past twelve months.\(^{27}\) Third, the employee’s place of employment must have at least fifty employees within a seventy-five mile radius.\(^{28}\) Furthermore, an employee must meet additional requirements to take FMLA leave for an immediate family member’s or the employee’s own serious health condition.\(^{29}\) Under the FMLA, a “‘serious health condition’ means an
illness, injury, impairment, or physical or mental condition that involves either: inpatient care...or subsequent treatment in connection with inpatient care; or continuing treatment by a health care provider.\footnote{30} If an employee meets these requirements, the employer must allow the employee to take unpaid leave of up to twelve weeks per twelve-month period.\footnote{31}

In addition to providing an employee with the opportunity to take unpaid leave, the FMLA provides an employee with job protection while on leave.\footnote{32} Once an employee returns to work from FMLA leave, the employee is entitled to be reinstated.\footnote{33} If the original job is unavailable, then the employee must be given “an equivalent job with equivalent pay, benefits, and other terms and conditions of employment.”\footnote{34} However, the “employee has no greater right to reinstatement or to other benefits and conditions of employment than if the employee had been continuously employed during the FMLA leave period.”\footnote{35} Thus, an employer may deny reinstatement if it can prove that the “employee would not otherwise have been employed at the time of reinstatement.”\footnote{36} For example, if the employee was hired to work on a specific project and the project is completed before the employee returns from FMLA leave, the employee is not entitled to reinstatement.\footnote{37}

In general, there are two types of FMLA lawsuits.\footnote{38} First, employees often sue if there is a refusal on the part of the employer to provide an FMLA benefit to the employee, and the employee is entitled to leave time, continued health insurance benefits, or reinstatement.\footnote{39} Second, employees also sue based on allegations that the employee has been discriminated against because the employee utilized FMLA leave.\footnote{40}

If an employer violates the FMLA, the statute provides for certain remedies.\footnote{41}

\footnotetext[30]{Id. (emphasis excluded).}
\footnotetext[31]{U.S. Dept. of Labor, Leave Benefits, supra note 23; U.S. Dept. of Labor, Family and Medical Leave Act, supra note 8.}
\footnotetext[32]{29 C.F.R. § 825.214 (2008). See also U.S. Dept. of Labor, Family and Medical Leave Act, supra note 8.}
\footnotetext[33]{29 C.F.R. § 825.214 (2008) See also U.S. Dept. of Labor, Fact Sheet # 28, supra note 21, at 3.}
\footnotetext[35]{29 C.F.R. § 825.216(a) (2008).}
\footnotetext[36]{Id.}
\footnotetext[37]{29 C.F.R. § 825.216(a)(3) (2008).}
\footnotetext[39]{Id.}
\footnotetext[40]{Id.}
\footnotetext[41]{29 U.S.C. § 2617 (2006).}
(a)(1)…Any employer who violates section 105 [29 U.S.C. § 2615] shall be liable to any eligible employee affected…for damages equal to…the amount of [(1)] any wages, salary, employment benefits, or other compensation denied or lost to such employee by reason of the violation; or [(2)] in a case in which wages, salary, employment benefits, or other compensation have not been denied or lost to the employee, any actual monetary losses sustained by the employee as a direct result of the violation, such as the cost of providing care, up to a sum equal to 12 weeks…of wages or salary for the employee.42

In sum, this section of the statute allows for damages such as back pay or other expenses.43 Additionally, this section allows for interest on the above amounts as well as liquated damages.44 Liquidated damages “essentially double a plaintiff’s backpay award.”45 However, a judge may choose to not award or to limit liquidated damages if the employer shows “that it acted in good faith and reasonably believed it had complied with the FMLA’s requirements.”46

Moreover, the statute expressly allows for equitable remedies, and employers may be liable “for such equitable relief as may be appropriate, including employment, reinstatement, and promotion.”47 Equitable remedies are derived from the legal doctrine of equity, which arises when typical legal remedies are in some way inadequate and the court must fashion some other method of achieving justice.48 The Ninth Circuit stated in *Traxler v. Multnomah County* that “the court’s power under the FMLA to award front pay, as an alternative to reinstatement, is derived solely from [this] statutory provision.”49 While this section does not expressly grant the remedy of front pay, courts have found that a front pay award is an appropriate equitable remedy in lieu of reinstatement.50 For example, the Ninth Circuit noted that reinstatement may be inappropriate when the employee’s relationship with the employer has broken down.51

43. See id.
46. Id.
49. Traxler v. Multnomah County, 596 F.3d 1007, 1012 (9th Cir. 2010) (citing 29 U.S.C. §2617(a)(1)(B) (2006)).
50. Id. at 1011.
51. Id. at 1012.
2. Legislative History

The FMLA has a unique purpose when compared to other employment-related statutes.\(^{52}\) While employment discrimination statutes such as Title VII or the Age Discrimination in Employment Act (“ADEA”) seek to prevent discrimination in hiring or firing employees or in the workplace environment,\(^{53}\) the FMLA seeks to protect an employee’s job while on leave and allow an employee to take unpaid leave.\(^{54}\) The FMLA is designed to allow employees “to balance their work and family [lives].”\(^{55}\)

The legislative history indicates that the primary purpose of the FMLA is job protection.\(^{56}\) In congressional hearings, the Senate and House sub-committees\(^{57}\) discussed the benefits of enacting the FMLA, including job protection.\(^{58}\) Senator Christopher J. Dodd noted that the FMLA was designed to help employees “balance work and family responsibilities”\(^{59}\) by providing them with “short-term job security…in times of family or medical emergency.”\(^{60}\) Senator Dodd noted that providing temporary, job-protected leave would be beneficial because it would save employers money in not having to hire or train other employees.\(^{61}\)

Senator Ted Kennedy echoed Senator Dodd’s remarks in his testimony at the Senate Hearing.\(^{62}\) In particular, he noted that when employees have the opportunity to care for their children or ill family members, they can be more productive at work.\(^{63}\) Senator Kennedy also stated that this productivity increases even more when employees no longer fear job loss because they take too much time off from work to care for their loved ones.\(^{64}\)

Testimony at the House Hearing reiterated Senators Dodd and Kennedy’s remarks.\(^{65}\) Then-Secretary of Labor Robert B. Reich\(^{66}\) testified at both the

---

\(^{52}\) See 29 C.F.R. § 825.101 (2008).


\(^{55}\) See 29 C.F.R. § 825.101(a) (2008).

\(^{56}\) See FMLA Senate Hearing, supra note 9; FMLA House Hearing, supra note 9.

\(^{57}\) The congressional hearings discussed in this Note are the hearings before the Subcommittee on Children Family, Drugs, and Alcoholism of the Senate Committee on Labor and Human Resources (“Senate Hearing”) on January 22, 1993 and the Subcommittee on the Labor-Management Relations on the House Committee on Education and Labor (“House Hearing”) on January 26, 1993.

\(^{58}\) See FMLA Senate Hearing, supra note 9; FMLA House Hearing, supra note 9.

