JOURNAL OF LAW, ECONOMICS & POLICY

VOLUME 13 | **WINTER 2017** | **NUMBER 1**

EDITORIAL BOARD
2016-2017

Blake Browning
Editor-in-Chief

Doan Phan
Executive Editor

Peter Zuk
Publications Editor

Colleen Conners
Managing Editor

Gina Lee
Senior Articles Editor

Abby Chestnut
Senior Notes Editor

David Showalter
Senior Research Editor

Dimitri Hasandras
Articles Selection Editor

Erin Gallagher
Symposium Editor

MEMBERS

Julian Flamant
Erich Grome
Steven Horn
Allison Curry
Lindsey Edwards
Dodd Fuller
Anshika Mehrotra
Christian Nee
Easton Saltsman
Kyle Spencer
Bonnie Kelly
Ray Samara
Brian Silver
Samantha Trussell

BOARD OF ADVISORS

Lisa E. Bernstein
Judge Guido Calabresi
Robert D. Cooter
Richard A. Epstein
Mark F. Grady
Bruce H. Kobayashi
A. Douglas Melamed
Eric Posner
Roberta Romano
Steven M. Shavell
Vernon L. Smith
Thomas S. Ulen
Henry N. Butler
Lloyd R. Cohen
Robert C. Ellickson
Judge Douglas H. Ginsburg
Michael S. Greve
Francesco Parisi
Judge Richard A. Posner
Hans-Bernd Schafer
Henry E. Smith
Gordon Tullock
W. Kip Viscusi
Todd J. Zywicki

JOURNAL OF LAW, ECONOMICS & POLICY

VOLUME 13 **WINTER 2017** **NUMBER 1**

CONTENTS

ARTICLES

1 WHAT THEORY AND THE EMPIRICAL EVIDENCE TELL US ABOUT PROXY ACCESS
Bernard S. Sharfman

39 THE SENSE IN COASE'S CRITICISM OF PIGOU: THE *CETERIS PARIBUS* CASE FOR INTERVENTION
David Campbell

55 TOWARDS A COHERENT AND WORKABLE ANTITRUST POLICY ON VERTICAL RESTRAINTS
Daniel J. Gifford & Robert T. Kudrle

STUDENT COMMENTS

91 ILLUMINATING THE OFF-LABEL FABLE: HOW OFF-LABEL PROMOTION MAY ACTUALLY HELP PATIENTS
Colleen Conners

113 REDEFINING LINCOLN'S LAW: HOW TO SHAPE THE THEORY OF IMPLIED CERTIFICATIONS POST-*ESCOBAR*
Doan Phan

WHAT THEORY AND THE EMPIRICAL EVIDENCE TELL US ABOUT PROXY ACCESS

Bernard S. Sharfman

INTRODUCTION

Traditionally, the default rules of corporate and securities law have provided a publicly traded company's board of directors with exclusive authority to decide whether shareholder proposals on proxy access—the ability of certain privileged shareholders to have their own slate of director nominees included in the company's proxy solicitation materials for purposes of voting at the annual meeting—are to be included in the company's proxy solicitation materials. However, the U.S. Securities and Exchange Commission (SEC) has recently amended its rules to allow such proposals to be included whether or not the board approves. That change has resulted in a new issue that now needs to be addressed: What are the consequences of potentially providing a small group of privileged shareholders, in addition to the board,[1] the power to decide which nominees for election to the board of directors are to be included in a public company's proxy solicitation materials (the proxy statement and proxy voting card)?[2]

During the 2015 proxy season, the Office of the Comptroller of New York City (comptroller), the custodian and investment adviser to the New York City Pension Funds, submitted seventy-five of the 108 proxy access proposals that were received by publicly traded companies.[3] From the comptroller's perspective, the effort was successful. Of the seventy-five precatory proposals the comptroller submitted,[4] sixty-six went to a

---

[1] According to Stephen Bainbridge, "There is no more basic question in corporate governance than 'who decides.'" Stephen M. Bainbridge, *Executive Compensation, Who Decides?*, 83 TEX. L. REV. 1615, 1650 (2005) (in the context of the board versus shareholders).

[2] For purposes of this study, a *public company* can be defined as a for-profit corporation that is publicly traded on a national exchange or over the counter but does not have a controlling shareholder.

[3] R.J. LEHMANN, R STREET POLICY INSTITUTE, R STREET POLICY STUDY NO. 38, PROXY ACCESS: SHAREHOLDER DEMOCRACY OR CREEPING MERCANTILISM 2 (2015), http://www.rstreet.org/policy-study/proxy-access-shareholder-democracy-or-creeping-mercantilism/.

[4] Precatory proxy access proposals are preferable because of the significant challenges involved in drafting a binding bylaw and at the same time trying to make sure it does not exceed Rule 14a-8's 500-word limit. *See*, *Proxy Access and Advance Notice Bylaws in the Wake of Invalidation of the SEC's Proxy Access Rule: An Approach to Private Ordering*, LATHAM & WATKINS LLP CORP. GOVERNANCE COMMENT, Nov. 2011, at 5, https://www.lw.com/upload/pubContent/_pdf/pub4437_1.pdf. Also, it has been suggested that a precatory proposal will garner more votes than a mandatory proposal because shareholders will take a precatory proposal less seriously. *See* Lawrence Hamermesh, *Precatory Proxy*

shareholder vote, with 55% average support.[5] Of those sixty-six, forty-one received majority support.[6] Moreover, at six companies where withdrawal of the proxy access proposal was negotiated, management agreed to adopt proxy access or put forward a management-sponsored proposal next year.[7] Overall, 100 companies faced nonbinding proxy access proposals in 2015, with sixty gaining the majority support of shareholders.[8] Moreover, 115 companies in 2015 adopted a binding proxy access bylaw.[9] As a result, 117 of the S&P 500 companies (21% of the index) now have a binding proxy access bylaw in place.[10] It has been estimated that another 200 companies will face proxy access proposals in 2016,[11] including thirty-six by the comptroller and forty by California State Teachers' Retirement System.[12]

As nonbinding proxy access proposals gain traction with shareholders and a number of boards begin to adopt binding proxy access bylaws in response to shareholder pressure, it may be only a matter of time before the SEC puts universal proxy access back on its agenda.[13] Universal proxy access, a recurring topic of the SEC focus for more than seventy years, would automatically allow certain privileged shareholders to place their nominees for the board into the proxy solicitation materials of almost all public companies without the need for a charter amendment or bylaw.

The possibility that the SEC will renew its interest in universally mandated (universal) proxy access was signaled by a recent empirical study of proxy access by staff economists in the SEC's Division of Economic and Risk Analysis. The study stated the issue as follows: "The fundamental question is whether private market forces, through the shareholder proposal process, would be able to realize (and perhaps surpass) the enhancements in

---

Access Proposals, INST. DEL. CORP. & BUS. L. (Nov. 15, 2011), http://blogs.law.widener.edu/delcorp/2011/11/15/precatory-proxy-access-proposals/#sthash.zXGjV6Qg.dpbs.

5 ACTIVIST INSIGHT, ACTIVIST INVESTING: AN ANNUAL REVIEW OF TRENDS IN SHAREHOLDER ACTIVISM 40 (Josh Black ed., 2016), http://www.srz.com/The_Activist_Investing_Annual_Review_2016/.

6 *Id.*

7 Barry B. Burr, *Board Support Rising but Concerns Remain, Study Says*, PENSIONS AND INVESTMENTS (July 27, 2015), http://www.pionline.com/article/20150727/PRINT/307279997/board-support-rising-but-concerns-remain-study-says.

8 Joe Cahill, *Four Companies Doing the Right Thing for Shareholders*, CRAIN'S CHICAGO BUSINESS (Nov. 25, 2015), http://www.chicagobusiness.com/article/20151125/ISSUE10/151129911/four-companies-doing-the-right-thing-for-shareholders.

9 Che Odom, *NYC Pension Funds, CalPERS Prep for Proxy Access Blitz*, BLOOMBERG BNA (Jan. 28, 2016), http://www.bna.com/nyc-pension-funds-n57982066676/.

10 ACTIVIST INSIGHT, *supra* note 5, at 40.

11 Cahill, *supra* note 8.

12 ACTIVIST INSIGHT, *supra* note 5, at 41.

13 Cydney Posner, *Is the SEC Considering Reproposing Mandatory Proxy Access Rules?*, PUBCO@COOLEY (Aug. 3, 2015, 3:57 PM), http://cooleypubco.com/2015/08/03/is-the-sec-considering-reproposing-mandatory-proxy-access-rules/.

shareholder value that could result from universally-mandated proxy access."[14]

This statement is surprising, because it relies on the premise that both shareholder-initiated proxy access and universal proxy access are superior to the historical approach of board-initiated proxy access in terms of shareholder wealth enhancement and firm performance. However, that premise has not been empirically verified and, as argued in this study, is incorrect. Another signal is a recent study done by the CFA Institute.[15] The objective of the study was to address the issues raised by the D.C. Court of Appeals in *Business Roundtable v. SEC*[16] when it vacated the SEC's universal proxy access rule in 2011 and to thereby encourage the SEC to revisit universal proxy access.[17] Finally, the Council of Institutional Investors, a nonprofit association representing the interests of public pension funds and labor union–related entities, continues to promote the idea of universal proxy access through its 2015 policy statement "that proxy access is a *fundamental* right of *long-term* shareowners."[18]

This study makes three primary arguments. First, the SEC's current regime of proxy access, by no longer allowing companies to exclude shareholder proposals on proxy access from their proxy solicitation materials,[19] should not be understood as an enhancement to the "private ordering" of a company's governance arrangements. Rather, this regime acts as a federal barrier to the more efficient approach of board-initiated proxy access. Therefore, this study recommends that the SEC return to its traditional approach to proxy access, allowing a board to omit shareholder proposals on proxy access from a company's proxy materials at its discretion. Second, the superiority of board decision-making in the context of proxy access creates a presumption that universal proxy access is an inefficient and unnecessary means of nominating and electing directors.

---

[14] Tara Bhandari et al., *Public versus Private Provision of Governance: The Case of Proxy Access* (U.S. Sec. & Exch. Comm'n, Working Paper, July 24, 2015), http://www.sec.gov/dera/staff-papers/working-papers/public-vs-private-provision-of-governance.pdf.

[15] Chiara Trabucchi et al., *Proxy Access in the United States: Revisiting the Proposed SEC Rule,* CFA INSTITUTE (2014), http://www.cfapubs.org/doi/pdf/10.2469/ccb.v2014.n9.1. For a critique of this study, *see* BERNARD. S. SHARFMAN, R STREET INSTITUTE, R STREET SHORTS NO. 2, CRITIQUING THE CFA INSTITUTE'S REPORT ON PROXY ACCESS 1 (2016), http://www.rstreet.org/wp-content/uploads/2016/03/RSTREETSHORT21.pdf (finding that the report's shortcomings should disqualify it "from being used as support for mandatory proxy access; for shareholder proposals on proxy access; for board discussions about whether a proxy-access bylaw should be implemented; and, perhaps most importantly, for board discussions about whether a proxy-access bylaw needs to be rescinded").

[16] Bus. Roundtable v. SEC, 647 F.3d 1144 (D.C. Cir. 2011).

[17] Trabucchi et al., *supra* note 15, at 1.

[18] COUNCIL OF INSTITUTIONAL INVESTORS, PROXY ACCESS: BEST PRACTICES 2 (2015) (emphasis added), http://www.cii.org/files/publications/misc/08_05_15_Best%20Practices%20-%20Proxy%20Access.pdf.

[19] 17 C.F.R. § 240.14a-8(i)(8) (2011).

Third, that presumption can be rebutted with empirical evidence that consistently shows, at a high level of statistical significance, that universal proxy access is wealth-enhancing for shareholders. That premise is required as the null hypothesis to be tested and can be stated as follows: the "preservation of managerial discretion" in the nomination of directors is wealth-enhancing for shareholders. However, the empirical evidence does not currently exist to reject the null hypothesis. As a result, it would be reasonable for the SEC to keep universal proxy access off its agenda.

The issue of proxy access must also be understood in the larger context of shareholder empowerment (the shifting of decision-making authority from the board of directors and executive management to shareholders).[20] Proxy access is clearly the corporate governance arrangement that is the current focus of those who advocate for shareholder empowerment. The problem with shareholder empowerment is that it tries to shift the balance between authority and accountability too far in the direction of accountability without proper theoretical or empirical justification. Rather, the balance should be heavily weighted toward authority for a public company (as subsequently discussed) to make the most efficient decisions.[21] Moreover, shareholder empowerment takes a one-size-fits-all approach without taking into consideration that all firms are different and that the optimal corporate governance arrangements at one company will not necessarily be the same at another.[22] Finally, there is no end in sight for shareholder empowerment.[23] The trend is toward "*creeping shareholder activism*, a constant movement toward shareholder empowerment without regard for what is lost in the process in terms of efficient decision making."[24] A line must be drawn somewhere, and proxy access is as good a place as any to start pushing back against the negative aspects of the shareholder empowerment movement.

The discussion that follows, when it references state corporate law, has been pragmatically framed in the context of Delaware corporate law. Delaware is the state where the majority of the largest U.S. companies are incorporated,[25] and its corporate law often serves as the authority that other

---

[20] Bernard S. Sharfman, *What's Wrong with Shareholder Empowerment?*, 37 IOWA J. CORP. L. 903, 903 (2012).

[21] *Id.*

[22] *Id.* at 907.

[23] *Id.* at 908.

[24] *Id.* (emphasis added).

[25] *See* LEWIS S. BLACK, JR., WHY CORPORATIONS CHOOSE DELAWARE 1 (2007), corp.delaware.gov/whycorporations_web.pdf (stating that Delaware is the "favored state of incorporation for U.S. businesses"). According to the State of Delaware website, Delaware is the legal home to "[m]ore than 50% of all publicly-traded companies in the United States including 64% of the Fortune 500." *About Agency*, STATE OF DELAWARE, http://corp.delaware.gov/aboutagency.shtml (last visited Oct. 15, 2014).

U.S. states look to when developing their own statutory and case law.[26] Therefore, the primary examples are from Delaware, but the study is meant to be global in nature.

In Part I, this study describes the SEC's current regime of proxy access. Part II discusses how the SEC's current regime of proxy access harms private ordering. That discussion leads to a recommendation that the SEC's current regime of proxy access be rescinded. Part III discusses universal proxy access. That section takes the perspective of an impartial SEC commissioner who is trying to decide whether to implement universal proxy access. This decision is difficult for a commissioner and the commission as a whole to make because the issue of universal proxy access resides in the world of corporate governance, not in securities regulation and its focus on disclosure. Moreover, the world of corporate governance arrangements rests on the foundation of state corporate law—again, not an area of expertise for the commission—and on the private ordering approach to such arrangements. If such an approach can be understood to be generally wealth-enhancing for shareholders, then a commissioner must tread very carefully when considering implementing universal proxy access. Such careful consideration requires that a vote for universal proxy access be supported by empirical evidence that consistently shows (from study to study) its value over time and that is not simply supportive of proxy access at any one point in time. This section argues that such empirical evidence does not currently exist. The final part concludes by summarizing this study's findings and recommendations.

I. THE SEC'S CURRENT REGIME OF PROXY ACCESS

From 1947 to 2011, the SEC's proxy rules on shareholder proposals allowed a public company to exclude any proposal from a company's proxy solicitation materials that related "to an election for membership on the company's board of directors or analogous governing body."[27] At a minimum, that exclusion meant that shareholders could not place their nominees in a public company's proxy solicitation materials without first getting the approval of the board. Notably, that requirement meant that universal proxy access was not allowed. Moreover, for most of that lengthy time period, with one significant exception,[28] the SEC also interpreted the

---

[26] *See* Nadelle Grossman, *Director Compliance with Elusive Fiduciary Duties in a Climate of Corporate Governance Reform*, 12 FORDHAM J. CORP. & FIN. L. 393, 397 (2007).

[27] AFSCME v. Am. Int'l Group, Inc., 462 F.3d 121(2d Cir. 2006) (citing 17 C.F.R. § 240.14a-8(i)(8)). The SEC's source of authority for this long-standing exclusion comes from Section 14(a) of the Securities Exchange Act of 1934. 15 U.S.C. § 78n-1 (2012).