\(^{59}\) FMLA Senate Hearing, supra note 9, at 1 (statement of Christopher J. Dodd, Senator).

\(^{60}\) Id.

\(^{61}\) Id. at 2; see id. at 3 (statement of Ted Kennedy, Senator).

\(^{62}\) See id. at 3-5 (statement of Ted Kennedy, Senator).

\(^{63}\) Id. at 3.

\(^{64}\) Id.

\(^{65}\) See FMLA House Hearing, supra note 9, at 5 (statement of William D. Ford, U.S. House Representative) (“The bill would protect people when they take time off from work for important family responsibilities. Parents should not have to choose between the jobs they need and the children who need them…this legislation is an affirmation of human dignity and family life.”) (quoting comments from the U.S. Conference of Catholic Bishops); id. at 7 (statement of Matthew
Senate Hearing and House Hearing. In his testimony, Secretary Reich cited two purposes that the FMLA would serve. First, it would provide American workers with minimum standards of fairness and decency by ensuring that they could take care of their loved ones without worrying about job loss. Second, it would help protect America’s workforce by ensuring a good relationship between employers and employees, which, in turn, would help productivity and commitments at companies. Thus, in passing the FMLA, Congress intended that employees be allowed to take time off of work to care for their families or themselves without fearing job loss.

B. Front Pay

This section will first address the remedy of reinstatement because front pay is only awarded when reinstatement is inappropriate. It will then examine the remedy of front pay and describe how courts determine when it is available and how judges and juries calculate the amount of the award. In addition, throughout this section, the remedies of reinstatement and front pay will be evaluated under the FMLA and other employment-related statutes, such as the ADEA or Title

G. Martinez, U.S. House Representative) (“Men and women ought to be able to care for their loved ones without the penalty of losing their livelihood); id. at 7-8 (statement of Donald M. Payne, U.S. House Representative) (“Without the passage of this legislation the majority of Americans risk losing their jobs if they are forced to take leave to care for a new or ill child, for their parents or even for themselves. This measure would provide some basic relief for these families.”); id. at 8 (statement of Eliot L. Engel, U.S. House Representative) (“With increases in the numbers of two-income families and single, working parents, it is becoming incredibly difficult for working individuals to meet the needs of many family responsibilities, not to mention family crises. Placing workers in a situation where they must chose between their job versus their family is an archaic policy.”); id. at 8 (statement of Ron Klink, U.S. House Representative) (“The days of the single breadwinner at work and another parent at home are gone. So many families have both parents working now, and there is little flexibility for emergencies at home. For single parents, there is none. The [FMLA] will help provide flexibility for working families. They will no longer be forced to choose between their jobs and their loved ones, or fear the loss of their jobs because of [their] illness.”).


67. See FMLA House Hearing, supra note 9, at 18-23 (statement of Robert B. Reich, Secretary of Labor); FMLA Senate Hearing, supra note 9, at 21-24 (statement of Robert B. Reich, Secretary of Labor).

68. FMLA Senate Hearing, supra note 9, at 6 (statement of Robert B. Reich, Secretary of Labor).

69. Id.

70. Id.

71. See FMLA Senate Hearing, supra note 9; FMLA House Hearing, supra note 9.

72. Traxler v. Multnomah County,, 596 F.3d 1007, 1012 (9th Cir. 2010).
VII. Courts often use the same analysis when examining these remedies under other employment-related statutes.73

Under employment-related statutes such as the FMLA, an employee has a right to be reinstated to the same or a similar position to the one that the employee had before he or she took leave.74 Because reinstatement is an equitable remedy, courts have held that a judge should determine its availability.75 Although reinstatement is the preferred remedy, it might not be appropriate in all circumstances.76 In Maxfield v. Sinclair International, the Third Circuit provided two examples of when reinstatement is not appropriate.77 First, reinstatement is inappropriate when, at the time of judgment, there is no position available.78 Second, it is also inappropriate when animosity has damaged the relationship between the parties.79

In Pollard v. E.I. du Pont de Nemours & Co.,80 the Supreme Court defined front pay as “money awarded for lost compensation during the period between judgment and reinstatement or in lieu of reinstatement.”81 Front pay is intended to make the former employee whole and it is supposed to put the former employee in the place he or she would have been if not terminated.82 Additionally, courts have characterized front pay as an equitable remedy under the FMLA, as well as under Title VII, the ADEA, and the Americans with Disabilities Act (“ADA”).83

The determination of a front pay award is a two-step process.84 First, a judge determines whether front pay, as an alternative to reinstatement, is available.85 Second, either the judge or a jury then calculates the amount of the front pay

73. See e.g., Traxler, 596 F.3d at 1012-13 (quoting Duke v. Uniroyal, Inc., 928 F.2d 1413, 1424 (4th Cir. 1991) (ADEA claim)); Arban v. West Publ’g Corp., 345 F.3d 390, 406 (6th Cir. 2003) (citing Simpson v. Ernst & Young, 100 F.3d 436, 444 (6th Cir. 1996) (ERISA claim); Roush v. KFC Nat’l Mgmt. Co., 10 F.3d 392, 398 (6th Cir. 1993) (ADEA claim)).
76. Maxfield, 766 F.2d at 796.
77. Id. at 788.
78. Id.
79. Id.
81. Id. at 846. See also Doom, supra note 3, at 531 (quoting Pollard, 532 U.S. at 846).
82. Doom, supra note 3 at 530.
83. Doom, supra note 3 at 535. See also Pollard, 532 U.S. at 849-50.
84. Downey v. Strain, 510 F.3d 534, 544 (5th Cir. 2007); Smith v. Diffee Ford-Lincoln-Mercury, Inc., 298 F.3d 955, 964-66 (10th Cir. 2002); Cline v. Wal-Mart Stores, Inc., 144 F.3d 294, 307 & n.3 (4th Cir. 1998).
85. Arban v. West Publ’g Corp., 345 F.3d 390, 406 (6th Cir. 2003); Cline, 144 F.3d at 307 & n.3; Roush v. KFC Nat’l Mgmt. Co., 10 F.3d 392, 398-99 (6th Cir. 1993).
award. The circuits are split on whether it is proper for the judge or the jury to make this calculation.

As to the first step, a judge must determine that an order of reinstatement is inappropriate. Reinstatement may be inappropriate when the plaintiff’s former position is no longer available or if the employer-employee relationship is broken down by hostility or animosity. In these circumstances, a judge may determine that front pay is an available award.

As to the second step, in the majority of circuits the judge calculates the front pay award; however, in the Sixth Circuit the jury determines the amount. The front pay award is calculated by discounting the “present value of the difference between the earnings he would have received in his old employment and the earnings he can be expected to receive in his present and future…employment.” In addition, judges or juries consider a variety of other factors when determining the appropriate amount and length of the front pay award. Although the factors may vary among the circuits, the following list of factors provides an example of what judges or juries may consider: “the relationship of damages to the employer’s wrongdoings, the employee’s years to retirement,…[the] employee’s lost opportunities,” “the presence of intimidating behavior towards the plaintiff, the effect of the illicit action on the complainant’s emotional health, and the feasibility of reinstatement in light of the employer-(former) employee relationship.”

86. Traxler v. Multnomah County, 596 F.3d 1007, 1014 (9th Cir. 2010); Arban, 345 F.3d at 406-07.
87. Traxler, 596 F.3d at 1014; Arban 345 F.3d at 406-07.
88. Traxler, 596 F.3d at 1012 (citing Whittlesey v. Union Carbide Corp., 742 F.2d 724, 728 (2d Cir. 1984)); Doom, supra note 3 at 529-30.
89. Traxler, 596 F.3d at 1012.
90. Id
91. Compare Arban, 345 F.3d at 406 (finding that the while the court determines the availability of a front pay award under the FMLA, the jury determines the amount of the award) with Traxler, 596 F.3d at 1014 (finding that the court determines the availability and amount of the front pay award under the FMLA) (the court’s view is in accord with the Fourth, Fifth, and Tenth Circuits).
92. McKnight v. Gen. Motors Corp., 908 F.2d 104, 116 (7th Cir. 1990). See also Doom, supra note 3 at 529 (quoting McKnight908 F.2d at 116).
93. See e.g. Arban, 345 F.3d at 406; Downey v. Strain, 510 F.3d 534, 544 (5th Cir. 2007) (citing Reneau v. Wayne Griffin & Sons, Inc., 945 F.2d 869, 871 (5th Cir. 1991) (other factors include “(1) the length of prior employment, (2) the permanency of the position held, (3) the nature of the work, (4) the age and physical condition of the employee, (5) possible consolidation of jobs, and (6) the myriad other non-discriminatory factors which could validly affect the employer/employee relationship.”)).
94. Doom, supra note 3, at n.99 (citing Brian H. Redmond, Annotation, Award of Front Pay Under State Job Discrimination Statutes, 74 A.L.R. 4th 746, at § 6 (1989)).
Moreover, a plaintiff has a duty to mitigate damages. Thus, a plaintiff must attempt to find other comparable employment before a judge or jury can award front pay. However, this duty is limited because the unemployed person is not required to take a lower paying or less prestigious position, go into a new career, or take a menial position. Additionally, the employer has the burden of proving that a plaintiff failed to mitigate his or her damages.

III. CIRCUIT SPLIT

While all circuits agree that the judge should determine the availability of a front pay award, there is a split among the circuits on who should determine the amount of the award. Almost all of the circuits find that the judge should determine the amount, while only the Sixth Circuit holds that it is an issue for the jury to determine.

A. Judge Determines Amount of Front Pay

1. Judicial Determination of Front Pay in The Ninth Circuit

In Traxler v. Multnomah County, the Ninth Circuit evaluated, for the first time, whether the court or the jury should determine the amount of a front pay award given under the FMLA. The Multnomah County Sheriff’s Office (“County”) hired Jill Traxler (“Traxler”) in 1987. Traxler received several promotions throughout her tenure with the County, eventually attaining the position of Human Resources Manager I in 1998. Traxler took medical leave under the FMLA twice, in 2002 and 2005. The County eliminated Traxler’s position in July 2005. She was placed on paid administrative leave and demoted to a lower paying position.

During this time, Traxler took additional FMLA leave. The County fired Traxler in September 2005, after a poor performance review which Traxler

---

96. Buckman, supra note 45, at § 2(a).
97. Id.
99. Id.
100. See e.g. Traxler v. Multnomah County, 596 F.3d 1007, 1011 (9th Cir. 2010); Arban v. West Publ’g Corp., 345 F.3d 390, 406 (6th Cir. 2003).
101. Traxler, 596 F.3d at 1011(discussing how other circuits have held).
102. Arban, 345 F.3d at 406.
103. Traxler, 596 F.3d at 1010.
104. Id. at 1010.
105. Id.
106. Id.
107. Id.
108. Id.
alleged was due to the exercise of her rights to FMLA leave. She then sued the County for violating the FMLA and the state’s equivalent statute. The jury found that the County violated the FMLA when it terminated Traxler and awarded her both front and back pay in the amounts of $1,551,000 and $250,000 respectively.

The County filed a motion for judgment as a matter of law, or alternatively, for a new trial. The County argued that it was an error for the court to allow the jury to calculate the front pay award. The district court agreed that it was an error “because front pay is an equitable remedy that must be determined by the court.” In the alternative, the court found that the award was grossly excessive. The district court vacated the front pay award and recalculated it to $267,000. Traxler appealed the district court’s decision to the Ninth Circuit Court of Appeals.

The Ninth Circuit began its analysis by evaluating the FMLA statute, specifically the section that governs remedies. The court noted that the first subsection on remedies provides for “damages for past costs including, among other things, wages, salary, employment benefits, or other compensation denied or lost,” as well as “non-wage actual monetary losses, interest, and liquidated damages.” The court also noted that the second subsection provides for future remedies by “allowing ‘for such equitable relief as may be appropriate, including employment, reinstatement, and promotion.’” The Ninth Circuit determined that because front pay is not specified in the first subsection, the court’s authority to grant the award is derived from the language in the second subsection, authorizing “such equitable relief as may be appropriate.”

The Ninth Circuit then noted that classifying front pay as an equitable remedy is consistent with other employment statutes. The Ninth Circuit looked to the Supreme Court’s decision in Pollard v. E.I. du Pont de Nemours & Co., which noted that unlike compensatory damages, front pay is the monetary

109. Traxler, 596 F.3d at 1010.
110. Id.
111. Id.
112. Id.
113. Id.
114. Id.
115. Traxler, 596 F.3d at 1010.
116. Id.
117. Id. at 1010-11.
120. Id.
121. Id. (quoting 29 U.S.C. § 2617(a)(1)(B)).
122. Id. at 1011-12.
123. Id. at 1012 (quoting 29 U.S.C. § 2617(a)(1)(B)).
124. Id. at 1012.
alternative to reinstatement and is offered to compensate an employee who has been wronged yet will not be able to return to their previously held position.\textsuperscript{126}

Additionally, the Ninth Circuit determined that because front pay is an equitable remedy, the statute does not support allowing the judge to determine the availability of front pay, and then the jury to determine that amount of the award.\textsuperscript{127} The court stated that this would split front pay into both an equitable and a legal component, which is impermissible under the statute.\textsuperscript{128}

The \textit{Traxler} court stated that determining when reinstatement is practical involves balancing equitable concerns.\textsuperscript{129} Therefore, because front pay is an alternative to reinstatement, it also involves evaluating equitable factors.\textsuperscript{130} Moreover, the court noted that judicial discretion is important when deciding what amount a plaintiff should be compensated for not being reinstated, and thus that determination is not appropriate for a jury.\textsuperscript{131}