[28] During the 1980s, the SEC denied several requests for no-action letters where a company was trying to exclude a proxy access proposal dealing with "procedural rules applying prospectively to future elections to the board." Michael E. Murphy, *The Nominating Process for Corporate Boards of*

relevant rule to mean that a board could exclude a shareholder proposal to establish a procedure by which certain shareholders could include their shareholder nominees in a company's proxy solicitation materials. Therefore, the implementation of proxy access has mainly existed under a corporate governance arrangement that can be referred to as *board-initiated proxy access*. Although it is subsequently argued in this study that the board, as well as shareholders, did not historically have the authority under Delaware corporate law to initiate proxy access through a bylaw, at the very least the board could initiate proxy access through the charter amendment process.

In practice, boards have resisted using their authority to initiate proxy access. Instead, the nomination of directors has been under the control of the board and its nominating committee. In 2006, it was reported that 99% of companies in the S&P 500 Index used a nominating committee.[29] Use of the committee meant that only candidates who had been screened and approved by the committee, with or without full board approval, would appear in the company's proxy solicitation materials for purposes of electing directors at the annual meeting.[30] Because shareholders could not place their slate of nominees in the company's proxy materials, the only alternative was to go through the cost-prohibitive process of entering into a proxy contest by creating their own proxy materials to nominate their slate. Therefore, board nominees—listed in the proxy materials and helped by the advantages of plurality voting—were always assured of winning an election except in those elections in which a proxy contest had been initiated.

However, in 2011, a dramatic change occurred in the way in which the SEC approached proxy access. By using authority granted to it by Section

---

Directors: A Decision-Making Analysis, 5 BERKELEY BUS. L.J. 131, 140 (2008) (citing Unicare Services, Inc., SEC No-Action Letter, 1980 SEC No-Act LEXIS 3289 (May 13, 1980); Mobil Corp., SEC No-Action Letter, 1981 LEXIS 3208 (Mar. 3, 1981); Union Oil Co. of Cal., SEC No-Action Letter, 1981 SEC No-Act. LEXIS 3001 (Jan. 29, 1981); Chittenden Corp., SEC No-Action Letter, 1987 SEC No-Act. LEXIS 1955 (Mar. 10, 1987)).

However, by the beginning of 1990, the SEC staff had changed its approach, allowing such exclusions. *Id.* (citing Bank of Boston, SEC No-Action Letter, 1990 SEC No-Act. LEXIS 206 (Jan. 26, 1990); Unocal Corp., SEC No-Action Letter, 1990 SEC No-Act. LEXIS 183 (Feb. 6, 1990); Amoco Corp., SEC No-Action Letter, 1990 SEC No-Act. LEXIS 242 (Feb. 14, 1990); Thermo Electron Corp., SEC No-Action Letter, 1990 SEC No-Act. LEXIS 549 (Mar. 22, 1990)). This hard-line approach was successfully challenged on grounds that the SEC never provided sufficient reasons for changing its position. *See* AFSCME v. Am. Int'l Group, Inc., 462 F.3d 121,129 (2d Cir. 2006). In response, the SEC quickly provided a release codifying the reasoning for the change by modifying the language of Section 14a-8(i)(8) to make it clear that the exclusion also applies to a shareholder proposal seeking a proxy access process. *See* Securities Act Release No. 34-56914 (2008), https://www.sec.gov/rules/final/2007/34-56914.pdf.

[29] *See* Murphy, *supra* note 28, at 147.

[30] *See* DEL.CODE ANN. tit. 8, § 141(c)(2) (West 2016).

971 of the recently enacted Dodd-Frank Act,[31] the SEC was able to modify Rule 14a-8(i)(8), the election exclusion rule, so that public companies could no longer exclude precatory or binding shareholder proposals on proxy access from their proxy solicitation materials.[32] The result is that there is now what can be referred to as *shareholder-initiated proxy access* in addition to board-initiated proxy access.

II. HOW THE SEC'S CURRENT REGIME OF PROXY ACCESS HARMS PRIVATE ORDERING

On its face, shareholder-initiated proxy access appears to be value-enhancing, because it provides a low-cost alternative to a proxy contest for certain shareholders seeking to get their slate of nominees voted on at an annual meeting. Yet proxy access can be value-enhancing for the corporation and shareholders only if proxy access actually enhances the private ordering of corporate governance and does not detract from it. As argued in this section, the SEC's current regime detracts from private ordering by its forced inclusion of shareholder proposals on proxy access in a company's proxy solicitation materials without providing the board the option to deny such inclusion. This argument is meant to correct the misunderstanding that the SEC's current regime of proxy access enhances the private ordering of corporate governance arrangements.[33] In essence, the SEC's current regime of proxy access creates a mandatory rule that overrides the approach to the private ordering of corporate governance arrangements that has traditionally existed under corporate law, board-initiated private ordering. The result is a less efficient decision-making process.

A. *What is Private Ordering under Corporate Law?*

According to Michael Jensen and William Meckling, "[c]ontractual relations are the essence of the firm, not only with employees but with

---

31 Section 971 of the Dodd-Frank Act provides the SEC with undisputed authority to promulgate proxy access rules as long as such rules can be justified on the grounds that they are "in the interests of shareholders and for the protection of investors." Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 971, 124 Stat. 1376, 1915, (2010).

32 17 C.F.R. § 240.14a-8(i)(8).

33 *See* Troy A. Paredes, Comm'r, U.S. Sec. & Exch. Comm'n, Statement at Open Meeting to Propose Amendments Regarding Facilitating Shareholder Director Nominations, (May 20, 2009), http://www.sec.gov/news/speech/2009/spch052009tap.htm; Joseph A. Grundfest, *The SEC's Proposed Proxy Access Rules: Politics, Economics, and the Law*, 65 BUS. LAW. 361, 376 (2010) (concurring with Paredes).

suppliers, customers, creditors, etc."[34] Most famously, Jensen and Meckling describe an organization, such as a public company, as a legal fiction that serves "as a nexus for a set of contracting relationships among individuals."[35] As explained by Jonathan Macey, "because firms consist of a complex web of contractual relationships, firm behavior depends critically on what those contracts provide. In turn, the contract provisions themselves depend on the outcome of the *bargaining process* that takes place between the contracting parties."[36]

Because corporate law primarily provides default, not mandatory, rules, this contractarian theory of the firm can also be applied to a firm's governance arrangements. Private ordering under corporate law is implemented through a process of creating, modifying, and repealing bylaws and charter amendments.[37] Private ordering is considered efficient because it allows for the implementation of market-driven corporate governance arrangements.[38] That is, it "allows the internal affairs of each corporation to be tailored to its own attributes and qualities, including its personnel, culture, maturity as a business, and governance practices."[39] In effect, "observed governance choices are the result of value-maximizing contracts between shareholders and management."[40]

However, private ordering is not a purely theoretical construct under corporate law. It is a structured approach that purposefully selects the board to take the lead in determining the optimal corporate governance arrangements. According to Michael Klausner, "[t]he contractarian theory of the firm . . . implies a theory of the role of corporate law: corporate law should merely provide a set of default rules that managers may adopt on

---

34 Michael C. Jensen & William H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305, 310 (1976).

35 *Id.*

36 Jonathan R. Macey, *Fiduciary Duties as Residual Claims: Obligations to Nonshareholder Constituencies from a Theory of the Firm Perspective*, 84 CORNELL L. REV. 1266, 1272 (1999) (emphasis added).

37 Jill E. Fisch, *Leave It to Delaware: Why Congress Should Stay Out of Corporate Governance,* 37 DEL. J. CORP. L. 731, 743 n.80 (2013).

38 According to Jonathan Macey,

> [B]ecause informal norms generate outcomes that are generally welfare-enhancing, while law *at best* generates outcomes that are mixed (and tend strongly towards the welfare-reducing), informal norms should come with a strong presumption of legitimacy. Formal legal rules are likely to be inefficient at best and amorally redistributive at worst. Thus, under a wide range of circumstances, such as when society is interested in maximizing utilitarian considerations, and when society is interested in resolving standard legal disputes within groups, lawmakers are unlikely to improve upon the customary rules the group develops through voluntary, private interaction.

Jonathan R. Macey, *Public and Private Ordering and the Production of Legitimate and Illegitimate Legal Rules*, 82 CORNELL L. REV. 1123, 1140-41 (1997) (emphasis added).

39 Paredes, *supra* note 33.

40 David F. Larcker et al., *The Market Reaction to Corporate Governance Regulation*, 101 J. FIN. ECON. 431, 431 (2011).

behalf of their firms, while leaving managers free to customize their companies' charters with legally enforceable rights and obligations."[41] That board-initiated private ordering approach permeates the thinking of the Delaware courts. For example, consider Leo Strine's discussion of the board's ability to unilaterally adopt a bylaw in *Boilermakers Local 154 Retirement Fund v. Chevron Corp.*:

> As our Supreme Court has made clear, the bylaws of a Delaware corporation constitute part of a binding broader contract among the directors, officers, and stockholders formed within the statutory framework of the DGCL [Delaware General Corporate Law]. This contract is, by design, flexible and subject to change in the manner that the DGCL spells out and that investors know about when they purchase stock in a Delaware corporation. The DGCL allows the corporation, through the certificate of incorporation, to grant the directors the power to adopt and amend the bylaws unilaterally.[42]

As a demonstration of how corporate law has supported board-initiated private ordering over an extensive period of time, Chancellor (now Chief Justice) Strine cited as his authority two Delaware Supreme Court opinions issued eighty years apart.[43]

B. *The Board's Control of Private Ordering*

Board-initiated private ordering of governance arrangements is an application of the most important default rule under corporate law,[44] the rule that provides the board with ultimate decision-making authority. For example, under Delaware corporate law, "[t]he business and affairs of every corporation organized under this chapter shall be managed by or under the direction of a board of directors, except as may be otherwise provided in this chapter or in its certificate of incorporation."[45] On its face, that statutory rule provides the board with unlimited managerial authority.

---

[41] Michael Klausner, *The Contractarian Theory of Corporate Law: A Generation Later*, 31 J. CORP. L. 779, 780 (2006).

[42] Boilermakers Local 154 Ret. Fund v. Chevron Corp., 73 A.3d 934, 939 (Del. Ch. 2013) (citing DEL. CODE ANN. tit. 8, § 109(a) (West 2016)).

[43] Airgas, Inc. v. Air Prods. & Chems., Inc., 8 A.3d 1182, 1188 (Del. 2010) and Lawson v. Household Fin. Corp., 152 A. 723, 726 (Del. 1930).

[44] Although default rules can be modified, "the default rule is tailored toward what the legislature believes most, but not all, of an organization's stakeholders would agree to if contracting were efficient." James D. Cox, *Corporate Law and the Limits of Private Ordering*, 93 WASH. U. L. REV. 257, 261 (2015) (*citing* Lucian A. Bebchuk, *Limiting Contractual Freedom in Corporate Law: The Desirable Constraints on Charter Amendments*, 102 HARV. L. REV. 1820, 1847 (1989)).

[45] DEL. CODE ANN. tit. 8, § 141(A) (WEST 2016).

Public companies never substantively modify the default rule,[46] and the lack of modification via charter amendment needs to be acknowledged as the first and most fundamental step in such a company's private ordering process. The default rule is so universally accepted that it most likely could have been written as a mandatory rule without restricting the contracting parties' abilities to enter into private ordering.[47] That is, if a bargaining process truly goes on between contracting parties in a public company, then there seems to be overwhelming support for allowing the board to retain ultimate decision-making authority. Conversely, if the contracting parties wanted to implement shareholder empowerment to enhance decision-making efficiency, one would expect that at least some public companies would have charter provisions that substantively weaken board authority.

Superior decision-making efficiency is the rationale that explains why the outcome of the bargaining process always allows Section 141(a) of the Delaware General Corporate Law (DGCL) to be incorporated without substantive modification into a public company's charter and that by extension allows the board to control the private ordering of corporate governance arrangements, including proxy access. Corporate law concentrates ultimate decision-making authority in the board because lawmakers recognize that a centralized, hierarchical authority is necessary for the successful management of a public company that can become extremely large in size. According to Robert Clark, hierarchies in large organizations lead to the "facilitation of cooperation in the carrying out of large-scale tasks."[48] According to Kenneth Arrow, information scattered over a large organization must be both filtered and transmitted to a centralized authority for a large organization to make informed decisions and minimize error in decision-making.[49]

Armen Alchian and Harold Demsetz argue that a centralized authority is necessary to eliminate the problems associated with having a large number of shareholders:

> If every stock owner participated in each decision in a corporation, not only would large bureaucratic costs be incurred, but many would shirk the task of becoming well informed on the issue to be decided, since the losses associated with unexpectedly bad decisions will be borne in large part by the many other corporate shareholders. More effective control of corporate activity is achieved for most purposes by transferring decision authority to a

---

46 *Id. See also* Aronson v. Lewis, 473 A.2d 805, 811 (Del. 1984) ("A cardinal precept of the General Corporation Law of the State of Delaware is that directors, rather than shareholders, manage the business and affairs of the corporation." (citing DEL. CODE ANN. tit. 8, § 141(A) (West 2016)).

47 Bernard S. Black, *Is Corporate Law Trivial? A Political and Economic Analysis*, 84 NW. U.L. REV. 542, 551 (1990).

48 ROBERT CHARLES CLARK, CORPORATE LAW 801 (1986).

49 KENNETH J. ARROW, THE LIMITS OF ORGANIZATION 68–70 (1974).

> smaller group, whose main function is to negotiate with and manage (renegotiate with) the other inputs of the team.[50]

In a public company, the board has a clear decision-making advantage over shareholders. As observed by Michael Dooley, for companies with a large number of shareholders, it is much more efficient for the board—the corporate actor that possesses overwhelming advantages in terms of information, including nonpublic information—to make corporate decisions than for shareholders or any other party that contracts with the corporation to do so.[51] In sum, what is desired by the contracting parties in terms of decision-making can be summarized in the following statement by Stephen Bainbridge: "[pr]eservation of managerial discretion should always be the null hypothesis."[52]

Indeed, the presumption that the board provides the corporation with superior decision-making is endorsed by the courts through the explanation of why courts apply the business judgment rule ("a presumption that in making a business decision, the directors of a corporation acted on an informed basis, in good faith and in the honest belief that the action taken was in the best interests of the company"):[53]

> The "business judgment" rule is a judicial creation that presumes propriety, under certain circumstances, in a board's decision. Viewed defensively, it does not create authority. In this sense the "business judgment" rule is not relevant in corporate decision making until after a decision is made. It is generally used as a defense to an attack on the decision's soundness. The board's managerial decision making power, however, comes from § 141(a). The judicial creation and legislative grant are related because the "business judgment" rule evolved to give recognition and deference to directors' business expertise when exercising their managerial power under § 141(a).[54]

The implementation of corporate governance arrangements is heavily influenced by that deference to board decision-making. Therefore, it should not be surprising that the board has sole discretion to initiate changes to the corporate charter, even though shareholder approval would be required.[55] Moreover, as already discussed, the board may unilaterally create its own bylaws if the charter provides it with such authority.[56]

---

50 Armen A. Alchian & Harold Demsetz, *Production, Information Costs, and Economic Organization*, 62 AM. ECON. REV. 777, 788 (1972).

51 Michael P. Dooley, *Two Models of Corporate Governance,* 47 BUS. LAW. 461, 467 (1992) (citing to ARROW, *supra* note 49, at 68-70).

52 Stephen M. Bainbridge, *The Business Judgment Rule as Abstention Doctrine*, 57 VAND. L. REV. 83, 109 (2004).

53 Smith v. Van Gorkom, 488 A.2d 858, 872 (Del. 1985) (citing Aronson v. Lewis, 473 A.2d 805, 812 (Del. 1984)).

54 Zapata Corp. v. Maldonado, 430 A.2d 779, 782 (Del. 1981).

55 DEL. CODE ANN. tit. 8, § 242(b)(1) (West 2014).

56 *Id.* § 109(a).

C. *The Role of Shareholders in Corporate Governance Arrangements*

Even though the board is provided the lion's share of authority to initiate changes in a public company's corporate governance arrangements, shareholders do have a role to play beyond approving charter amendments. Shareholders may propose and approve binding bylaws.[57] However, those bylaws, as well as bylaws in general, will not survive a legal challenge if they interfere with the board's substantive decision-making authority under DGCL Section 141(a).[58] According to the Delaware Supreme Court in *CA, Inc.*, "[i]t is well-established Delaware law that a proper function of bylaws is not to *mandate* how the board should decide specific substantive business decisions, but rather, to define the process and procedures by which those decisions are made."[59] Therefore, the threshold question for the legality of shareholder-initiated bylaws is "whether the Bylaw is one that establishes or regulates a process for substantive director decision-making, or one that mandates the decision itself."[60] For example, in *Gorman v. Salamone*, the Delaware Chancery Court invalidated a binding shareholder-initiated bylaw approved by written consent of the shareholders that would have granted shareholders the substantive decision-making authority of removing corporate officers without cause.[61] According to the court in *Gorman v. Salamone*:

> Valid bylaws focus on process, and "[w]hether or not a bylaw is process-related must necessarily be determined in light of its context and purpose." The Court may look to the intent and effect of a bylaw to determine whether it is a proper subject for stockholder action; "even facially procedural bylaws can unduly intrude upon board authority."[62]

In sum, unlike charter amendments, shareholder bylaws, when under court review, will be closely scrutinized to see whether they encroach on the substantive decision-making of the board.

D. *Proxy Access under Corporate Law*

As the Delaware Chancery Court said in *Gorman v. Salamone*, "even facially procedural bylaws can unduly intrude upon board authority."[63] That likelihood is the problem with shareholder proposals on proxy access,

---

57 *Id.*
58 CA, Inc. v. AFSCME Emp. Pension Plan, 953 A.2d 227, 234-37 (Del. 2008).
59 *Id.* at 234-35 (emphasis added).
60 *Id.* at 235.
61 Gorman v. Salamone, No. 10183-VCN, 2015 WL 4719681, at *5 (Del. Ch. July 31, 2015).
62 *Id.* (quoting *CA, Inc.*, 953 A.2d at 236-37).
63 *Id.*

whether they are crafted as binding or precatory bylaws. The proposals in and of themselves request that the board be relieved of some of its authority to participate in the director nomination process without a charter amendment to authorize such a change as required under DGCL Section 141(a). Indeed, proxy access, the ultimate objective of such shareholder proposals, allows for the placing of director nominees into the company's proxy solicitation materials without review and approval by the board nominating committee, the part of the corporation that is in the best position to determine which nominees are the most qualified candidates to serve as directors:

> The Board nominating committee has an informational advantage over even the most informed shareholders because of the inside information it has on how the current board interacts with each other and executive officers, expectations on how a particular nominee will meld with other board members and executive officers, and the needs of the corporation in terms of directors, based on both public and confidential information. Shareholders who want to take advantage of proxy access do not have this information available to them.[64]

Allowing proxy access undercuts the informational advantage held by "the nominating committee by failing to assign it any role in screening or approving shareholder nominations."[65] Most importantly, boards would be forced to abdicate their fiduciary duties because they would not be given the opportunity to deny placing shareholder nominees into the proxy solicitation materials after a review of their qualifications, the requirements of the company, how the candidates would interact with other board members, and so on. According to the Delaware Supreme Court in *CA, Inc.*, a bylaw cannot "violate the prohibition, which our decisions have derived from Section 141(a), against contractual arrangements that commit the board of directors to a course of action that would preclude them from fully discharging their fiduciary duties to the corporation and its shareholders."[66] In *CA, Inc.*, the court invalidated a bylaw that mandated "reimbursement of election expenses in circumstances that a proper application of fiduciary principles could preclude."[67] Proxy access creates an analogous situation as the proper application of the board's fiduciary duties could preclude the inclusion of shareholder nominees into a company's proxy solicitation materials.

But that is not the end of the story. Those shareholders who disagree with the invalidation of a bylaw "have two alternatives. They may seek to amend the Certificate of Incorporation to include the substance of the

---

64 Bernard S. Sharfman, *Why Proxy Access Is Harmful to Corporate Governance*, 37 J. CORP. L. 387, 402 (2012).

65 Murphy, *supra* note 28, at 144.

66 *CA, Inc.*, 953 A.2d at 238.

67 *Id.* at 240.

Bylaw; or they may seek recourse from the Delaware General Assembly."[68] In essence, the latter is what happened. A year after the decision in *CA, Inc.*, DGCL Section 112 was made part of Delaware's statutory corporate law.[69] Section 112 allows proxy access to be implemented through a shareholder bylaw proposal or board-created bylaw, no longer requiring a charter amendment as argued earlier. The legislature passed this legislation—despite the precedence that makes providing shareholders with such authority an outlier under Delaware corporate law—clearly in response to the prospect of SEC-mandated proxy access and not in an attempt to overrule the judicial approach taken in *CA, Inc.* According to a comment letter from the Corporate Law Section of the Delaware Bar Association to the SEC, there was no need for universal proxy access via Rule 14a-11 given the recent passage of DGCL Section 112.[70] Most importantly, whether intended or not, Section 112 extinguished any concerns that if the SEC eliminated the ability of boards to exclude shareholder bylaw proposals on proxy access, as it ultimately did, doing so would be deemed a violation of Delaware law and therefore could still be excludable under SEC's Rule 14a-8(i)(2).[71]

E. *The Problem of Shareholder Opportunism and Proxy Access*

The intrusion of the SEC's current regime of proxy access into the private ordering of corporate governance arrangements also implicates shareholder opportunism. While the board nominating committee and the board as a whole must adhere to their fiduciary duties of care and loyalty, such duties are not required of shareholders who submit proposals on proxy access or participate in the nomination of directors under a binding proxy access bylaw. Therefore, proxy access may allow certain shareholders, such as public pension funds and labor union–related entities, the ability to act opportunistically if they deem it to their advantage. As discussed below, several recent reports indicate that certain shareholders may be using shareholder proposals, including proposals on proxy access, as a means to act opportunistically.

In a paper sponsored by the Manhattan Institute, Tracie Woidtke examines the relationship between public pension funds engaged in

---

68 *Id.*

69 DEL. CODE ANN. tit. 8, § 112 (West 2009).

70 Letter from Del. State Bar Ass'n to Elizabeth M. Murphy, Sec'y, SEC, on facilitating shareholder director nominations (July 24, 2009), http://www.sec.gov/comments/s7-10-09/s71009-65.pdf. *See also* Mark J. Roe, *The Corporate Shareholder's Vote and Its Political Economy, in Delaware and in Washington*, 2 HARV. BUS. L. REV. 1, 18-20 (2012).

71 17 C.F.R. § 240.14a-8(i)(2) (This rule allows a public company to exclude a shareholder proposal from its proxy materials if it would "violate any state, federal, or foreign law to which it is subject.").

shareholder activism and firm value during 2001–2013.[72] She finds that "[o]wnership by public pension funds engaged in social-issue shareholder-proposal activism is negatively related to firm value" and that "[t]here is no significant relationship between public pension fund ownership and firm value for funds engaging in shareholder-proposal activism focused on corporate governance rules."[73] If the proposals were intended to enhance shareholder wealth, one would expect some positive relationship.

Tara Bhandari, Peter Iliev, and Jonathan Kalodimos found that firms targeted by the comptroller for the submission of proxy access proposals "did not exhibit statistically significant stock market underperformance relative to the control group."[74] If those proposals were related to enhancing shareholder wealth, one would expect the proposals to target underperforming companies, but they did not.

In another recent paper, John Matsusaka, Oguzhan Ozbas, and Irene Yi examined the use of shareholder proposals by labor union–related entities.[75] Essentially, they found that shareholder proposals were being used as "bargaining chips in contract negotiations."[76] According to the authors, "[u]nion proposal activity increases by one-quarter in years where the union is negotiating a new contract with the company, and by two-thirds when the negotiation is contentious as evidenced by a work stoppage."[77] The authors concluded that "[t]he evidence suggests that sometimes having more rights can be costly for shareholders."[78]

In sum, those studies suggest that enhanced shareholder power through the ability to initiate shareholder proposals, which now includes proxy access, can lead to a reduction in shareholder value.

---

72 TRACIE WOIDTKE, MANHATTAN INSTITUTE, PUBLIC PENSION FUND ACTIVISM AND FIRM VALUE: AN EMPIRICAL ANALYSIS 20 (2015).

73 *Id.* at 3.

74 Bhandari et al., *supra* note 14, at 15. *See also* Jonathan B. Cohn et al., *On Enhancing Shareholder Control: A (Dodd-) Frank Assessment of Proxy Access*, 71 J. FIN. 1623, 1626 (2016) (suggesting that unions and public pension funds participating in proxy access could lead to decreases in shareholder value).

75 John G. Matsusaka et al., *Opportunistic Proposals by Union Shareholders* (Marshall School of Business, University of Southern California, Research Paper No. CLASS15-25, 2016) http://ssrn.com/abstract=2666064. *See also* Ashwini K. Agrawal, *Corporate Governance Objectives of Labor Union Shareholders: Evidence from Proxy Voting*, 25 REV. FIN. STUD. 187, 189 (2012) (finding that "[w]hen a firm's unionized employees are no longer represented by the AFL-CIO, the AFL-CIO's pension funds become significantly less opposed to the firm's directors in subsequent board elections").

76 Matsusaka et al., *supra* note 75, at 14.

77 *Id.* at 26-27.

78 *Id.* at abstract.

F. *Empirical Evidence*

So far, the study by Bhandari, Iliev, and Kalodimos, an SEC staff working paper, provides the only empirical study on the value of proxy access proposals.[79] The researchers focused their study on the day the New York City comptroller unexpectedly announced the proxy access initiative to the public, November 6, 2014.[80] Bhandari and her colleagues found that the announcement led to a positive, statistically significant 0.53% abnormal return for the seventy firms that they used in their sample.[81] They also found that the firms targeted by the comptroller's initiative did not correlate with those that were perceived to have benefited the most from universal proxy access at the time the SEC stayed its proxy access rules.[82] As already discussed, such a finding is one indication that the comptroller was not using proxy access in the most value-enhancing manner.[83] Supporting this conclusion is the finding that the comptroller's choice of target firms was "not significantly associated with poor recent stock performance of the firm or the growth opportunities of the firm . . . ."[84]

Even though the study suggests otherwise,[85] the sample clearly suffers from selection bias, resulting in a lack of randomness, and is therefore not representative of the universe of public companies. Of the seventy-five companies targeted by the comptroller, thirty-three were targeted because they were in industries directly related to climate change, twenty-four were targeted for a lack of board diversity, and twenty-five were cited for having received "significant opposition to their 2014 advisory vote on executive compensation."[86] As a result, twenty of the seventy-five firms targeted as part of the BAP initiative were from the petroleum and natural gas industry, nine were from the utilities industry, and six more were from the coal industry.[87] This sample is significantly overweighted with firms that are either producing or consuming huge quantities of carbon-based fuels, and is therefore not representative of the current universe of US public companies. Moreover, the sample size is small, so it is not possible to support the overall result with cross-sectional analysis. Selection bias and small sample

---

79 Bhandari et al., *supra* note 14.

80 Five firms were removed from the sample because they had made earnings announcements on that day. *Id.* at 18.

81 *Id.* at 4.

82 *Id.* at 24.

83 *Id.*

84 *Id.* at 28.

85 *Id.* at 17.

86 Press Release from Scott M. Stringer, New York City Comptroller, NYC Pension Funds Launch National Campaign to Give Shareowners a True Voice in How Corporate Boards Are Elected (Nov. 6, 2014), http://comptroller.nyc.gov/newsroom/comptroller-stringer-nyc-pension-funds-launch-national-campaign-to-give-shareowners-a-true-voice-in-how-corporate-boards-are-elected/.

87 Bhandari et al., *supra* note 14, at 43.

size mean that the results cannot be generalized to the market as a whole or be used to draw strong conclusions.[88] In sum, the study suffers from a lack of external validity. That is, the results of the study may provide information about how proxy access may affect the firms in the sample, but they cannot be generalized to the thousands of other firms that make up the universe of public companies.

G. *Summary*

Mandated shareholder-initiated proxy access strays from the principle that it is not the role of the federal securities laws to determine how authority is to be distributed in a public company. That is the role of private ordering as sanctioned by state corporate law. In 1997, the SEC understood that principle when it discussed shareholder proposals in the context of the "ordinary business" exclusion:

> The shareholder proposal process affects the internal governance of corporations, and it is state law—not federal securities law—which is primarily concerned with corporate governance matters. In its current form, rule 14a-8 in fact defers to state law on the central question of whether a proposal is a proper matter for shareholder action. The "ordinary business" exclusion is based in part on state corporate law establishing spheres of authority for the board of directors on one hand, and the company's shareholders on the other.[89]

The discussion in this section offers several lessons. First, private ordering under corporate law is a structured approach that purposefully selects the board to take the lead in determining the optimal corporate governance arrangements.

Second, superior decision-making efficiency is the rationale that explains why the outcome of the bargaining process always allows DGCL Section 141(a) to be incorporated without substantive modification in a public company and by extension allows the board to control the private ordering of governance arrangements, including proxy access.

Third, shareholders may propose and approve binding bylaws. However, those bylaws, as well as bylaws in general, will not survive a legal challenge if they interfere with the board's substantive decision-making authority under DGCL Section 141(a).

---

88 ASWATH DAMODARAN, INVESTMENT VALUATION: TOOLS AND TECHNIQUES FOR DETERMINING THE VALUE OF ANY ASSET 121 (2012) ("Since there are thousands of stocks that could be considered part of this universe, researchers often choose to use a smaller universe. When this choice is random, this does limited damage to the results of the study. If the choice is biased, it can provide results which are not true in the larger universe.").

89 Amendments to Rules on Shareholder Proposals, 34 Fed. Reg. 39093 (proposed Sept. 18, 1997) (to be codified at 17 C.F.R. pt. 240).

Fourth, shareholder proposals on proxy access in and of themselves request that the board be relieved of some of its authority to participate in the director nomination process without a charter amendment to authorize such a change as required under DGCL Section 141(a) and *CA, Inc*.

Fifth, DGCL Section 112 is an outlier under Delaware corporate law and was passed in response to the prospect of SEC-mandated proxy access and not to overrule the judicial approach taken *CA, Inc.*

Sixth, proxy access may allow certain shareholders, such as public pension funds and labor union–related entities, to act opportunistically if they deem it to be their advantage without the burden of fiduciary duties to the company or other shareholders. Moreover, once a binding bylaw is in place, the direct nomination of directors by shareholders may not follow a value-maximizing path. For example, what would stop the comptroller from continuing to target firms for the direct nomination of directors using the same criteria that it used in its proxy access initiative? Therefore, the negative wealth effects from inefficiency in the targeting of firms would arguably only expand as potentially opportunistic shareholders used the direct nomination of director candidates as a negotiating tool to accomplish their non-wealth-maximizing goals. Those implications suggest that fiduciary duties may be required of shareholders who nominate directors through proxy access.[90]

The lessons learned in this section also lead to a policy recommendation that the SEC reinstate the ability of boards to exclude shareholder proposals on director elections under Rule 14a-8(i)(8), thereby allowing boards to once again have sole authority to determine whether the proxy access process should be initiated.

III. UNIVERSAL PROXY ACCESS

Despite the SEC's public recognition of the value provided by a board nominating committee,[91] the SEC has made several attempts over the years to change that paradigm and institute universal proxy access. The SEC's first attempt occurred in 1942.[92] However, the rule proposal was quickly withdrawn after strong opposition from corporate management.[93] In 1977

---

[90] Roberta Karmel, former SEC commissioner, was the first to suggest the idea of fiduciary duties for those shareholders who take advantage of proxy access. Roberta S. Karmel, *Should a Duty to the Corporation Be Imposed on Institutional Shareholders?*, 60 BUS. LAW. 1, 20-21 (2004). *See also* Iman Anabtawi & Lynn Stout, *Fiduciary Duties for Activist Shareholders*, 60 STAN. L. REV. 1255 (2008).

[91] A SEC task force organized in 1977 reported this to Congress. *See* Jill E. Fisch, *The Destructive Ambiguity of Federal Proxy Access*, 61 EMORY L.J. 435, 441 (2012) ("[A]s the SEC task force reported to the Senate, due to the emergence of nominating committees, a shareholder nomination rule was unnecessary.").

[92] *See* Karmel, *supra* note 90, at 440.