The Ninth Circuit then evaluated \textit{Traxler}'s argument that the monetary award of front pay is a legal remedy, not an equitable one.\textsuperscript{132} In making this argument, \textit{Traxler} relied on the Ninth Circuit's decision in \textit{Cassino v. Reichhold Chemicals, Inc.},\textsuperscript{133} which held that when the court determines that reinstatement is not possible, the jury should determine the amount of the front pay award.\textsuperscript{134} The Ninth Circuit rejected this argument, stating that this was dicta and noting that courts can use advisory juries, but the court has the ultimate power to determine the amount of the front pay award.\textsuperscript{135} Thus, the Ninth Circuit affirmed the district court's decision to vacate and modify the front pay award, holding that "under the FMLA, front pay is an equitable remedy that must be determined by the court, both as to the availability of the remedy and the amount of any award."\textsuperscript{136}

2. Judicial Determination of Front Pay in the Fourth Circuit

The Fourth Circuit, as well as the Fifth and Tenth Circuits, has also found that under the FMLA, the judge should determine the availability and the amount of a front pay award.\textsuperscript{137} In \textit{Cline v. Wal-Mart Stores, Inc.}, the Fourth Circuit

\begin{itemize}
\item \textsuperscript{126} \textit{Traxler}, 596 F.3d at 1012 (quoting \textit{Pollard}, 532 U.S. at 853 n.3) (internal quotations omitted)).
\item \textsuperscript{127} Id. at 1012.
\item \textsuperscript{128} Id.
\item \textsuperscript{129} Id.
\item \textsuperscript{130} See id.
\item \textsuperscript{131} Id. at 1012.
\item \textsuperscript{132} \textit{Traxler}, 596 F.3d at 1013.
\item \textsuperscript{133} Id. \textit{See also} \textit{Cassino v. Reichhold Chemicals, Inc.}, 817 F.2d 1338 (9th Cir. 1987).
\item \textsuperscript{134} \textit{Traxler}, 596 F.3d at 1013 (quoting \textit{Cassino}, 817 F.2d at 1347).
\item \textsuperscript{135} Id.
\item \textsuperscript{136} Id. at 1011.
\item \textsuperscript{137} Downey \textit{v. Strain}, 510 F.3d 534, 544 (5th Cir. 2007); \textit{Cline v. Wal-Mart Stores, Inc.}, 144 F.3d 294, 307 & n.3 (4th Cir. 1998); \textit{see also} \textit{Smith v. Diffie Ford-Lincoln-Mercury, Inc.}, 298 F.3d 955, 964-66 (10th Cir. 2002) (On remand, the district judge was instructed to make new
discussed the issue of a front pay award under the FMLA. In 1993, Wal-Mart Stores, Inc. ("Wal-Mart") hired Keith Cline ("Cline") as a night maintenance worker. Less than a year later, Wal-Mart promoted Cline to night maintenance supervisor, and soon after, Cline took FMLA leave for a serious medical condition. Before Cline returned to work, Wal-Mart hired another employee to replace Cline as night maintenance supervisor.

In November 1994, Cline was able to return to work without medical restrictions, but learned that Wal-Mart had demoted him for alleged health reasons. Cline repeatedly threatened to sue Wal-Mart because of his demotion. In January 1995, Wal-Mart fired Cline for clocking in early to a meeting. Another employee also clocked in early and although Wal-Mart stated that he was also fired, that employee was eventually rehired and promoted to a supervisory position.

Cline sued Wal-Mart, inter alia, for violating the FMLA by failing to reinstate him to the supervisory position and for retaliation against his threats to sue. The district court granted Cline’s motion for judgment as a matter of law on the issue of reinstatement and the jury found that Wal-Mart violated the FMLA by retaliating against Cline. The jury awarded Cline damages for lost wages, front pay, back pay, and liquidated damages on his FMLA restoration and retaliation claims. The district court denied Wal-Mart’s motion for a new trial on the FMLA retaliation claim and denied Wal-Mart’s motion for judgment notwithstanding the verdict on the retaliation claim and the jury’s award of front pay.

On appeal to the Fourth Circuit, Wal-Mart argued that it was inappropriate for the district court to submit the issue of front pay to the jury. The Fourth Circuit found that front pay is an equitable remedy because the award is an alternative to reinstatement, which is an equitable remedy. The court looked

138. Cline, 144 F.3d at 307.
139. Id. at 298.
140. Id.
141. Id. at 299.
142. Id.
143. Id.
144. Cline, 144 F.3d at 299.
145. Id.
146. Id.
147. Id. at 300.
148. Id.
149. Id.
150. Cline, 144 F.3d at 300.
151. Id. at 307.
152. Id.
to its decision in *Duke v. Uniroyal*,\(^\text{153}\) where it noted that a court should make all determinations of a front pay award.\(^\text{154}\) Additionally, the *Duke* court held that a judge should calculate the amount of a front pay award, because the varying factual circumstances surrounding an award of front pay do not provide the jury with an adequate legal standard to guide it when making the determination of the amount of the award.\(^\text{155}\) The *Cline* court found that its analysis in *Duke* also applied to front pay awards under the FMLA.\(^\text{156}\) Thus, it vacated and remanded the amount of the front pay award, stating that the district court needed to recalculate the amount of the award.\(^\text{157}\)

3. Judicial Determination of Front Pay in the Eighth Circuit Under the ADEA

In *Newhouse v. McCormick & Co., Inc.*,\(^\text{158}\) the Eighth Circuit addressed the issue of whether a judge or a jury should determine the amount of the front pay award in the ADEA context.\(^\text{159}\) Although the plaintiff in *Newhouse* brought an ADEA claim against his employer, the Eighth Circuit’s analysis applies equally in the FMLA context.\(^\text{160}\) McCormick & Co., Inc. (“McCormick”) interviewed Richard Newhouse (“Newhouse”), a 61-year-old man, for a sales representative position.\(^\text{161}\) Although Newhouse had held the position in the past and was qualified for the job, McCormick hired a younger man with less experience.\(^\text{162}\) Newhouse applied for full-time employment elsewhere, but was unsuccessful and instead opted to retire and receive social security benefits.\(^\text{163}\) Had Newhouse been able to obtain full-time employment with McCormick, he would not have retired at that age.\(^\text{164}\)

Newhouse sued McCormick for age discrimination under the federal and state statutes.\(^\text{165}\) After trial, the jury returned a verdict for Newhouse “award[ing] him $59,426.76 in back pay and $206,359 in front pay from the date of the verdict until the date Newhouse normally would have retired,” as well as


\(^{154}\) *Cline*, 144 F.3d at 307 & n.3 (citing *Duke*, 928 F.2d at 1424).

\(^{155}\) Id. (quoting *Duke*, 928 F.2d at 1424).

\(^{156}\) Id.

\(^{157}\) Id.

\(^{158}\) *Newhouse v. McCormick & Co., Inc.*, 110 F.3d 635 (8th Cir. 1997).

\(^{159}\) Id. at 641-643; see e.g., *Traxler v. Multnomah County*, 596 F.3d 1007, 1016 (9th Cir. 2010) (quoting *Pollard v. E.I. du Pont de Newmours & Co.*, 532 U.S. 843 (2001)).