[93] *Id.* at 441.

and again in 1992, the SEC seriously studied proxy access but did not pursue its implementation.[94] In 2003, the SEC once again proposed universal proxy access but with a twist.[95] Universal proxy access would not be required unless a specific triggering event occurred at the company.[96] For example, if one or more directors had received a 35% withhold vote, then proxy access would be mandatory.[97] That rule was also withdrawn because of opposition from corporate management and the desire of commentators to wait until the impact of the recently passed Sarbanes-Oxley Act of 2002 had become known.[98] Most recently, on June 10, 2009, the SEC introduced another universal proxy access proposal.[99] However, it was not until after the enactment of the Dodd-Frank Act—which contained a provision, Section 971, providing the SEC with express authority to implement universal proxy access[100]—did the SEC go ahead and adopt Rule 14a-11,[101] a rule that was to become effective on November 15, 2010. In response, the Business Roundtable and the US Chamber of Commerce filed a lawsuit in the D.C. Circuit Court of Appeals seeking that the court vacate the rules.[102] The three-judge panel of the D.C. Circuit Court of Appeals

---

[94] *Id.* at 441-42.

[95] Security Holder Director Nominations, 68 Fed. Reg. 60784–85 (proposed Oct. 23, 2003) (to be codified at 17 C.F.R. pt. 240, 249, 274).

[96] Fisch, *supra* note 91, at 442.

[97] *Id.*

[98] *Id.* at 444.

[99] Facilitating Shareholder Director Nominations, 33 Fed. Reg. 9046, 34 Fed. Reg. 60089 (proposed June 10, 2009) (to be codified at 17 C.F.R. pt. 200, 232, 240, 249, 274).

[100] Section 971 of the Dodd-Frank Act provides the following:

> (a) PROXY ACCESS.—Section 14(a) of the Securities Exchange Act of 1934 (15 U.S.C. 78n(a)) is amended—(1) by inserting "(1)"after "(a)"; and (2) by adding at the end the following:
>
> "(2) The rules and regulations prescribed by the Commission under paragraph (1) may include—"(A) a requirement that a solicitation of proxy, consent, or authorization by (or on behalf of) an issuer include a nominee submitted by a shareholder to serve on the board of directors of the issuer; and "(B) a requirement that an issuer follow a certain procedure in relation to a solicitation described in subparagraph (A)."
>
> (b) REGULATIONS.—The Commission may issue rules permitting the use by a shareholder of proxy solicitation materials supplied by an issuer of securities for the purpose of nominating individuals to membership on the board of directors of the issuer, under such terms and conditions as the Commission determines *are in the interests of shareholders and for the protection of investors.*

Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, §971, 124 Stat. 1376, 1915 (2010) (emphasis added). Providing such statutory authority to the SEC in the Dodd-Frank Act was necessary to erase any doubts that the SEC had the authority to promulgate proxy access rules that arose in an over 20-year-old decision involving the Business Roundtable. *See* Fisch, *supra* note 91, at 438 (citing Bus. Roundtable v. SEC, 905 F.2d 406 (D.C. Cir. 1990)) (arguing that the SEC does not have authority to interfere with the substantive features of company voting rights as established under state corporate law).

[101] 17 C.F.R. § 240.14a-11.

[102] Brief of Petitioner at 1-2, Bus. Roundtable v. SEC, 647 F.3d 1144 (D.C. Cir. 2011) (No. 10-1305), 2010 WL 3770710.

unanimously decided to vacate Rule 14a-11 after determining that the SEC had violated the Administrative Procedure Act (APA) by promulgating the rule in violation of the act's "arbitrary and capricious" standard of review.[103] That violation was the result of the SEC's failure "adequately to consider the rule's effect upon efficiency, competition, and capital formation" as required by statute.[104]

The discussion in the preceding section of this study suggests that allowing the nomination of directors to be controlled by the board and its nominating committee has significant value for shareholders. If so, why has the SEC revisited universal proxy access over the years?

Whatever the reason, one could still make the case that universal proxy access is efficient if the empirical evidence demonstrated its efficiency. After all, the benefits of proxy access may exceed the value of board-initiated private ordering if the absence of proxy access has led to significant "managerial rent extraction,"[105] or what is more commonly referred to as *agency costs*, created by the separation of share ownership from the management of a public company.[106] According to Macey, "Ultimately, the best way of evaluating the relative desirability of an enabling regime of corporate law, as opposed to a mandatory regime, is by examining the relevant empirical evidence."[107] This empirical approach is totally consistent with what is required of the SEC under its statutory obligations as discussed in *Business Roundtable*.[108] If the SEC decides to make another attempt at universal proxy access, then it will need to do a comprehensive review of the empirical work available to show that it has met "its statutory responsibility to determine the likely economic consequences of" a new rule "and to connect those consequences to efficiency, competition, and capital formation."[109]

---

[103] 647 F.3d at 1156.

[104] *Id.* at 1146 (citing § 3(f) of the Exchange Act and § 2(c) of the Investment Company Act of 1940). According to the Court,

> Here the Commission inconsistently and opportunistically framed the costs and benefits of the rule; failed adequately to quantify the certain costs or to explain why those costs could not be quantified; neglected to support its predictive judgments; contradicted itself; and failed to respond to substantial problems raised by commenters.

Id. at 1148–49.

[105] Larcker et al., *supra* note 40, at 431.

[106] *See* ADOLPH A. BERLE & GARDINER C. MEANS, THE MODERN CORPORATION AND PRIVATE PROPERTY (1932); Jensen & Meckling, *supra* note 34.

[107] Jonathan R. Macey, *Corporate Law and Corporate Governance: A Contractual Perspective*, 18 J. CORP. L. 185, 207 (1993).

[108] 647 F.3d at 1148.

[109] *Id.*

A. *Enhancing Shareholder Wealth as the Objective of Universal Proxy Access*

Before reviewing the existing empirical evidence, it is critical to determine the objective of universal proxy access. After all, a review of empirical studies would be worthless without an understanding of what a researcher would be looking for in those studies. Fortunately, Section 971(b) of the Dodd-Frank Act takes a shareholder-centric approach to any new SEC rule on universal proxy access. That is, any such rule must be "in the interests of shareholders and for the protection of investors."[110] Somewhat surprisingly, the meaning of that statutory requirement was not discussed in the original SEC release finalizing Rule 14a-11,[111] or in the *Business Roundtable* decision that vacated the rule.[112] So, how will the SEC interpret the statutory language if it tries to implement another universal proxy access rule? For starters, it is not enough to say that "proxy access is a fundamental right of long-term shareowners."[113] Although "that sentiment may have a vaguely constitutional ring to it,"[114] it offers the SEC no tangible guidance in terms of what is meant by "in the interests of shareholders and for the protection of investors."[115] Moreover, long-term shareholders have never had a fundamental right to proxy access under either corporate or federal securities law.[116] The only requirement in terms of director nominations is that at least some shareholders must have the power to nominate directors at the annual meeting itself.[117] In addition, it is not adequate to say that "[o]ne of the key tenets of the Federal proxy rules on which the Commission has consistently focused is whether the proxy process functions, as nearly as possible, as a replacement for an actual in-person meeting of shareholders."[118] This statement may be true, but it is not relevant to interpreting the new language found in the Dodd-Frank Act.

However, there is no reason for the SEC not to interpret the language of the Dodd-Frank Act, as it pertains to universal proxy access, to mean enhancing shareholder wealth. That approach is consistent with a number of established ways of looking at corporate governance. For example,

---

[110] Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 971, 124 Stat. 1376, 1915, (2010).

[111] *See generally* Facilitating Shareholder Director Nominations, 75 Fed. Reg. 56,668 (Sept. 16, 2010) (to be codified at 17 C.F.R. pt. 200, 232, 240 and 249).

[112] 647 F.3d at 1151-52.

[113] COUNCIL OF INSTITUTIONAL INVESTORS, *supra* note 18, at 1.

[114] BERNARD. S. SHARFMAN, R STREET INSTITUTE, R STREET SHORTS NO. 12, PUBLIC-PENSION FUNDS PLAY WITH NEWEST TOY IN CORPORATE GOVERNANCE 2 (2015) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2643622.

[115] Dodd-Frank Act § 971(b).

[116] *Id.*

[117] Lawrence A. Hamermesh, *Director Nominations*, 39 DEL. J. CORP. L. 117, 150-51 (2014).

[118] Facilitating Shareholder Director Nominations, 75 Fed. Reg. 56,670.

under a nexus of contracts understanding of the firm, shareholders are the sole claimants on the residual cash flows generated by the firm, because other parties transacting with the corporation can adequately protect themselves by contract. That is, shareholders have the greatest risk of ending up with nothing as a result of their dealings with the corporation. The board may have ultimate authority to act and make decisions under the default rules of corporate law, *but* that authority is given by shareholders only if the board acts to enhance shareholder value. Moreover, having a board and executive management target the enhancement of shareholder wealth means that all other parties that have contracted with the corporation must be paid off before the shareholders receive a residual, if any.[119] Therefore, these other contracting parties should be supportive of enhancing shareholder wealth as the objective of corporate authority. As stated by Henry Manne, the result is an example of "pure positive economics" and should be accepted as such.[120] In sum, that objective is what all parties to the corporate contract agree to and what the courts should be expected to enforce.

From a sustainability perspective, that approach is consistent with how the management of public companies should go about making decisions to create value.[121] Value is created when a firm generates enough cash inflows to cover its cash outflows.[122] "The timing of the inflows and outflows must then be discounted by the proper interest rate to determine if they have a positive net present value."[123] If the net present value is positive, "then the firm has value."[124] Moreover, continuously making investments with present values expected to be positive should lead to sustainable value creation.[125]

However, if management wants to make sure that sustainable value creation has the best chance of occurring, then it should also have the responsibility of trying to maximize the net present value as part of its decision-making calculus at any point in time.[126] The process of maximization can be referred to as *long-term value creation*, a process that management should be striving to implement.[127] As a result, long-term

---

119 FRANK H. EASTERBROOK & DANIEL R. FISCHEL, THE ECONOMIC STRUCTURE OF CORPORATE LAW 38 (1991) (noting that "maximizing profits for equity investors assists the other 'constituencies' automatically").

120 Email from Henry G. Manne, professor emeritus of law, Geo. Mason Univ., to Bernard S. Sharfman (Dec. 29, 2012) (on file with author).

121 Bernard S. Sharfman, *Activist Hedge Funds in a World of Board Independence: Creators or Destroyers of Long-Term Value?*, 2015 COLUM. BUS. L. REV. 813, 832 (2015).

122 *Id.* at 831.

123 *Id.* at 832.

124 *Id.*

125 *Id.*

126 *Id.*

127 *Id.*

value creation equals maximizing a firm's net present value and thus also equates to shareholder wealth maximization.[128]

The benefits of that equivalency have been described in terms of sustainability by Eugene Fama and Michael Jensen:

> Contracts [entered into by organizations such as public companies] that direct decisions toward the interests of residual claimants . . . add to the survival value of organizations. Producing outputs at lower cost is in the interests of residual claimants because it increases net cash flows, but lower costs also contribute to survival by allowing products to be delivered at lower prices.[129]

B. *The Empirical Evidence*

Critical to deciding whether the SEC should use its authority to implement universal proxy access is a proper evaluation of the empirical evidence that is available. For that purpose, this section takes the perspective of an impartial SEC commissioner who is trying to decide whether to implement universal proxy access. This decision is difficult for the commissioner because the issue of universal proxy access resides in the world of corporate governance, not securities regulation, and its focus on disclosure. Moreover, the world of corporate governance arrangements rests primarily in the hands of state corporate law—again not an area of expertise for the commissioner—and its private ordering approach to such arrangements. If such an approach can be understood to be generally wealth enhancing for shareholders, then the commissioner must tread very carefully when considering the implementation of universal proxy access. Such careful consideration requires that a vote for universal proxy access be supported by empirical evidence that consistently shows (from study to study) its value over time and that does not simply support proxy access at any one point in time. This section argues that such empirical evidence does not currently exist.

So far, the empirical evidence on proxy access comes exclusively from event studies. Event studies investigate the effect of new information—the event—on the expected stock returns of a targeted cross-section of firms.[130] The null hypothesis to be tested is whether the mean abnormal return at the time of the event is equal to zero.[131] That is, event studies are used to determine "whether there is an abnormal stock price effect associated with

---

128 *Id.*

129 Eugene F. Fama & Michael C. Jensen, *Separation of Ownership and Control*, 26 J. L. & ECON. 301, 303 (1983).

130 Roberta Romano, *Less is More: Making Institutional Investor Activism a Valuable Mechanism of Corporate Governance*, 18 YALE J. ON REG. 174, 187, n. 37 (2001).

131 S.P. Kothari & Jerold B. Warner, *Econometrics of Event Studies*, *in* HANDBOOK OF CORPORATE FINANCE: EMPIRICAL CORPORATE FINANCE 9 (B. Espen Eckbo ed., vol. 1, 2006).

an unanticipated event"[132] on a targeted sample of firms that may have been uniquely affected by the event.

In the context of using event studies to evaluate proxy access, one can think of the standard null hypothesis, as previously described, as corresponding to the following statement: The "preservation of managerial discretion" in the nomination of directors is wealth enhancing for shareholders.[133] Therefore, if the SEC feels compelled to interject universal proxy access into the governance of public companies, then the SEC has the burden to demonstrate that the available empirical studies provide sufficient evidence to show that proxy access consistently generates abnormal returns to the extent that the SEC has comfort that this null hypothesis has been rejected. Moreover, because event studies report what the market thinks only at a point in time, significant consistency between studies is needed to provide comfort that the burden has been met over a period of time (stationarity), not just at any one point in time. If that burden cannot be met, then the SEC once again risks being found by a court to have "relied upon insufficient empirical data" when the SEC concludes that its universal proxy access rule "will improve board performance and increase shareholder value by facilitating the election of dissident shareholder nominees."[134] Finally, although not a statutory requirement, the threshold of evidence must be high because of the political reality that if the SEC implements a mandatory rule, that rule will be very difficult to unwind, even if the evidence becomes overwhelming that the rule is generally harmful to shareholders.

So far, all the event studies on proxy access can be characterized as being natural experiments.[135] That is, those event studies provide statistical analysis on proxy access–related events that are understood to be exogenous shocks to the stock market, such as when the court in *Business Roundtable* vacated the SEC's universal proxy access rule on July 22, 2011. A summary of those studies follows.

---

132 S.V.D. Nageswara Rao & Sreejith. U, *Event Study Methodology: A Critical Review*, 3 MACROTHEME REV. 40, 40 (2014).

133 This null hypothesis is derived from Stephen Bainbridge's argument that the "[p]reservation of managerial discretion should always be the null hypothesis." Stephen M. Bainbridge, *The Business Judgment Rule as Abstention Doctrine*, 57 VAND. L. REV. 83, 109 (2004).

134 Bus. Roundtable v. SEC, 647 F.3d 1144, 1150 (D.C. Cir. 2011).

135 The term *natural experiments* may be defined as "a naturally occurring state (event) resulting from a social or political situation and thus not intentionally set up by the researcher." Jennifer Gippel et al., *Endogeneity in Accounting and Finance Research: Natural Experiments as a State-of-the-Art Solution*, 51 ABACUS 143, 156 (2015). The "naturally occurring state, often comes about from a social or political situation such as a government policy change." *Id.* at 158. Moreover, "[n]atural experiments are not 'true' experiments . . ." *Id.* "This is because the so called *naturally occurring state* is not intentionally set up by the researcher and so the treatment group is not randomly assigned. Such experiments are more like observational studies where the researcher cannot manipulate the environment, although the researcher must choose the comparison or control group." *Id.*

1. Larcker, Ormazabal, and Taylor

David F. Larcker, Gaizka Ormazabal, and Daniel J. Taylor evaluated thirteen proxy access–related events that occurred between April 2007 and June 2009.[136] The authors characterized eight of those events as increasing the likelihood of proxy access regulation, and they identified the other five as decreasing the likelihood.[137] Overall, they found "a weak negative reaction to proxy access regulation,"[138] an average abnormal event day return of −0.32%.[139] Also, on a cross-sectional basis they found "strong evidence that abnormal returns are increasingly negative for firms with a greater number of large institutional blockholders."[140] Financial firms were excluded from their sample.[141] Abnormal returns were computed on the day of the event relative to the Center for Research in Security Prices value-weighted market index that excluded dividends and distributions.[142]

As the first empirical study on proxy access, of course, Larcker and his colleagues enabled subsequent studies the opportunity to provide their share of criticism. Ali Akyol, Wei Fen Lim, and Patrick Verwijmeren found the approach in Larcker, Ormazabal, and Taylor problematic because the average abnormal return of a large portfolio of US firms compared with the US market index should be close to zero.[143]

Bo Becker, Daniel Bergstresser, and Guhan Subramanian criticized Larcker and his colleagues' report for using events that are of "questionable importance."[144] For example, they questioned whether April 24, 2007, the date the SEC announced that it was scheduling a series of discussions on proxy access, was an event that increased the likelihood of proxy access or was even directionally clear because "the *AFSCME* decision permitted proxy access on a company-by-company basis."[145] Moreover, the authors suggest that many of the events used by Larcker, Ormazabal, and Taylor were predicted in advance.[146] For example, it is commonly known that the Delaware legislature gives great deference to the recommendations of the Corporate Law Section of the Delaware Bar Association when amending the DGCL. Therefore, when the Corporate Law Section voted in favor of a

---

136 Larcker et al., *supra* note 40, at 442–43.

137 *Id.* at 432.

138 *Id.*

139 *Id.* at 444.

140 *Id.* at 432-33.

141 *Id.* at 438.

142 *Id.* at 439, n. 39.

143 Ali C. Akyol et al., *Shareholders in the Boardroom: Wealth Effects of the SEC's Proposal to Facilitate Director Nominations*, 47 J. FIN. & QUANTITATIVE ANALYSIS 1029, 1031 (2012).

144 Bo Becker et al., *Does Shareholder Proxy Access Improve Firm Value? Evidence from the Business Roundtable's Challenge*, 56 J. L. & ECON. 127, 137 (2013).

145 *Id.*

146 *Id.*

shareholder access amendment on February 26, 2009, its implementation in Delaware became very likely.[147] Nevertheless, as Becker and his colleagues pointed out, Larcker, Ormazabal, and Taylor do not include the recommendation from the Corporate Law Council on February 26, 2009 but do include as events that decreased the possibility of proxy access three subsequent and anticipated steps in the Delaware statutory process to make the shareholder access amendment law.[148] According to Becker and his colleagues, because the marketplace may have fully anticipated those subsequent events, the events studied may have had no value.[149]

2. Akyol, Lim, and Verwijmeren

Ali C. Akyol, Wei Fen Lim, and Patrick Verwijmeren evaluated seventeen proxy access–related events between September 2006 and September 2010.[150] Nine of the seventeen events can also be found in Larcker and his colleagues' study.[151] In contrast to Larcker, Ormazabal, and Taylor, Akyol and his colleagues used a global market index (excluding U.S. firms) and a Canadian market index as benchmarks.[152] Because the SEC's rule affects only U.S. companies, an abnormal return for the U.S. portfolio compared with either the world index or the Canadian index on event days should demonstrate the value relevance of the proxy access–related events.[153] The authors used a sample of firms totaling 4,719.[154] They observed a statistically significant daily abnormal return of −0.70% for ten events that increased the probability of proxy access and a significantly positive return of 0.80% for the seven events in which the probability of proxy access declined.[155] They also observed that the overall wealth effect for proxy access was −0.70% relative to the global index and

---

147 *Id.*

148 *Id.* at 137-38.

149 *Id.* at 138.

150 Akyol et al., *supra* note 143, at 1036.

151 *Id.* at 1035. Akyol and his colleagues also did a follow-up study that focused on how a firm's governance characteristics affected the returns from proxy access–related events. *See* Ali C. Akyol et al., *Governance Characteristics and the Market Reaction to the SEC's Proxy Access Rule*, 12 INT'L REV. FIN. 175 (2012). In that study, they find that announcement effects are positively related to the fraction of independent directors and the ratio of non-cash-based compensation but negatively related to board size. According to the authors, "We find that the market reaction to the proxy access rule is more negative for firms that have a larger board, a lower fraction of independent directors, a lower non-cash-based compensation to total compensation ratio, a CEO who is also a chairman, and a higher number of eligible institutional investors. Overall, our results suggest that the one-size-fits-all proxy access rule is not perceived as value increasing by the marginal shareholder." *Id.* at 194.

152 Akyol et al., *supra* note 143, at 1030.

153 *Id.*

154 *Id.* at 1038.

155 *Id.* at 1030.

−0.60% relative to the Canadian market index.[156] They found the overall wealth effect to be statistically significant at the 5% level under both the global market index and the Canadian market index.[157] Their findings are also consistent with Larcker and his colleagues in showing that the firms with more eligible institutional investors had the most negative wealth effects.[158] As with Larcker, Ormazabal, and Taylor, Becker and his colleagues criticized Akyol, Lim, and Verwijmeren for using some event dates that may be of "questionable importance."[159]

3. Cohn, Gillan, and Hartzell

Jonathan Cohn, Stuart Gillan, and Jay Hartzell focused on three proxy access–related events before Rule 14a-11 was vacated.[160] The first event was the June 16, 2010 announcement by Senator Christopher Dodd that he was submitting a proposal amending the Dodd-Frank Act to direct the SEC to require that an investor or group of investors own at least 5% of a firm's shares for two years before gaining access to a firm's proxy for purposes of nominating directors.[161] At the time, the SEC had proposed a much less stringent tiered system, with minimum holdings of 1%, 3%, and 5%, respectively, for firms with market capitalizations greater than $700 million (large accelerated filers), between $75 million and $700 million (accelerated filers), and less than $75 million (nonaccelerated filers).[162] The second event was the dropping of Dodd's proposal on June 24, 2010, which led to a restoration of the SEC's proposed thresholds as the likely outcome of proxy access.[163] The third event occurred on October 4, 2010, when the SEC voluntarily stayed implementation of the universal proxy access rule in response to the *Business Roundtable* litigation.[164]

Cohn and his colleagues compared the abnormal returns of accelerated filers and large accelerated filers who were expected to be affected by these events with those of nonaccelerated filers who were not expected to be affected by the first two events and were expected to be affected, to a lesser extent, by the third event because they had already been given a three-year exemption from proxy access. Financial firms were excluded from the study's database, which included 3,102 firms.[165]

---

156 *Id.* at 1044.
157 *Id.* at 1043.
158 *Id.* at 1031, 1055.
159 Becker et al., *supra* note 144, at 137.
160 Cohn et al., *supra* note 74.
161 *Id.* at 2, 15.
162 *Id.* at 2.
163 *Id.*
164 *Id.* at 3.
165 *Id.* at 12.

As expected, Cohn, Gillan, and Hartzell found about zero abnormal returns for nonaccelerated filers for all three events.[166] However, for accelerated and large accelerated filers the results were mixed, even though directionally they show that the market favored proxy access.[167] For the first event, Cohn and his colleagues found a statistically insignificant negative response for both accelerated and large accelerated filers.[168] For the second event, the researchers found a statistically significant (10% level) positive large response for accelerated filers (2.01%) but a statistically insignificant positive response for large accelerated filers.[169] For the third event, they found a statistically significant (10% level) negative response for accelerated filers (1.68%) and a statistically insignificant negative response for large accelerated filers.[170] However, when Cohn and his colleagues combined the returns found in all three events, the statistical significance increased dramatically to the 1% level for both accelerated and large accelerated filers.[171] That combination enabled Cohn, Gillan, and Hartzell to conclude that universal proxy access is value-enhancing for shareholders.[172]

4. Nonstationarity

Cohn and his colleagues raise the issue of nonstationarity. *Nonstationarity* refers to how the stock market may react differently to the same events at different points in time.[173] That is, the stock market may provide "one result for a period and a diverse outcome for another period."[174] Nonstationarity "is due to the change in perception of investors over a period of time."[175] That change is consistent with an efficient market in which the market price is an unbiased estimate of the true value of the investment but is not necessarily a correct one at any point in time.[176] As a result, it is possible that as the market becomes more informed about the real impact of an event, the market may change its opinion on how the event affects shareholder value.[177]

---

166 *Id.* at 44.
167 *Id.*
168 *Id.*
169 *Id.*
170 *Id.*
171 *Id.*
172 *Id.* at 31.
173 Rao & Sreejith, *supra* note 132, at 44.
174 *Id.*
175 *Id.*
176 DAMODARAN, *supra* note 88, at 153.
177 Yaniv Konchitchki & Daniel E. O'Leary, *Event Study Methodologies in Information Systems Research*, 12 INT'L J. ACCT. INFO. SYS. 99, 108 (2011).

Cohn, Gillan, and Hartzell found it disconcerting that their study's results differ significantly from the results found by both Larcker and his colleages and Akyol and his colleagues.[178] Cohn and his colleagues suggest that the differences could have been the result of nonstationarity:

> We cannot reject the possibility that the differences between our results and those of Larcker, Ormazabal, and Taylor (2010) [and those of Akyol and his colleauges as well[179]] are driven by changes over time in the market's beliefs about the value of shareholder control. One possible trigger of such a change is the financial crisis. However, we exclude from our analysis financial firms, where concerns about mismanagement were likely to have been the strongest after the crisis. Moreover, our events take place in 2010, after the worst part of the crisis period, and complementary results regarding subsequent events in contemporaneous working papers (Becker, Bergstresser, and Subramanian (2013) and Jochem (2012)) suggest that our characterization of the perceived value of proxy access persisted at least until mid-2011.[180]

Even if the market had a positive perspective on proxy access on the dates studied by Cohn, Gillan, and Hartzell, it is easy to see how the market may change its mind over time. The market's valuation of proxy access is based on very little data. Proxy access has yet to be used to nominate a director, let alone provide information traders with enough data to evaluate how board members who were nominated by shareholders have enhanced or subtracted from firm performance and market value. This lack of data makes the results of any empirical study at this time and for a number of years in the future susceptible to nonstationarity. Therefore, it may be a long time before the market gets a handle on the ultimate value, positive or negative, of proxy access.

5. The Role of Activist Hedge Funds

Most importantly, to understand why Cohn, Gillan, and Hartzell found universal proxy access to be wealth enhancing for shareholders on the event dates, one needs to understand the significant role that activist hedge funds may have played in the results. On all three event dates, Cohn and his colleagues found the market response to be stronger for firms that had a known activist hedge fund as a shareholder.[181] Moreover, chronologically, sitting between the second and third events was a fourth event that the researchers also analyzed, the SEC's passage of the final proxy access rule on August 25, 2010.[182] Even though passage was anticipated, uncertainty

---

178 Cohn et al., *supra* note 74, at nn.6-7.

179 *Id.* at n.6.

180 *Id.* at n.7 (citing Becker et al., *supra* note 144, and Jochem, *infra* note 212).

181 *Id.* at 1626.

182 *SEC Adopts New Measures to Facilitate Director Nominations by Shareholders,* SEC NEWS DIGEST, Aug. 25, 2010, https://www.sec.gov/news/digest/2010/dig082510.htm.

remained about whether the minimum holding period was going to be two or three years.[183] As it turned out, the final rule had a three-year holding period.[184] Cohn, Gillan, and Hartzell found that the three-year holding period had a negative effect for activist hedge funds relative to others.[185] The finding makes sense because for activist hedge funds to maximize returns to their investors, they cannot hold the stock of their targets for a long period of time.[186] That is, to maximize their profits, they must maximize the number of their interventions.[187] That statement is consistent with the argument that anything that negatively affects hedge fund activism should negatively affect the stock market because hedge fund activism is the only form of activism that has been shown to be significantly wealth-enhancing for shareholders over time.[188] In sum, Cohn, Gillan, and

---

[183] Cohn et al., *supra* note 74, at 1627.

[184] *Id.*

[185] *Id.*

[186] Holding periods of activist hedge funds are estimated to average around 20 months. Alon Brav et al., *Hedge Fund Activism, Corporate Governance, and Firm Performance*, 63 J. FIN. 1729, 1732 (2008).

[187] Sharfman, *supra* note 121, at 825.

[188] *See* Brav et al., *supra* note 186, at 1731. *See also* Nicole M. Boyson & Robert M. Mooradian, *Corporate Governance and Hedge Fund Activism*, 14 REV. DERIVATIVES RES. 169, 175-78, 201 (2011) (examining data from 1994–2005 and finding that hedge fund activism improved by short-term and long-term performance of companies); Christopher P. Clifford, *Value Creation or Destruction? Hedge Funds as Shareholder Activists*, 14 J. CORP. FIN. 323, 324 (2008) (finding that in a control group containing hedge funds that filed Schedule 13Gs, "firms targeted by hedge funds for active purposes earn larger, positive [returns] than firms targeted by hedge funds for passive purposes"); Robin Greenwood & Michael Schor, *Investor Activism and Takeovers*, 92 J. FIN. ECON. 362, 374 (2009) (finding that "activists are most successful at creating value when they are able to [force] a change in control"); Dionysia Katelouzou, *Myths and Realities of Hedge Fund Activism: Some Empirical Evidence*, 7 VA. L. & BUS. REV. 459, 479 (2013) (examining empirical results consistent with these studies but focusing on hedge fund activity outside the United States); April Klein & Emanuel Zur, *Entrepreneurial Shareholder Activism: Hedge Funds and Other Private Investors*, 64 J. FIN. 187, 213, 217–18 (2009) (focusing on activist campaigns by both hedge funds and other types of entrepreneurial activists and finding that both types of campaigns produced average abnormal returns for target shareholders); Alon Brav et al., *Shareholder Power and Corporate Innovation: Evidence from Hedge Fund Activism* (Ind. Univ., Kelley Sch. of Bus., Research Paper No. 2014-05, 2014) (finding a link between improvements in innovation efficiency and hedge fund activism at firms with a diverse set of patents as a result of the activism leading to a more targeted approach to innovation); C. N. V. Krishnan et al., *The Second Wave of Hedge Fund Activism: The Importance of Reputation, Clout, and Expertise* (Vand. L. Sch., Law & Economics Working Paper No. 15-9, 2015), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2589992 (discussing that hedge fund activism continues to generate positive, abnormal stock returns for an account period using a dataset collected from 2008 through mid-2014); Marco Becht et al., *The Returns to Hedge Fund Activism: An International Study* (European Corporate Governance Institute, Finance Working Paper No. 402/2014, 2014), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376271; Shane C. Goodwin, *Myopic Investor Myth Debunked: The Long-Term Efficacy of Shareholder Advocacy in the Boardroom* 10-13 (Harvard Bus. Sch., Working Paper 2014), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2450214 (reporting excess returns for activist hedge

Hartzell's findings that universal proxy access is wealth-enhancing on the event dates tested leads to the conclusion that those results may have been driven by the wealth-enhancing effects of hedge fund activism. Ironically, that value was driven out by the SEC's three-year holding period, a holding period that has become standardized in shareholder proposals on proxy access and in the binding proxy access bylaws implemented by boards.[189]

6. Campbell, Campbell, Sirmon, Bierman, and Tuggle

Joanna Tochman Campbell, T. Colin Campbell, David G. Sirmon, Leonard Bierman and Christopher S. Tuggle focused their study exclusively on one event date, August 25, 2010, the day the SEC approved its universal proxy access rule.[190] Their sample of firms was taken from the S&P 500.[191] After Campbell and her colleagues removed firms that had confounding events, such as dividends, earnings announcements, and missing data, the sample contained 392 firms.[192] Like Akyol and his colleauges, they used a market index of Canadian firms (but not a global index) to calculate expected returns.[193] They found a relatively large abnormal return of 0.83%.[194]

---

funds that gain board representation); Nicole M. Boyson et al., *Serial Activists* (Northeastern U. D'Amore-McKim Sch. Of Bus., Research Paper No. 2727371, 2016), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2727371&download=yes (using data from 2001–2013 to demonstrate that target performance is best when experienced hedge fund activists are involved). *But see* K. J. Martijn Cremers et al., *Hedge Fund Activism and Long-Term Firm Value*, (Working Paper, 2015), http://ssrn.com/abstract=2693231; Yvan Allaire & François Dauphin, *The Game of "Activist" Hedge Funds: Cui bono?*, INT'L J. DISCLOSURE & GOVERNANCE (2015), https://igopp.org/en/the-game-of-activist-hedge-funds-cui-bono/. For a sharp critique of the Cremers et al. study, *see* Lucian Bebchuk et al., *The Long-Term Effects of Hedge Fund Activism: A Reply to Cremers, Giambona, Sepe, and Wang*, HARVARD LAW SCHOOL FORUM ON CORPORATE LAW AND FINANCIAL REGULATION (Dec. 10, 2015), https://corpgov.law.harvard.edu/2015/12/10/the-long-term-effects-of-hedge-fund-activism-a-reply-to-cremers-giambona-sepe-and-wang/.

[189] It is important to note that the SEC's Rule 14a-11 has provided the parameters for current proxy access proposals. Such proposals typically allow shareholders who have owned 3% of the company's voting common stock for at least three years, on an individual but more commonly on an *aggregated* basis, access to a publicly traded company's proxy-solicitation materials for purposes of nominating their own candidates for board membership. That rule means that, at those companies for which a proxy access bylaw has been approved by shareholders or the board, certain shareholders, either individually or in a group, may now include their own slate of nominees (typically up to 20% or 25% of the board seats that are up for election) in the companies' proxy materials.