\(^{160}\) See e.g., *Traxler*, 596 F.3d at 1012-13 (quoting *Duke*, 928 F.2d at 1424 (4th Cir. 1991) (ADEA claim)); *Arban v. West Publ’g Corp.*, 345 F.3d 390, 406 (6th Cir. 2003) (citing *Simpson v. Ernst & Young*, 100 F.3d 436, 444 (6th Cir. 1996) (ERISA claim); *Roush v. KFC Nat’l Mgmt. Co.*, 10 F.3d 392, 398 (6th Cir. 1993) (ADEA claim)).

\(^{161}\) *Newhouse*, 110 F.3d at 637.

\(^{162}\) Id.

\(^{163}\) Id. at 638-39.

\(^{164}\) Id. at 639.

\(^{165}\) Id.
liquidated damages and attorney’s fees for his federal age discrimination claim.166 The jury also awarded Newhouse back pay, front pay, and attorney’s fees on his state age discrimination claim.167

After denying McCormick’s “motions for judgment as a matter of law, for a new trial, or alternatively, to alter or amend the judgment,” the district court ordered a new trial on the issue of damages, unless Newhouse agreed to a reduction of the front pay award.168 Newhouse agreed to a reduction of the front pay award, and thus, the district court reduced the award to $158,365.96.169 McCormick appealed, arguing that it was error for the district court to determine that front pay was an available remedy in lieu of reinstatement and to allow the jury to calculate the amount of the front pay award.170

On appeal to the Eighth Circuit, McCormick argued that it was error for the district court to determine that reinstatement was not appropriate.171 The district court determined that the lawsuit caused tension between Newhouse and McCormick, including Newhouse’s would-be supervisor and that this tension made reinstatement inappropriate.172 The Eighth Circuit found that the district court did not abuse its discretion when it found that reinstatement was inappropriate and thus the court ordered front pay for two reasons.173 First, the court noted that the tension caused by the litigation was sufficient to make reinstatement inappropriate.174 Second, the court determined that reinstatement was inappropriate because Newhouse was receiving social security benefits.175

At the earlier trial in district court, McCormick objected to Newhouse’s request that the front pay issue be sent to the jury, arguing that because front pay was an equitable remedy, the judge should determine the amount of the award.176 On appeal, McCormick argued that the district court abused its discretion by allowing the jury to determine the amount of the front pay award.177 The district court, in rejecting McCormick’s argument, found that because the jury calculates damages in other types of cases, it was appropriate for it to calculate the front pay award here.178

In considering McCormick’s argument, the Eighth Circuit first examined its previous decisions on the issue of front pay.179 It noted that it previously held

166. Id.
167. Newhouse, 110 F.3d at 639.
168. Id.
169. Id.
170. Id. at 639, 641-42.
171. Id. at 641.
172. Id.
173. Newhouse, 110 F.3d at 641-42.
174. Id.
175. Id. at 642.
176. Id.
177. Id.
178. Id.
179. Newhouse, 110 F.3d at 642.
that a district judge had discretion to calculate the award him or herself, or to submit the issue to the jury.\textsuperscript{180} The court also noted that in \textit{Doyne v. Union Electric Co.},\textsuperscript{181} it declined to decide whether a jury is permitted to calculate the amount of a front pay award.\textsuperscript{182}

The Eighth Circuit then evaluated the circuit split on this issue.\textsuperscript{183} It first reviewed the law of the circuits that adopted the jury determination approach.\textsuperscript{184} It noted that at that time in the Third, Fifth, Sixth, and Ninth Circuits, the jury calculated the amount of front pay.\textsuperscript{185} The Eighth Circuit then stated that “[t]hese courts have adopted this view without any analysis, merely reciting the one statement of dicta from \textit{Maxfield [v. Sinclair Int’l]} in which the Third Circuit, without citation to any authority, stated that ‘the amount of damages available as front pay is a jury question.’”\textsuperscript{186}

The Eighth Circuit next examined the law of the circuits that adopted the judicial determination approach.\textsuperscript{187} It noted that the Second, Fourth, Seventh, and Tenth Circuits had held that only the judge could calculate the front pay award.\textsuperscript{188} The Eighth Circuit then evaluated the Second Circuit’s decision in \textit{Dominic v. Consol. Edison Co. of New York, Inc.},\textsuperscript{189} finding its analysis most persuasive.\textsuperscript{190}

In \textit{Dominic}, the Second Circuit provided two reasons for its decision that only the judge may calculate the amount of the front pay award.\textsuperscript{191} First, the Second Circuit noted that the ADEA statute only allows for a jury trial on legal damages, not equitable remedies such as front pay.\textsuperscript{192} Second, it stated:

There is much overlap between the facts relevant to whether an award of front pay is appropriate and those relevant to the size of the award. For example, both questions turn in part on the ease with which the employee will be able to find other employment. To divide the fact-

\begin{footnotesize}
\textsuperscript{180} \textit{Id.}
\textsuperscript{181} \textit{Doyne v. Union Elec. Co.}, 953 F.2d 447 (8th Cir. 1992).
\textsuperscript{182} \textit{Newhouse}, 110 F.3d at 642 (citing \textit{Doyne}, 953 F.2d at 451).
\textsuperscript{183} \textit{Id.} at 642-43.
\textsuperscript{184} \textit{Id.} at 642.
\textsuperscript{185} \textit{Id.} All circuits but the Sixth have since adopted the judicial determination approach. \textit{See} Traxler v. Multnomah County, 596 F.3d 1007, 1013 (9th Cir. 2010) (stating “[n]ot surprisingly, only one circuit is squarely in opposition--the Sixth Circuit); \textit{see also} Donlin v. Phillips Lighting N. Am. Corp., 581 F.3d 73, 86, 88 & n.11 (3d Cir. 2009) (the district court used a jury’s advisory verdict on the issue of front pay which the court modified); Deloach v. Delchamps, Inc., 897 F.2d 824 (5th Cir. 1990) (The court in \textit{Newhouse} referred to a 5th circuit decision from 1989 in its analysis).
\textsuperscript{186} \textit{Newhouse}, 110 F.3d at 642 (quoting \textit{Maxfield v. Sinclair Int’l}, 766 F.2d 788, 796 (3d Cir. 1985)).
\textsuperscript{187} \textit{Id.}
\textsuperscript{188} \textit{Id.}
\textsuperscript{189} \textit{Dominic v. Consol. Edison Co. of New York, Inc.}, 822 F.2d 1249 (2d Cir. 1987)
\textsuperscript{190} \textit{Newhouse}, 110 F.3d at 643.
\textsuperscript{191} \textit{Id.} at 643 (citing \textit{Dominic}, 822 F.2d at 1257-58).
\textsuperscript{192} \textit{Id.}
\end{footnotesize}
finding responsibilities in such circumstances would be anomalous and would risk inconsistent decisions...For example, a judge might find front pay appropriate, but the jury might award only a nominal sum based on its belief that the employee could secure immediate employment.193

The Newhouse court agreed with the Second Circuit’s analysis and adopted the judicial determination approach, holding that only a judge can calculate the amount of the front pay award.194 It determined that calculating front pay is an equitable issue and that the district court should not have submitted the issue to the jury.195 However, the Eighth Circuit did not remand the issue because the district court stated that it would have awarded Newhouse a reduced front pay award of $84,062.196 The Eighth Circuit found this front pay award was appropriate.197 Despite the Eighth Circuit’s Newhouse decision and the decisions of other circuits which have taken the same approach when deciding this issue, the Sixth Circuit has come to a different conclusion.