[190] Joanna Tochman Campbell et al., *Shareholder Influence over Director Nomination via Proxy Access: Implications for Agency Conflict and Stakeholder Value*, 33 STRATEGIC MGMT. J. 1431, 1432 (2012).

[191] *Id.* at 1440.

[192] *Id.*

[193] *Id.*

[194] *Id.* at 1444.

What is quite disconcerting about the study is the way the authors characterized the event as one of significant uncertainty.[195] They cited the 3–2 vote of the SEC and the absence of "widely held expectation of the outcome of the vote in advance" as attesting to this uncertainty.[196] However, this characterization of the event is incorrect. At the time, observers noted a lack of uncertainty about how the vote would go, long before the vote was ever taken. According to Becker, Bergstresser, and Subramanian, the majority view was that Section 971 of the Dodd-Frank Act made universal proxy access inevitable.[197] Moreover, according to a *Wall Street Journal* article that ran on Aug. 5, 2010, "people familiar with the matter" believed that the SEC would approve the proxy access rule,[198] making the event largely anticipated.[199] Moreover, as discussed by Cohn and his colleagues, although the decision was anticipated, the holding period was uncertain, leading one to expect a negative effect on shareholder wealth as a result of the three-year holding period.[200]

Consistent with the announcement being anticipated, Akyol and his colleagues found the event to have a small statistically insignificant negative effect of 30 basis points on shareholder wealth.[201] So why are Campbell and her colleagues' returns so large and positive (eighty-three basis points) even though the event may have been highly anticipated and expected to be negative? That conflict is yet to be resolved.

7. Becker, Bergstresser, and Subramanian

Bo Becker, Daniel Bergstresser, and Guhan Subramanian focused on October 4, 2010, the day when the SEC announced that it would stay implementation of its proxy access rules in response to the *Business Roundtable* litigation.[202] A news report indicated that the SEC stays are not a common occurrence, lending credence to choosing the stay as an

---

[195] *Id.* at 1442. What is also striking is the authors' mischaracterization of the overlap and differences between shareholder primacy, director primacy, team production, and stakeholder theory. *Id.* at 1443. However, a critique of this mischaracterization is beyond the scope of this study.

[196] *Id.* at 1442.

[197] Becker et al., *supra* note 144, at 132.

[198] Akyol et al., *supra* note 143, at 1034 (citing Kara Scannell, *SEC Set to Open Up Proxy Access*, WALL ST. J. (Aug. 5, 2010, 12:01 AM), https://www.wsj.com/articles/SB10001424052748704741904575409680246527908).

[199] Akyol et al., *supra* note 143, at 1053.

[200] Cohn et al., *supra* note 74, at 1626.

[201] Akyol et al., *supra* note 143, at 1044.

[202] Becker et al., *supra* note 144, at 139.

unanticipated event.[203] Becker and his colleagues used a sample taken from the S&P Composite 1500 consisting of 1,388 firms.[204] They also used measures of institutional ownership as a proxy for measuring a firm's vulnerability to proxy access.[205] In doing so, they divided their sample into deciles ranging from low to high levels of institutional ownership.[206] The bottom decile has institutional ownership of 25.3%, the top decile averages almost 70%.[207] Becker, Bergstresser, and Subramanian found that the average equal-weighted return was a negative 124 basis points on the event day, but firms in the top decile dropped forty-four basis points more than the firms in the bottom decile.[208] This finding implies that firms most susceptible to proxy access experienced significantly greater losses than firms that were most protected.[209] Therefore, the study indicates that universal proxy access is wealth-enhancing for shareholders.

It is interesting to note that Becker and his colleagues based their study of proxy access on the SEC's voluntary stay of Rule 14a-11. If the study is right, then one must ask why the stay surprised the market. One likely scenario is that the market must have believed that the request for a stay would fail at least one of the four factors the SEC used to evaluate a stay.[210] By granting the stay, the SEC must have determined that all four factors had been satisfied. The four factors are:

> 1) whether there is a strong likelihood that a party will succeed on the merits in a proceeding . . . (or, if the other factors strongly favor a stay, that there is a substantial case on the merits);
>
> 2) whether, without a stay, a party will suffer imminent, irreparable injury;
>
> 3) whether there will be substantial harm to any person if the stay were granted; and
>
> 4) whether the issuance of a stay would likely serve the public interest.[211]

---

[203] Bhandari et al., *supra* note 14, at 12, n.22. Note that several other recent motions to stay SEC rules, including rules related to mutual fund governance, conflict minerals, resource extraction, and securities issuance under Regulation A, were denied.

[204] Becker et al., *supra* note 144, at 143.

[205] *Id.* at 129.

[206] *Id.* at 143.

[207] *Id.*

[208] *Id.* at 143-44.

[209] *Id.* at 129.

[210] S.E.C, Order Denying Stay, Exchange Act Release, No. 68197, at 5 (2012), https://www.sec.gov/rules/other/2012/34-68197.pdf.

[211] *Id.* at 3.

The most notable factor is the first, the SEC's determination that there was a strong likelihood that the petitioners would succeed on the merits. If the market was truly surprised by the stay, then perhaps the surprise reflected a market that had underestimated the probability that the plaintiffs could win in court.

8. Jochem

Torsten Jochem focused exclusively on the event date of July 22, 2011, the day when the D.C. Circuit Court of Appeals vacated Rule 14a-11.[212] He estimated the shareholder wealth effects of greater proxy access by comparing the average daily abnormal returns of various portfolios of firms that were expected to have been affected by the event to the abnormal returns of the corresponding control portfolios that were not expected to be affected by the repeal.[213] Overall, he did "not find any significant shareholder wealth effect for the U.S. market as a whole."[214] However, whenever he does find statistically significant results, they are consistent with proxy access being wealth-enhancing.[215] Moreover, he found a "monotonically increasing valuation decline the smaller a firm" became.[216] However, Jochem did not find any significant effects on large capitalized firms or on firms with capitalization below $75 million.[217]

The greatest effect was on firms with capitalization between $75 million and $184 million.[218] He found a difference of twenty-three to thirty-eight basis points between the smallest quintile of affected firms and the largest quintile of firms.[219] He also found weak negative effects at firms where investors were eligible to immediately take advantage of proxy access.[220] Additionally he found a negative effect where a company has implemented a large number of antitakeover provisions such as staggered boards.[221] Finally, he did not find any effect on firms caused by the presence of public pension funds and labor union–related funds as shareholders.[222]

212 Torsten Jochem, *Does Proxy Access Increase Shareholder Wealth? Evidence from a Natural Experiment* 1 (U. of Pittsburgh, Working Paper, Apr. 2012).

213 *Id.* at 12.

214 *Id.* at 1.

215 *Id.* at 22.

216 *Id.*

217 *Id.* at 23.

218 *Id.*

219 *Id.*

220 *Id.*

221 *Id.* at 22.

222 *Id.* at 23.

9. Stratmann and Verret

Thomas Stratmann and J. W. Verret tested whether small firms with a market capitalization of between $25 million and $75 million performed differently on August 25, 2010, the date the SEC approved universal proxy access, compared with a control group of firms with a market capitalization of between $75 million and $125 million.[223] Stratmann and Verret's sample consisted of small firms because there were several surprises involving firms with capitalization of less than $75 million (nonaccelerated filers).[224] First, small firms obtained only a temporary three-year exemption from Rule 14a-11 instead of an expected total exemption.[225] Second, small firms did not receive any exemption from changes to Rule 14a-8.[226] Third, such firms would face only a 3% ownership threshold for proxy access use[227] rather than the 5% ownership threshold envisioned by the SEC's earlier proposal.[228] Stratmann and Verret found that their sample had statistically significant (1% level) negative abnormal returns of 75.3 basis points compared with the control group.[229]

Unlike the case in the other studies discussed, Stratmann and Verret argued that they could "use a control group to *precisely* identify the effect of the event."[230] Their control group consisted of firms that were only slightly larger in market value than the sample firms and were not affected by the SEC announcement like the small firms because their market capitalization was greater than $75 million.[231] In essence, they had almost a perfect control group.

Stratmann and Verret then estimated how much the SEC announcement on August 25, 2010 lowered the stock market capitalization of their sample firms.[232] They estimated the loss at $347 million after multiplying the average loss in stock market value of the sample firms (75.3 basis points) by the average market capitalization of the sample of those firms ($47 million) and by multiplying that product by the number of firms in the sample (980 firms).[233] Most importantly, the Stratmann and Verret study raises an important question regarding the effect of universal proxy

---

223 Thomas Stratmann & J. W. Verret, *Does Shareholder Proxy Access Damage Share Value in Small Publicly Traded Companies?*, 64 STAN. L. REV. 1431, 1458 (2012).

224 *Id.* at 1454.

225 *Id.*

226 *Id.*

227 *Id.*

228 *Id.* at 1450.

229 *Id.* at 1460-61.

230 *Id.* at 1436 (emphasis added).

231 *Id.* at 1458.

232 *Id.* at 1462.

233 *Id.*

access in general. That is, if it can be shown so clearly that small firms are negatively affected by a proxy access event under optimal testing conditions, why would the same result not hold for medium to large firms, which appear to be more susceptible to proxy access because of their greater institutional ownership?

C. *Implications of Empirical Studies for Universal Proxy Access*

In *Business Roundtable*,[234] a primary factor in the court's decision to vacate the SEC's universal proxy access rule was a determination that the SEC failed to adequately consider its statutory obligations because it relied on "insufficient empirical data"[235] in assessing "the economic effects of a new rule."[236] More specifically, the SEC was found to have "relied upon insufficient empirical data" by completely discounting studies "because of questions raised by subsequent studies, limitations acknowledged by the studies' authors, or [its] own concerns about the studies' methodology or scope."[237] Therefore, all these empirical studies, including that of Bhandari, Iliev, and Kalodimos,[238] will need to be considered and weighed in the promulgation of any new universal proxy access. Excluding any of them from the SEC's benefit-cost analysis will be inconsistent with the holding of *Business Roundtable*. If any are excluded, the SEC will be repeating the same mistake it made when it implemented Rule 14a-11, a mistake that will not hold up under judicial review.

So far, the empirical evidence provides weak and conflicting evidence on the value of universal proxy access. Some studies support the value of proxy access, while others do not. The results are not consistent between studies and across points in time. Moreover, contradictions between studies have yet to be reconciled. In addition, as Cohn and his colleagues point out, there is a concern that the results suffer from a lack of stationarity, a major drawback when all the empirical evidence comes in the form of event studies. [239] Finally, the findings in Cohn suggest that the value of universal proxy access is driven primarily, or at least significantly, by the wealth-enhancing effects of hedge fund activism. But this value is almost completely undone by requirements that shareholders must have held the stock in question for a minimum of three years, a holding period that is

234 Bus. Roundtable v. S.E.C, 647 F.3d 1144 (D.C. Cir. 2011).

235 *Id.* at 1150.

236 *Id.* at 1148. For a summary of the SEC's statutory obligations, *see* Memorandum from the Div. of Risk, Strategy, & Fin. Innovation of the SEC & the Office of Gen. Counsel of the SEC, to Staff of the Rulewriting Divs. & Offices, Current Guidance on Economic Analysis in SEC Rulemakings 1–4 (Mar. 16, 2012), https://www.sec.gov/divisions/riskfin/rsfi_guidance_econ_analy_secrulemaking.pdf.

237 647 F.3d 1144, 1150–51 (quoting 75 Fed. Reg.at 56,762-63).

238 *See generally*, Bhandari et al., *supra* note 14.

239 Cohn et al., *supra* note 74, at 17-18.

found in the SEC's voided universal proxy access rule and that, ironically, has become standardized in shareholder proposals on proxy access and in the proxy access bylaws implemented by boards.[240] In sum, there is not yet enough information to reject the null hypothesis concerning universal proxy access: the "preservation of managerial discretion" in the nomination of directors is wealth-enhancing for shareholders.

CONCLUSION

Until recently, the default rules of corporate and securities law have provided the board with exclusive authority to decide whether shareholder proposals on proxy access are to be included in a public company's proxy solicitation materials. However, that exclusive authority is no longer the rule because the SEC currently allows such proposals to be included. Because there is weak and conflicting empirical evidence about the value of shareholder-initiated proxy access, this study recommends that the SEC's current regime of proxy access be abandoned and urges the SEC not to put universal proxy access back on its agenda. These recommendations are efficiency-based: the Board—the locus of authority with the expertise and access to information that is not accessible to shareholders—is simply in the best position to determine whether proxy access is wealth-enhancing for shareholders. The shareholders themselves are not in such a position.

Dating at least to 1942, when the SEC first attempted to mandate proxy access for all public companies, board-initiated proxy access has been the default rule. Moreover, during this time period, it has been empirically observed that proxy access has rarely been initiated at public companies. If the parties to the corporate contract believed proxy access to be wealth-enhancing, then one would have expected to see at least some experimentation, either by boards allowing shareholders to include their nominees in the proxy solicitation materials or through the charter modification process.

It is still possible that proxy access is wealth-enhancing for shareholders but that board members are unanimously resistant because implementing proxy access is not in their personal interest. That is, the argument can be made that agency costs are inhibiting the implementation of proxy access and that universal rules allowing shareholder-initiated proxy access or providing universal proxy access are required to overcome those agency costs. To test whether that theory is correct, one must start with the null hypothesis that the "preservation of managerial discretion" in the nomination of directors is wealth enhancing for shareholders. For one to reject this null hypothesis with empirical analysis, it would appear reasonable to require multiple empirical studies that consistently reject this

[240] SHARFMAN, CRITIQUING THE CFA INSTITUTE'S REPORT ON PROXY ACCESS, supra note 15, at 2.

null hypothesis with statistically significant results over a significant period of time. This approach appears to be consistent with what was required by the court in *Business Roundtable*.

Yet such empirical evidence has yet to reject the null hypothesis. For the current regime of shareholder-initiated proxy access, the empirical evidence is limited to one event study, a study whose results cannot be considered representative of the entire universe of U.S. public companies. For universal proxy access, the empirical evidence provides conflicting results and conclusions and is subject to a lack of stationarity. Therefore, at this point, it would be reasonable for the SEC (a) to allow public companies to once again exclude shareholder proposals on proxy access from their proxy solicitation materials and (b) to continue to keep universal proxy access off its agenda.

THE SENSE IN COASE'S CRITICISM OF PIGOU: THE *CETERIS PARIBUS* CASE FOR INTERVENTION

David Campbell[*]

INTRODUCTION: COASE AND THE CRITICISM OF PIGOU

In a contribution which is of great interest to those studying the theory of regulation, Roger Backhouse and Steven Medema, two of our leading historians of economics, the latter arguably the leading authority on the work of the late Ronald Coase,[1] have sought to further revise our evaluation of Coase's criticism of A. C. Pigou.[2] This criticism, which is not merely one of the cornerstones of law and economics, but one of the pivots on which regulatory theory and practice as such turned in the 1970s, has extremely wide ranging implications. I will not address all aspects of Backhouse and Medema's argument, which broaden the issue out into an examination of the conceptions of private and public economic action underlying welfare economics and public choice economics. I will argue only that they go too far in defending Pigou and are thereby in danger of causing us to lose the great sense in Coase's criticism of him. I intend this argument to appeal across a wide range of standpoints on economic and legal policy, indeed to all standpoints that might reasonably hope to command majority support in liberal-democratic societies based on a market economy. But it is written from the position of one who is committed to a far more extensive welfare state than Coase was for, as I have previously argued,[3] even those taking such a position must learn from the reasons why Coase abandoned his own early socialism.

---

* Professor of Law, Lancaster University Law School, UK. This article was presented to The Centre for Law and Society, Lancaster University Law School in January 2015 and to The Transnational Law Institute, Dickson Poon School of Law, King's College London in November 2015. I am grateful for comments made at those presentations, especially those of Peer Zumbansen, and for the comments of Matthias Klaes and Guy Oakes.

1 Though now over twenty years old, the centrepiece remains Medema's biography of Coase, though in a number of respects Medema's views have shifted. STEVEN G. MEDEMA, RONALD H. COASE (1994).

2 Roger E. Backhouse & Steven G. Medema, *Economists and the Analysis of Government Failure: Fallacies in the Chicago and Virginia Interpretations of Cambridge Welfare Economics*, 36 CAMBRIDGE J. ECON. 981 (2012). *See also* STEVEN G. MEDEMA, THE HESITANT HAND 59-72 (2009); Steven G. Medema, *Pigou's 'Prima Facie Case': Welfare Economics in Theory and Practice*, *in* NO WEALTH BUT LIFE: WELFARE ECONOMICS AND THE WELFARE STATE IN BRITAIN, 1880-1945 42 (Roger E. Backhouse & Tamotsu Nishizawa eds., 2010).

3 David Campbell & Matthias Klaes, *The Principle of Institutional Direction: Coase's Regulatory Critique of Intervention,* 29 CAMBRIDGE J. ECON. 263, 264 (2005).

What might be called the first phase of the interpretation of *The Problem of Social Cost* was dominated by an attempt to utilize the "Coase Theorem" as a direct guide to public policy formulation. With the benefit of hindsight, this was a bad mistake, not least because it was entirely contrary to Coase's own views as they developed over the course of writing that article.[4] Once this mistake had been overcome and Coase's own, as it were, "anti-Coaseanism"[5] had been appreciated, it seemed clear that Coase had made a number of telling points against Pigou.

In particular, though very substantial work of interpretation proved to be needed before the centrality of the criticism of Pigou to *The Problem of Social Cost*[6] became clear, it eventually became understood that Coase had argued that Pigou's case for intervention rested on the logical fallacy that identification of a "market failure" itself justified intervention. In and of itself, this should never be enough. Any such justification must also show that intervention will produce a superior level of welfare. This second stage of the argument for intervention was, Coase argued, often ignored because it was assumed, rather than properly argued, that the public institutions necessary to successfully carry out the intervention were, or could be made, available and would work as envisaged. Coase later called policy made on the basis of this assumption "blackboard economics;" the policy will work on the blackboard, unfortunately it may well be that it cannot be put into practice.[7] Proper choice of economic policy required, *inter alia*, a concept of "government failure" to balance the concept of market failure that Pigou, though he did not use the term, had made central to welfare economics.[8]

The unexpected significance of this, which even became the basis of a line of left-wing advocacy of Coase, the first contribution to which was made by Schlag,[9] was that Coasean arguments were able to be marshalled against the concept of deregulation taken (almost) literally. Coase, as we shall see, very much wished to reduce the extent of intervention, but he was insistent that "[a] call for sensible regulation is not a call for no regulation."[10] Coase generally worked with a concept of "regulation" as "the establishment of the legal framework within which economic activity

---

4 Steven G. Medema, *Legal Fiction: The Place of the Coase Theorem in Law and Economics*, 15 ECON. & PHIL. 209, 209-10 (1999).

5 Robert C. Ellickson, *The Case for Coase and Against 'Coaseanism'*, 99 YALE L.J. 611 (1989).

6 David Campbell, *Ronald Coase's The Problem of Social Cost*, 35 U. QUEENSLAND L.J. 75 (2016).

7 Ernest W. Williams, Jr. & Ronald H. Coase, *Discussion*, 54 AM. ECON. REV. 192, 195 (1964).

8 *Id.*

9 Pierre Schlag, *An Appreciative Comment on Coase's 'The Problem of Social Cost': A View from the Left*, 1986 WIS. L. REV. 919, 921 (1986). Schlag reviews the development of this line of argument in Pierre Schlag, *Coase Minus the Coase Theorem: Some Problems with Chicago Transaction Cost Analysis*, 99 IOWA L. REV. 175 (2013).

10 Ronald H. Coase, *Social Cost and Public Policy*, *in* EXPLORING THE FRONTIERS OF ADMINISTRATION 33, 40 (George A. Edwards ed., 1970).

is carried out."[11] The great width of this concept has been criticised for causing a loss of focus, but it is the foundation of a productive relationship between economics and law because it has the great advantage of drawing attention to the necessity of a public regulatory framework, even for private economic action.[12] On this basis, the literal notion of "deregulation" is shown to be wholly untenable. Defensible economic and legal policy (other than doing nothing) has to take the form of an even-handed choice between market, firm and government (or hybrids) as "alternative form[s] of economic organisation,"[13] all of which involve extensive, active and continuous work of regulatory design.

It is certainly the case that Coase's later views were overwhelmingly conservative. As he put it in 1974:

> [T]here have been more serious studies made of government regulation of industry in the last fifteen years or so, particularly in the United States, than in the whole preceding period . . . The main lesson to be drawn from these studies is clear: they all tend to suggest that the regulation is ineffective or that, when it has a noticeable impact, on balance the effect is bad . . . I have come to the conclusion that the most probable reason we obtain these results is that the government is attempting to do too much—that it operates on such a gigantic scale that it has reached the stage at which, for many of its activities, as economists would say, the marginal product is negative.[14]

Even though I am, as I say, committed to a far more extensive welfare state than Coase, I have long found Coase's views such as these to be a very wise check on the extent of our recourse to Pigouvian intervention.[15] While one might often disagree with Coase, one had to recognise that aspiration towards even-handedness was central to his work; a quality certain other influential Chicagoans certainly failed to exhibit to the same extent.[16] I have thought it very significant that there is little mention in the immense

---

11 Ronald H. Coase, *Advertising and Free Speech*, 6 J. LEGAL STUD. 1, 5 (1977): "The term 'regulation' . . . is often confined to the work of the [executive], but regulation is also the result of legislative and judicial actions, and it seems ill-advised not to take these into consideration."

12 Ronald H. Coase, *The Institutional Structure of Production*, *in* ESSAYS ON ECONOMICS AND ECONOMISTS 11 (1994).

13 Ronald H. Coase, *The Problem of Social Cost*, *in* THE FIRM, THE MARKET AND THE LAW 95, 115 (1986).

14 Ronald H. Coase, *Economists and Public Policy*, *in* ESSAYS ON ECONOMICS AND ECONOMISTS, *supra* note 11, at 61-62.

15 David Campbell, *Luhmann Without Tears: Complex Economic Regulation and the Erosion of the Market Sphere*, 33 LEGAL STUD. 162, 183 (2013) (reviewing Gunther Teubner, NETWORKS AS CONNECTED CONTRACTS (2011)).

16 *See* David Campbell, *On What is Valuable in Law and Economics*, 8 OTAGO L. REV. 489 (1996); David Campbell, *Welfare Economics for Capitalists: The Economic Consequences of Judge Posner*, 33 CARDOZO L. REV. 2233 (2012).

literature[17] of the way that the smoke nuisance, an example of a "harm" which was very important in Pigou's thinking and to Coase's critique of that thinking in *The Problem of Social Cost*, was actually dealt with in that article. Though the difficulties of devising intervention against the smoke harm are set out in a way which indeed makes them seem very daunting,[18] to the extent that it makes any policy recommendation at all, *The Problem of Social Cost* concludes that it "would seem particularly likely" that "governmental . . . regulation [would] lead to an improvement in economic efficiency" in the case of the "smoke nuisance [when] a large number of people is involved and the costs of handling the problem through the market or the firm may be high."[19]

But now following earlier work by Medema[20] and others such as Hovenkamp[21] and Simpson,[22] Backhouse and Medema argue that Coase's criticism has much less power than was thought, and indeed is very unfair. Pigou did not in fact think that the identification of an externality justified intervention. In his two books which constitute the foundation of modern welfare economics, *Wealth and Welfare*[23] and *The Economics of Welfare*,[24] Pigou explicitly stated that such identification raised only a "*prima facie*" case for intervention, and this case could "become more than a *prima facie*" case only after consideration of "the qualifications . . . which governmental agencies may be expected to possess for intervening advantageously."[25] Rather than assume the perfection of those agencies, Pigou concluded that they might not have the requisite capacity to bring about a welfare improvement, and thus intervention would not be justified. The point could not be put more succinctly than it was in the analytic contents of *The Economics of Welfare*, "the mere failure of private industry when left free

---

17 Despite the subtle way the smoke harm was handled by Stigler in his textbook treatment of Coase, that was the actual source of the Coase Theorem. *See* GEORGE J. STIGLER, THEORY OF PRICE 113-14 (3d ed. 1996).

18 COASE, *supra* note 13, at 151-53.

19 *Id.* at 118. *See also* Ronald H. Coase, *The Federal Communications Commission*, 2 J. L. & ECON. 1, 29 (1959).

20 *See, e.g.,* Nahid Aslanbeigui & Steven G. Medema, *Beyond the Dark Clouds: Pigou and Coase on Social Cost*, 30 HIST. POL. ECON. 601 (1998).

21 *See, e.g.,* Herbert Hovenkamp, *Coase, Institutionalism, and the Origins of Law and Economics*, 86 IND. L. J. 499 (2011).

22 A. W. Brian Simpson, *The Story of* Sturges v. Bridgman, in PROPERTY STORIES 12 (Gerald Korngold & Andrew P. Morriss eds., 2009). *See also* David Campbell & Matthias Klaes, *What Did Ronald Coase Know About the Law of Tort?*, 39 MELBOURNE UNIV. L. REV. 793 (2016).

23 ARTHUR C. PIGOU, WEALTH AND WELFARE 247 (1912).

24 ARTHUR C. PIGOU, THE ECONOMICS OF WELFARE 331 (Transaction Books 2002). The first edition of *The Economics of Welfare*, an expanded revision of *Wealth and Welfare*, was published in 1920. Substantial revisions were made in 1924 and 1928 until the text was established in the 4th ed. of 1932. A '"5th ed."' in 1952 made no changes to the main text but added 11 appendices. *See infra* text accompanying note 44.

25 *Id.* at 332.

from public interference, to maximise the national dividend, does not of itself warrant intervention; for this *might* make things worse."[26]

It is unarguable that this undermines Coase's criticism of Pigou in *The Problem of Social Cost*,[27] and Backhouse and Medema build on this by showing that Henry Sidgwick and Alfred Marshall, Pigou's predecessors in founding the Cambridge School of Economics,[28] had made just the same point, and so "the Cambridge welfare economists were far from naïve about . . . government intervention."[29] One of Coase's most telling strictures has been that "[u]ntil we reali[z]e that we are choosing between social arrangements which are all more or less failures, we are not likely to make much headway."[30] But Backhouse and Medema show that there is much to indicate that the Cambridge economists saw the value of a similar "choice of evils" approach.[31] Putting to one side Sidgwick's[32] and Marshall's[33] views to this effect, Pigou unarguably displayed an awareness of the possible shortcomings of government institutions which seem to anticipate points now central to public choice economics, such as the inadequacy of governmental competence, the self-interest of public servants, and regulatory capture. In an important chapter of *The Economics of Welfare* on "Intervention by Public Authorities," to which we shall return, Pigou tells us that:

> It is not sufficient to contrast the imperfect adjustments of unfettered private enterprise with the best adjustment that economists in their studies can imagine. For we cannot expect that any public authority will attain, or will even wholeheartedly seek, that ideal. Such

---

26 *Id*. at xix.

27 In Ronald H. Coase, *The Firm, the Market and the Law*, *in* THE FIRM, THE MARKET AND THE LAW, *supra* note 13, at 20. Coase acknowledged that he had not dealt with the *prima facie* case argument in *The Problem of Social Cost* and tried to do so in a way the discussion of which is the central issue of this article. *See infra* text accompanying note 44. Even after reading (I believe) his every published word on this issue, it is unclear to me whether Coase was actually aware of the *prima facie* case argument when he wrote *The Problem of Social Cost*, and it seems most likely he wasn't. In relationship to this point, Aslanbeigui and Oakes' criticisms of Coase's method of reading of Pigou are at their most telling. ASLANBEIGUI & OAKES, ARTHUR CECIL PIGOU 165-66 (2015). Though I do not want to discuss this at all, it also does seem that Coase's criticism of Pigou may have been tinged by a dislike of aspects of Pigou's personal and professional conduct, in forming an unfavourable opinion about which view Coase was not alone.

28 Roger E. Backhouse, *Sidgwick, Marshall and the Cambridge School of Economics*, 38 HIST. POL. ECON 1, 15-44 (2006).

29 Backhouse & Medema, *supra* note 2, at 985.

30 Williams Jr. & Coase, *supra* note 7, at 195.

31 Backhouse & Medema, *supra* note 2, at 983-85.

32 Henry Sidgwick, *Economic Socialism*, *in* MISCELLANEOUS ESSAYS AND ADDRESSES 200, 206 (Forgotten Books 2012).

33 ALFRED MARSHALL, INDUSTRY AND TRADE 1, 496 (Cosimo Classics 2011).

> authorities are liable alike to ignorance, to sectional pressure and to personal corruption by private interest.[34]

It was only against a background of undoubted improvements in the quality and capacity of the British local and central state since "the days of Adam Smith"[35] that Pigou was able to advocate increased intervention. As he put it in all editions of *The Economics of Welfare*: "[T]here is now a greater likelihood that any given piece of interference, by any given governmental authority, will prove beneficial than there was in former times."[36]

THE *CETERIS PARIBUS* CASE FOR INTERVENTION

The just quoted claim about a growth in government capacity in *The Economics of Welfare* was prefigured in *Wealth and Welfare*,[37] and Backhouse and Medema draw our attention to its statement there. They claim that Pigou found evidence for this claim in the development of "[q]uasi-governmental entities . . . not directly subject to political control . . . [and] well suited to the regulation of industry . . . [which are insulated] from electoral pressures and provide the continuity in policy-making absent from government."[38] Pigou is very brief about this in *Wealth and Welfare*, telling us only that the serious disadvantages of government institutions "can, be, in great measure, obviated by the recently developed invention of 'Commissioners.'"[39] In *The Economics of Welfare*, he says at somewhat greater length that "regular governmental agencies" have:

> disadvantages [which] are all serious. But all of them can be, in great measure, obviated [they] can be overcome, perhaps even more effectively, by the recently developed devices of Commissions or ad hoc Boards, that is to say, bodies of men appointed for the express purpose of industrial operation or control. An example of a Commission for operation is afforded by the Railway Department of New South Wales or the Port of London authority in this country, and one of a commission for control by the Interstate Railway Commission of the United States.[40]

Backhouse and Medema acknowledge that "[t]his optimism regarding the prospects for future government intervention might be taken as justifying the charge that the Cambridge welfare economists were naïve."[41] But, in essence, in light of their evidence about the *prima facie* case, they accept

---

34 PIGOU, *supra* note 24, at 332.

35 Arthur C. Pigou, *State Action and Laisser-faire*, *in* ECONOMICS IN PRACTICE 107, 126 (1935).

36 PIGOU, *supra* note 24, at 333.

37 PIGOU, *supra* note 2323, at 249.

38 Backhouse & Medema, *supra* note 2, at 992.

39 PIGOU, *supra* note 23, at 250.

40 PIGOU, *supra* note 24, 334.

41 Backhouse & Medema, *supra* note 2, at 992.

only a very watered down version of this charge of naïveté; perhaps they do not accept it at all. But surely they have not done enough here, for they were, in my opinion, obliged at this point to address what I believe is one of Coase's two most successful criticisms of Pigou himself,[42] as opposed to the later Pigouvian tradition of welfare economic arguments for intervention, which is set out in the introduction Coase wrote for *The Firm, the Market and the Law*, the first, published in 1986, of two books of selections he made from his papers.[43]

Coase says that Pigou's reference to commissions, a position which Backhouse and Medema refer to as optimistic, seems "laughable to us today."[44] This is sufficiently illustrative of the differences of their evaluations of this aspect of Pigou. Coase's reasons for saying this are, however, that:

> Pigou's belief . . . was first expressed in Wealth and Welfare in 1912 and repeated in all [five] editions of The Economics of Welfare without change. Pigou never seems to have thought it necessary to inquire whether his optimistic opinion about these commissions was justified by events in the subsequent forty years (the 1952 reprint [of the fourth edition] is the last edition to contain new material). In all editions the Interstate Commerce Commission is referred to as the Interstate Railway Commission, and this body, created in 1887, is always described as "recently developed", which does not suggest any real interest in the subject.[45]

These points are, in my opinion, entirely accurate and a devastating criticism of Pigou's "manner of working" in regard of institutional analysis,[46] which justifies Coase's belief that his original criticism of Pigous was "essentially correct."[47] As we shall see, Pigou did make numerous comments on the capacity of the state in *Wealth and Welfare*, *The Economics of Welfare* and in other works, but they are all, when not outright superficial, not remotely detailed enough to actually be regarded as adequate institutional analyses of the possibility of getting the blackboard proposal to work. They are, I submit, *ceteris paribus* economics.