B. Jury Determines Amount of Front Pay

In Arban v. West Publishing Corp.198 the Sixth Circuit discussed the calculation of a front pay award.199 Lawyer’s Cooperative Publishing (“LCP”) hired Daniel Arban (“Arban”) in 1995 as a sales representative, and subsequently merged with the West Publishing Company (“West”) in 1996.200 During his tenure at LCP and West, Arban suffered from health problems.201

In February 1998, West found that Arban made misrepresentations with his sales accounts and in April, sent him a warning letter concerning the situation.202 In December 1998, management met to discuss the violations and Arban’s supervisor suggested termination.203 Other members of management agreed, but decided not to terminate Arban’s employment until after the holidays.204

In late December, Arban suffered from a flare up of his health condition and after visiting with his doctor, he requested FMLA leave through human resources.205 After contacting the disability insurance carrier, the insurance

193. Id. (quoting Dominic, 822 F.2d at 1257) (alteration in original).
194. Id.
195. Id.
196. Newhouse, 110 F.3d at 643.
197. Id.
199. Id. at 405-07.
200. Id. at 394-395.
201. Id. at 395.
202. Id.
203. Id. at 395-396.
204. Arban, 345 F.3d at 396.
205. Id.
carrier approved his short-term disability benefits. He visited his doctor again and was advised to take three weeks off from work. Arban notified his manager of this by telephone and e-mail.

While on leave, West tried to get account information from Arban so that another sales representative could work on his account. Arban provided some information, but West repeatedly asked for more. Arban stated that he thought working while on leave would violate his FMLA benefits and refused to provide any more information. In early January, while he was still on leave, Arban’s manager contacted him and informed him that West was terminating his position. However, Arban resigned before being terminated.

Arban then sued West, alleging that his termination violated the FMLA. The jury found that West violated the FMLA, but the district court did not award front pay. The district court determined that front pay was not available because Arban provided insufficient and speculative evidence on lost wages and compensation. Additionally, Arban found other employment. The district court stated that based on this evidence, a reasonable jury would not award front pay. The district court denied West’s motion for judgment as a matter of law. West appealed the decision to the Sixth Circuit Court of Appeals based upon the jury instructions relating to Arban’s termination, while Arban appealed based upon the issue of front pay.

The Sixth Circuit discussed whether it was error for the district court to refuse to submit the issue of front pay to the jury. The Sixth Circuit first noted that although a judge should determine whether front pay is available, a jury should determine the amount of the award. The Sixth Circuit then discussed how a judge determines when front pay is available. The court stated “the question to be answered is whether front pay damages are needed in a particular

206. *Id.*
207. *Id.* at 397.
208. *Id.*
209. *Id.* at 397-98.
210. *Arban*, 345 F.3d at 397-98.
211. *Id.*
212. *Id.* at 398-99.
213. *Id.* at 400.
214. *Id.*
215. *Id.*
216. *Arban*, 345 F.3d at 406.
217. *Id.*
218. *Id.*
219. *Id.* at 400.
220. *Id.* at 400, 404.
221. *Id.* at 405. The district court concluded the FMLA does not provide for a remedy of front pay.
222. *Arban*, 345 F.3d at 406.
223. *Id.*
case to make the plaintiff whole. The court found that judges should evaluate the following six factors:

- (1) an employee’s duty to mitigate,
- (2) the availability of employment opportunities,
- (3) the period within which one by reasonable efforts may be re-employed,
- (4) the employee’s work and life expectancy,
- (5) the discount tables to determine the present value of future damages and
- (6) other factors that are pertinent on prospective damage awards.

After reviewing the evidence, the Sixth Circuit found that Arban’s evidence in favor of a front pay award was speculative because his request was based on what he thought he would make and the Social Security Wage Index. Thus, the Sixth Circuit held that it was not an abuse of discretion for the district court to decline to submit the front pay issue to the jury, while maintaining that in cases where the court decides that front pay is available, the calculation of the award should be completed by the jury.

District courts in the Sixth Circuit have noted that the Sixth Circuit’s decision in Arban requires district courts to “use a bifurcated method for front pay questions.” First, the court must decide on the availability of front pay as a remedy, and then the jury must calculate the amount of any award.

Other circuits, which had previously held that juries should calculate the amount of a front pay award, have adopted the judicial determination approach only in the past decade. These include the Third, Fifth, and Ninth Circuits. In Newhouse, the Eighth Circuit addressed the issue of whether a judge or jury

---

224. Id. (quoting Wilson v. Int’l Bro. of Teamsters, 83 F.3d 747, 756-57 (6th Cir. 1996) (internal quotations omitted)).
225. Id. (quoting Roush v. KFC Nat’l Mgmt. Co., 10 F.3d 392, 399 (6th Cir. 1993) (internal quotations omitted)).
226. Id. at 407.
227. Id.

229. Case, 2010 U.S. Dist. LEXIS 41525 at *2. In Case, the Western District of Michigan denied the defendant’s motion to strike the plaintiff’s jury demand because a jury is required to determine the amount of a front pay award under both the FMLA and the ADA. Id. at *2-3. Cf. Madden v. Chattanooga City Wide Service Dept., 549 F.3d 666, 679-80 (6th Cir. 2008) (analyzing whether a district court abused its discretion in awarding a plaintiff front pay). In Madden, the district court did not analyze whether a judge or jury should determine the amount of a front pay award. See generally id. The district court held a bench trial to determine whether the plaintiff was wrongfully terminated under Title VII and state law. Id. at 671. Although the judge calculated the amount of the front pay award in Madden, this should not be taken as a departure from Arban, because this was a bench trial, in which the judge decides questions of fact, such as the amount of a front pay award.

230. See Maxfield v. Sinclair Int’l, 766 F.2d 788, 796 (3rd Cir. 1985). See also supra Part III.A.
231. See supra Part III.A.
should determine the amount of the front pay award, noting that the Third, Fifth, Sixth, and Ninth Circuits never provided analysis for why they followed the jury determination approach.232

IV. ANALYSIS

A. Critique of Jury Determination of the Front Pay Award

There are two arguments for allowing juries to determine the amount of a front pay award. First, juries already decide questions of fact, including calculations of other types of damages.233 Second, juries already review the issue of front pay on an advisory basis.234

As to the first argument, while some critics argue that complex factual issues are not appropriate for juries,235 others argue that juries already determine complex factual issues including determining the amount of damages in personal injury cases and pain and suffering awards.236 In addition, because juries often determine facts in a variety of complex cases, including medical malpractice cases, securities litigation, and products liability cases, it is argued that juries should also be able to evaluate the necessary factors when calculating a front pay award.237 Supporting this argument is the fact that juries often calculate the amount of future damages awards in personal injury cases and consider factors similar to those involved in calculating front pay awards.238

The jury system is important to the American judicial system and it is important for juries to be able to decide fact issues to ensure that litigants are being evaluated by the standards of their community and not the court alone. Because juries reflect community values, jurors ensure that judicial decisions are based on these values.239 Furthermore, jurors are able to decide complex factual issues without setting a precedent.240 Although jurors may be inexperienced in legal issues, that “very inexperience is an asset because it secures a fresh perception of each trial, avoiding the stereotypes said to infect the judicial

233. See supra Part III.A.&B.
234. See supra Part III.A.&B.
236. Id. at 1494.
237. Id.
238. Id. at 1494-95.
240. Id.
In addition, the collective wisdom of a jury exceeds the wisdom of a single judge. For instance, “[a] judge can thus draw only on that rather more narrow and unrepresentative life experience in determining what is ‘fair’ or ‘reasonable,’ whereas juries can draw on the varied experiences of several people from different walks of life.”

However, a majority of courts have consistently held that it is inappropriate for a jury to calculate the front pay award in an FMLA case. Although juries determine other complex issues, these courts hold that the issue of front pay under the FMLA is too speculative for a jury to determine. Arguably, calculating the amount of a future damages award in personal injury cases, or back pay awards in other cases, is distinguishable from determining the amount of a front pay award under the FMLA. Future damages in personal injury cases are less speculative, because usually the plaintiff no longer has the same earning capacity as before the injury. In contrast, a plaintiff injured under the FMLA typically retains the same earning capacity, which makes the calculation of a front pay award more speculative.

As to the second argument, judges often use advisory juries to suggest awards of front pay. Judges will submit the issue of the amount of the front pay award to the jury and have the jury calculate the amount of the award on an advisory basis. The judge will then make a final determination on what is an appropriate amount. Therefore, if juries are already calculating front pay awards on an advisory basis, there is a strong argument that the issue of front pay should be submitted to the jury for final determination.

However, submitting the issue of a front pay award to a jury for only advisory purposes is vastly different than allowing a jury to determine the amount of a front pay award. When the award is submitted to the jury on an advisory-only basis, the judge still maintains control of the award, but with an additional opinion from others who heard all of the evidence. While using advisory juries is a beneficial tool for judges, the arguments in favor of allowing a judge to determine the amount of a front pay award outweigh the argument that a jury should make the final determination.

242. Id. at 242.
243. Id.
244. See supra Part III.A.
246. Id.
247. Traxler, 596 F.3d at 1013.
248. See e.g., Traxler v. Multnomah County, 596 F.3d 1007, 1013 (9th Cir. 2010); Donlin v. Philips Lighting N. Am. Corp., 581 F.3d 73, 78 & n.1 (3d Cir. 2009); Mota v. Univ. of Tex. Houston Health Sci. Ctr., 261 F.3d 512, 526 (5th Cir. 2001).
249. Traxler, 596 F.3d at 1013.
B. Analysis and Proposal of Judicial Determination of the Front Pay Award

There are two arguments in favor of allowing only a judge to calculate the amount of a front pay award. First, because front pay is an alternative to reinstatement and is thus an equitable remedy, only a judge should determine the amount of the award. Second, only a judge can adequately evaluate the variety of factors necessary when determining the appropriate amount and duration of a front pay award.

First, some courts have held that because front pay is an equitable remedy, only a judge can determine the amount of the award. The Ninth Circuit stated that because the FMLA statute did not explicitly provide for front pay, it was available under the provision allowing for equitable relief. In addition, the Ninth Circuit determined that, as an alternative to reinstatement, front pay is also an equitable remedy. Although critics have argued that front pay is a legal remedy because it is a monetary award, the argument is not persuasive. The majority of circuits have consistently found that under the FMLA, front pay is available under the equitable provision of the statute, through the equitable remedy of reinstatement.

As to the second argument, some courts have held that only the judge can adequately evaluate the factual circumstances necessary to calculate the front pay award. For example, the Fourth Circuit stated in Duke v. Uniroyal, Inc. that because a front pay award involves looking at a variety of future circumstances which can be difficult to predict, the calculation of the award will always be speculative. In addition, the Fourth Circuit reasoned that because of the speculative nature of the award, a jury is unable to properly evaluate these circumstances. The Second Circuit’s analysis in Dominic v. Consol. Edison Co. of New York, Inc. is persuasive on this issue. As discussed in Section III.A., the Second Circuit stated that it was inappropriate for a jury to calculate the amount of front pay. In Dominic, the Second Circuit provided further guidance on what is necessary for a judge to determine the amount of a front pay award. The factors that judges use when determining whether an award of front pay is

251. See supra Part II.B.
252. See e.g., Duke, 928 F.2d at 1423-24.
253. Traxler, 596 F.3d at 1011.
254. Id. at 1011-12.
255. Id.
256. Mehden, supra note 235 at 1484-86.
257. See supra Part III.A.
259. Id. at 1423.
260. Id.
262. See Dominic, 822 F.2d at 1257-58. See also supra Part III.A.
263. Dominic, 822 F.2d at 1257; supra Part III.A.
264. Dominic, 822 F.2d at 1257.
appropriate are similar to the factors considered when calculating the amount of
an award.\footnote{Id.} One factor to consider for both questions is whether the employee
can find alternative employment.\footnote{Id. at 1257-58.} Because of the potential disparity in awards
from case to case, it is inappropriate to split the decision of the front pay award
between the judge and jury by allowing the judge to determine the availability of
the award and the jury to determine the amount.\footnote{Id. at 1257.} The Second Circuit stated
that this would risk inconsistent decisions.\footnote{Id.} Without consistency, plaintiffs
could bring factually similar cases and receive different results. Two different
juries within different circuits could arrive at significantly different front pay
awards. Furthermore, this situation could arise in two different cases within the
same circuit.

By allowing only the judge to calculate the front pay award, the factors listed
in Section II. B. will be consistently evaluated. Recall that these factors include
“the relationship of damages to the employer’s wrongdoings, the employee’s
years to retirement,…[the] employee’s lost opportunities,”\footnote{Doom, supra note 3, at n.99 (citing Brian H. Redmond, Annotation, Award of Front Pay Under State Job Discrimination Statutes, 74 A.L.R. 4th 746, at § 6 (1989)).} “the presence of
intimidating behavior towards the plaintiff, the effect of the illicit action on the
complainant’s emotional health, and the feasibility of reinstatement in light of
the employer-(former) employee relationship.”\footnote{Dede Koffie-Lart & Christopher J. Tyson, Sixth Annual Review of Gender and Sexuality
Law: IV. Employment Law Chapter: Title IV of the Civil Rights Act of 1964, 6 GEO. J. GENDER & L. 615, 632 (2005).} When these factors are
evaluated more consistently, there will be more consistent awards of front pay.
This will provide prospective plaintiffs with more guidance when bringing suits
under the FMLA, because they will have a better idea of the potential remedies
and outcome of their case. It will also allow judges to use other decisions as
guidance in calculating front pay awards. Moreover, it will ensure fairness
among the circuits so that similarly situated plaintiffs are not getting a windfall
in one circuit while getting a nominal award in the next.\footnote{See Duke v. Uniroyal Inc., 928 F.2d 1413, 1423 (4th Cir. 1991). See also supra Part III.A.}

Because of the difficulty involved in calculating a front pay award, the judge
is the appropriate one to determine the amount of the award. The Fifth Circuit’s
decision in Downey v. Strain\footnote{Downey v. Strain, 510 F.3d 534, 544-45 (5th Cir. 2007).} is illustrative. As discussed in Section III.A.,
the district court in Downey had to balance a variety of issues in deciding the
appropriate length of the front pay award.\footnote{Id.} The Downey court evaluated
several factors, including the uncertainty of working in an elected official’s
office, previous requests to be transferred out of the department, and the
speculative nature of a front pay award.\textsuperscript{274} When considering these factors with
the plaintiff’s request for a five year award and the defendant’s request for a two
or three year award,\textsuperscript{275} and in light of the Fourth Circuit’s analysis in \textit{Duke},\textsuperscript{276} the
difficulty in deciding which award is more appropriate is evident. The varieties
of potential circumstances that may be considered are nearly infinite,\textsuperscript{277} and
because these circumstances are dependent upon speculation on future events, it
is better that a more experienced judge, rather than a jury, evaluates these factors
and calculates the award.

Although some degree of speculation will be inherent even when a judge
determines the amount of the award, a judge is still the appropriate one to
determine the amount of the award. Judges are more knowledgeable on the law
and have more experience in determining whether the amount of a front pay
award is appropriate. These qualities make it more appropriate that a judge
determine the amount of the award because a judge can more consistently apply
the relevant factors and therefore provide more consistency among the
circuits.\textsuperscript{278}

Although the Third Circuit noted in \textit{Maxfield} that “an award for future lost
earnings is no more speculative than awards for lost earning capability routinely
made in personal injury and other types of cases,”\textsuperscript{279} it is still inappropriate to
allow a jury to calculate the amount of the award. First, it is important to note
that since \textit{Maxfield}, the Third Circuit has adopted the judicial determination
approach.\textsuperscript{280} Second, in \textit{Duke}, the Fourth Circuit distinguished future wages
awards from front pay awards.\textsuperscript{281} A future wages award is “often determined
with reasonable certainty and awarded as legal damages in circumstances where
the earning capacity of a plaintiff is destroyed or damaged.”\textsuperscript{282} A front pay
award is more speculative because the plaintiff still has an earning capacity.\textsuperscript{283}
More factors must be considered as to when, or if, the plaintiff may find
employment again.\textsuperscript{284} In sum, because plaintiffs bringing FMLA claims are
often still capable of working, the speculative nature of the variety of

\textsuperscript{274} Id.
\textsuperscript{275} Id.
\textsuperscript{277} See id.
\textsuperscript{278} See supra Part II.A.
\textsuperscript{279} Maxfield v. Sinclair Int’l, 766 F.2d 788, 796 (3rd Cir. 1985).
\textsuperscript{280} See Traxler v. Multnomah County, 596 F.3d 1007, 1013 (9th Cir. 2010) (stating “[n]ot
surprisingly, only one circuit is squarely in opposition—the Sixth Circuit); Donlin v. Philips
Lighting North Am. Corp., 581 F.3d 73, 86-88 & n.11 (3rd Cir. 2009) (The district court was not
originally required to submit the issue of front pay to an advisory jury because a bench trial is
sufficient to determine an equitable award like front pay).
\textsuperscript{281} Duke, 928 F.3d at 1423.
\textsuperscript{282} Id. at 1423.
\textsuperscript{283} Id. at 1423-24.
\textsuperscript{284} Id.
circumstances that may occur makes it inappropriate for a jury to decide the amount of front pay.285

Although judges may allow juries to determine the amount of a front pay award on an advisory basis, this still does not justify allowing the jury to calculate the award itself. An award of front pay is also more speculative than other damages because it involves consideration of a variety of factors. In addition, after termination in an FMLA case, the employee often still has a potential earning capacity and could work for an undetermined time into the future.

Because the FMLA is designed to protect an employee’s job while he or she is on leave,286 the FMLA is unique when compared to other employment related statutes because the purpose is to allow employees to care for their families while protecting the integrity of the workforce.287 The award of front pay directly correlates to the purpose of the statute. Because front pay is designed to make an employee whole when reinstatement is not viable,288 it hits at the heart of the primary purpose of the statute—job protection while an employee tends to family’s needs.289

Because the purpose of the statute is to provide job protection for employees as well as protect the economy and workforce, it is necessary that courts across the nation consistently apply the FMLA and award consistent damages, including front pay awards. If all circuits adopt the judicial determination approach, the remedy of front pay will be more consistently calculated. This will ensure that in similar factual situations, plaintiffs across the nation will receive similar awards. Additionally, to allow for fairness to both the employer and the employee, it is important that a judge have final say on determining the availability and the amount of the award. This will help to ensure that the employee is not given a windfall or under-award given the circumstances.

Courts, including the Ninth Circuit, often use advisory juries to calculate the amount of a front pay award.290 Judges should be encouraged to use advisory juries when calculating the amount of a front pay award. Having extra guidance on the issue of the amount of the award will ensure that judges do not miss important factors in the course of the trial and will allow for additional perspective on the issues.

285. See id.
286. See FMLA House Hearing; FMLA Senate Hearing, supra note 9; see also supra Part II.A.2.
287. See FMLA House Hearing, supra note 9; FMLA Senate Hearing, supra note 9; see also supra Part II.A.2.
288. See supra Part II.B.; see also Doom, supra note 3, at 530.
289. See FMLA House Hearing, supra note 9; FMLA Senate Hearing, supra note 9; see also supra Part II.A.2.
290. See Traxler v. Multnomah County, 596 F.3d 1007, 1013 (9th Cir. 2010).
V. CONCLUSION

The FMLA entitles covered employees to take job protected leave when they are ill or to care for new or sick family members. If an employer violates the FMLA, one of the remedies allowed by the statute is a front pay award. While all the circuits agree that only the judge can determine when a front pay award is available, the circuits are split on whether a judge or jury should determine the amount of the award. Most of the circuits have determined that the judge should calculate the amount of the award, while only the Sixth Circuit differs, holding that the jury should calculate the amount.

The FMLA is unique when compared to other employment related statutes because its primary purpose is job protection. Because the award of front pay directly correlates with this purpose, it is even more important that a judge calculates the amount of the front pay award. Additionally because front pay is an equitable remedy and because of the variety of factors that must be considered when calculating front pay, it is necessary that the award be calculated by one who is knowledgeable and experienced. This can only be guaranteed when a judge calculates the award. While judges should be encouraged to use advisory juries when calculating the amount of the award, the judge is the only one who should have the final decision on the amount of the front pay award.