Wealth and Welfare is a large book of almost 500 pages. *The Economics of Welfare* is an immense book, which in its first edition was almost double that size. In both, the single chapter, which is really the only substantial discussion of the *prima facie* case, is five pages in the former ("State Intervention") and seven pages in the latter ("Intervention by Public

---

[42] The other is his criticism of Pigou's treatment of sparks from steam locomotives. COASE, *supra* note 12, at 23. Though the late Professor Brian Simpson was strongly of the opinion that Coase had got the law about this wrong, I will merely say without argument that I believe my account of it is essentially right. David Campbell, *Of Coase and Corn: A (Sort of) Defence of Private Nuisance*, 63 MOD. L. REV. 197, 207 (2000); Campbell & Klaes, *supra* note 22, at 843 n.182.

[43] The other is Coase, ESSAYS ON ECONOMICS AND ECONOMISTS, *supra* note 12 (1994).

[44] COASE, *supra* note 13, at 23.

[45] *Id.*

[46] *Id.*

[47] *Id.* at 20.

Authorities").[48] What Pigou was actually doing by making this mere gesture[49] was using rhetoric that I have previously called "the *ceteris paribus* case for intervention:"[50]

> A lawyer might presume to add something to Coase's contribution to the analysis of policy formulation by saying something further about the way the mischief of blackboard economics actually is visited on policy decisions. For, of course, no one engaged in economic (and by extension legal and social) policy formulation ever admits that they are at the blackboard. Rather, they enter all sorts of caveats about the application of the blackboard outcomes and then proceed regardless with policy proposals the basic nature of which follows from their being formulated on blackboard assumptions. There is both an acknowledgement, as a caveat, of the institutional engineering which would be necessary to make those proposals effective, and a perfect disregard of that engineering in the proposals put forward. (The proposals would, of course, be changed in the process of ever accommodating that engineering.)[51]

Though I do not think it would add to the theoretical point I am trying to make to extensively discuss any figure other than Pigou, my opinion is that Backhouse and Medema, rather than rebut the charge of naiveté levelled against the Cambridge School, effectively trace the use of *ceteris paribus* reasoning back to its sources in Sidgwick and Marshall.

One might add to Coase's criticism of Pigou's overall positive view of the capacity of public bodies some evaluation of the evidence Pigou gave for maintaining that opinion. The only such evidence in the chapter on intervention in *The Economics of Welfare* is a brief quotation from what has become a famous 1907 article of Marshall's, based on his after-dinner address to that year's Congress of the Royal Economic Society, on economic chivalry.[52] Quoting Marshall, Pigou wrote:

> [D]uring the past century in England there has been 'a vast increase in the probity, the strength, the unselfishness, and the resources of government . . . And the people are now able to rule their rulers, and to check class abuse of power and privilege, in a way which was

[48] Pigou does apply what he says of the pros and cons of public bodies to a number of the specific issues he discusses, such as "Public Operation of Industries" in chapter 17 of *Wealth and Welfare* and chapter 22 of *The Economics of Welfare*. Backhouse & Medema, *supra* note 2, at 989. But this hardly supports Backhouse and Medema's argument because this application is always unmodified by its specific context and is, indeed, always little more than formulaic.

[49] Campbell, *supra* note 42, at 200-01.

[50] My use of *ceteris paribus* obviously must be distinguished from other uses, and, in this connection, particularly from the role Alfred Marshall gave it in partial equilibrium analysis. ALFRED MARSHALL, PRINCIPLES OF ECONOMICS 304 (8th ed. 1920).

[51] Campbell, *supra* note 43, at 204.

[52] Alfred Marshall, *The Social Possibilities of Economic Chivalry*, *in* MEMORIALS OF ALFRED MARSHALL 323 (Arthur C. Pigou ed., 1925). This version slightly revises the first published version in *The Economic Journal* for 1907. Marshall worked much of the material of this address into the revisions of the last chapter of the *Principles* from the fifth edition in 1907. MARSHALL, *supra* note 50, at 592-601.

> impossible before the days of general education and a general surplus of energy over that required for earning a living.'[53]

This "important fact" was the entire evidence which Pigou gave for the claim we have seen him make, "that there is now a greater likelihood that any given piece of interference, by any public authority, will prove beneficial than there was in former times."[54]

Now, it must be said that this passage in Pigou does not properly convey what Marshall meant. Marshall was trying to explain the Victorian growth of belief in collectivism, specifically the socialism of J. S. Mill, and he was doing so in order to caution that:

> [W]e can now safely venture on many public undertakings which a little while ago would have been . . . unworkable . . . on the other hand, this very enlargement opens out so many and so arduous new public duties that no Government . . . can nearly catch up the work that is specially its own.[55]

Though Marshall attempted to redefine *laissez faire* in such a way as to allow for the government doing "work . . . which none but government can do efficiently,"[56] and was in this sense a "socialist," the overall tenor of his paper is to caution against "the anti-social influences likely to result from governmental enterprise in matters where the private hand is competent for action" and the "[s]ocial disaster [which] would probably result from the full development of the collectivist programme."[57]

But it is not so much that Pigou's entire evidence for the quality of public bodies in the crucial chapter of *The Economics of Welfare* is based on a misinterpretation of an after dinner address by Marshall;[58] it is that a brief quotation from an after dinner address *is* his entire evidence. Believing that this was good enough, it is no wonder that Pigou got the name of the Interstate Commerce Commission wrong. I do not wish to be mean spirited about the work of a man whose achievement is indisputably enormous. Nevertheless, this clearly shows that shortcomings in the way cases are made for Pigou's interventionist approach "cannot be ascribed to

---

53 PIGOU, *supra* note 24, at 333 (quoting Marshall, *The Social Possibilities of Economic Chivalry*, *supra* note 50, 335-36).

54 Pigou's previously mentioned reference to changes since "the days of Adam Smith" would also seem to be derived from Marshall's chivalry essay. Pigou, *supra* note 34, at 126; Marshall, *The Social Possibilities of Economic Chivalry*, *supra* note 50, at 334.

55 Marshall, *The Social Possibilities of Economic Chivalry*, *supra* note 52, at 336.

56 *Id.*

57 *Id.* at 337, 339. What seems a penetrating and fair evaluation of the balance of Marshall's address is put forward in PETER GROENEWEGEN, A SOARING EAGLE: ALFRED MARSHALL 1842-1924, 608-09 (1995).

58 In other work, Pigou paid due regard to the criticism of collectivism in Marshall's views, the force of which he acknowledged. ARTHUR C. PIGOU, SOCIALISM VERSUS CAPITALISM 79-87 (1937).

a few slips in analysis [but stem] from defects in the current approach to problems of welfare economics."[59]

Coase concluded *The Problem of Social Cost* by telling us that "a change of approach" is needed.[60] Coase did not think he had managed to bring this change about and, rather bizarrely, the tone overall of his estimation of the reception of his own work was one of disappointment. A typical comment was: "My point of view has not in general commanded assent, nor has my argument, for the most part been misunderstood."[61] This is in part a point about theory. The Coase theorem works perfectly given the general competitive equilibrium assumption of zero transaction costs, an assumption Coase thought was unhelpful in economic policy formulation: 'The world of zero transaction costs has often been described as a Coasean world . . . [n]othing could be further from the truth. It is the world of modern economic theory, one which I was hoping to persuade economists to leave.'[62] Though there has, of course, been first-rate work done, much of it by legal theorists, on the application of Coasean bargaining to empirical situations,[63] it is unarguable that he has failed to carry his theoretical point with many, if not most, economists:

> The extensive discussion [of The Problem of Social Cost] in the journals has concentrated almost entirely on the 'Coase Theorem,' a proposition about the world of zero transaction costs. This response, though disappointing, is understandable. The world of zero transaction costs, to which the Coase Theorem applies, is the world of modern economic analysis, and economists feel quite comfortable handling the intellectual problems it poses, remote from the real world though they may be . . . if I am right, current economic analysis is incapable of handling many of the problems to which it purports to give answers.[64]

Like Coase, however, I do not think this is the most important issue. What is needed is that even those who accept the importance of the appreciation of positive transaction costs in policy formulation must actually do the work of institutional analysis, not merely acknowledge its necessity as an abstract, *ceteris paribus*, point: "a change [of theoretical] approach is not enough. Without some knowledge of what would be achieved with alternative institutional arrangements, it is impossible to choose sensibly among them."[65] Such knowledge can only come from detailed empirical study of the sort at which Coase himself could excel:[66]

---

59 COASE, *supra* note 13, at 153.

60 *Id.*

61 *Id.* at 1.

62 *Id.* at 174.

63 *See, e.g.,* ROBERT C. ELLICKSON, ORDER WITHOUT LAW (1991).

64 COASE, *supra* note 13, at 15.

65 *Id.* at 30.

66 As it happens, some of Coase's own work is at the moment being subjected to powerful criticism because his overall aim when applying empirical argument was in most cases illustrative rather than systematic, and in a number of instances was based on insufficiently comprehensive reference to

"If our discussions are to have any value, our theories must have an empirical basis."[67] The point is that it is not enough to make a theoretical case for empirical evaluation of the capacity of government. One must actually do the empirical work as an essential, integral component of policy formulation. In the course of doing the work, this blackboard economics that underpin so many welfare proposals formulated in the orthodox way will tend to be eliminated and the proposal refined if not abandoned. It is not good enough just saying that the difficulties of economic policy formulation and its legal implementation will be taken into account if one then, in the typical way, ignores those difficulties and proceeds on effectively the original, blackboard basis. This indeed is merely *ceteris paribus* reasoning.

THE VIEWS OF ASLANBEIGUI AND OAKES

I have argued that Backhouse and Medema fail to take account of Pigou's slips over the Interstate Commerce Commission and their implications. These slips are only partially addressed in an authoritative account of Pigou's life and work that appeared after their article which promises to redefine the entire study of Pigou: *Arthur Cecil Pigou* by Nahid Aslanbeigui and Guy Oakes.[68] I am giving a most unbalanced view of this important book when I focus only on what I regard as an unconvincing

---

secondary literature or official material. Even his extremely successful criticism of the standard treatment of lighthouses as public goods (RONALD H. COASE, *The Lighthouse in Economics*, *in* THE FIRM, THE MARKET AND THE LAW 187-213(1988)) has been shown to have defects of this sort. *See* Elodie Bertrand, *Coase's Empirical Studies and Their Interpretations: The Case of the Lighthouse*, *in* THE ELGAR COMPANION TO RONALD H. COASE 320 (Claude Ménard & Elodie Bertrand eds., 2016). The core of Coase's work that does display specialist, in-depth study of the relevant empirical context was on the broadcasting industry in general and the British Broadcasting Corporation in particular, and this work does seem to have stood the test of time. RONALD H. COASE, BRITISH BROADCASTING: A STUDY IN MONOPOLY (1950). *See also* Steve Pratten, *Coase on Broadcasting, Advertising and Policy* 25 CAMBRIDGE J. ECON. 617 (2001); Richard Collins & Zoe Sujon, *UK Broadcasting Policy: The 'Long Wave' Shift in Conceptions of Accountability*, *in* BROADCASTERS AND CITIZENS IN EUROPE 42 (Paolo Baldi & Uwe Hasebrink eds., 2007); Andrea Prat & David Strömberg, *The Political Economy of Mass Media, in* 2 ADVANCES IN ECONOMICS AND ECONOMETRICS: TENTH WORLD CONGRESS 181 (Daron Acemoglu, Manuel Arellano & Eddie Dekel eds., 2013).

[67] Ronald H. Coase, *The Conduct of Economics: The Example of Fisher Body and General Motors*, 15 J. ECON. & MGMT. STRATEGY 255, 277 (2006). Marshall himself is in a sense a difficult case, but his published interventions on concrete issues of economic policy involved no empirical study of its possibilities (nor a theory framed in light of such study). There was, to put it this way, no Marshall on economic policy comparable to the Marshall who complemented his work on economic theory with pathbreaking work on the empirical operation of the market economy. Coase much admired Marshall's "aim . . . to understand the working of the real economic system" and the limits this led him to place on abstract reasoning in economics, but his views here are on Marshall on the market economy in general and on industrial organization in particular. COASE, *supra* note 13, at 171.

[68] ASLANBEIGUI & OAKES, *supra* note 27.

"Addendum" which seeks to rebut Coase's criticism of Pigou.[69] In connection with the Interstate Commerce Commission, Aslanbeigui and Oakes say:

> Coase possessed an acute facility for uncovering errors in The Economics of Welfare that were embarrassing and inexcusable but ultimately trivial and inconsequential. It was irresponsible for Pigou to refer to Parliamentary commissions as 'recently created' even as late as the 1952 reprint of his book. He was not the most conscientious editor of his own work . . . For [this] and other mistakes, Coase skewered Pigou. However, they left his theory of policy analysis intact and undamaged.[70]

I must simply leave it to the reader to decide whether the points made by Coase, which I have sought to affirm, evidence a seriously inadequate appreciation of the necessity of detailed institutional analysis, or evidence merely an "embarrassing" but "ultimately trivial and inconsequential" failing on Pigou's part.

However, in the course of their 'Addendum' Aslanbeigui and Oakes try to show that Pigou did make appropriately determined efforts to provide institutional detail. Their evidence is, however, very largely the chapters on intervention and other occasional statements in *Wealth and Welfare* and *The Economics of Welfare* about the capacity of the state derived from Marshall that I have discussed.[71] This is supplemented by reference to a number of Pigou's other works: the collection of lectures which includes the lecture on *State Action and Laisser-faire* already discussed;[72] a lecture on housing;[73] an essay on charitable support of the poor,[74] and an essay on 'chivalrous' employers such as the Cadburys and the Rowntrees published in his *Essays on Applied Economics*.[75] The most lengthy, and telling, quotation from all this material which Aslanbegui and Oakes marshal to their argument is the one from the chapter of *The Economics of Welfare* on "Intervention by

---

[69] *Id.* at 165-71. This addendum should be read in conjunction with Chapter Four on Pigou's views on policy analysis.

[70] *Id.* at 171. I have omitted Aslanbeigui and Oakes' acknowledgement that Coase's inclusion of Pigou amongst those he criticised over the use of the lighthouse example (Coase, *supra* note 65, at 187-91) also hit home because, in respect of the issues discussed here, this would not merit separate discussion as the Pigou's treatment of the lighthouse merely is another instance, though not so acutely mistaken, of the attitude exhibited over the Interstate Commerce Commission.

[71] ASLANBEIGUI & OAKES, *supra* note 27, at 167-68.

[72] Pigou, *supra* note 35 (discussed by ASLANBEIGUI & OAKES, *supra* note 27, at 168-69).

[73] Arthur C. Pigou, *Some Aspects of the Housing Problem*, *in* R. SEEBOHM ROWNTREE AND ARTHUR C. PIGOU, LECTURES ON HOUSING 35 (1914) (discussed by ASLANBEIGUI & OAKES, *supra* note 27, at 169-70).

[74] Arthur C. Pigou, *Some Aspects of the Problem of Charity*, *in* THE HEART OF EMPIRE 236 (Charles F. G. Masterman ed.,1973) (discussed, with reference to the 1st ed. of 1901, by ASLANBEIGUI & OAKES, *supra* note 27, at 170).

[75] ARTHUR C. PIGOU, ESSAYS ON APPLIED ECONOMICS 12 (1965) (discussed, with reference to the 1st ed. of 1923, by ASLANBEIGUI & OAKES, *supra* note 27, at 170).

Public Authorities" which I have given myself.[76] They conclude that this material is overwhelming evidence that:

> Pigou embraced an approach to policy based on detailed, case-by-case investigations, prudence, and pragmatism. His emphasis on the social structure and cultural setting in which successful policies are embedded document the institutional bent in his thinking, providing grounds for including him amongst institutional economists, albeit with a lower case 'i'.[77]

As indeed is the case with respect to Backhouse and Medema, it is not with any comfortable feeling that I take a position opposed to that of such distinguished authorities as Aslanbeigui and Oakes. Nevertheless, I believe that, lacking a concept akin to *ceteris paribus* reasoning, they have confused statements about the necessity of careful institutional analysis, and even statements of intention to do that analysis, with actually doing it. I find none of the material to which Aslanbeigui and Oakes refer us to contain such analysis, though there certainly is a theoretical statement of its necessity, but I do find policy prescriptions that show little or, in fact, no sign of being affected by such analysis after the caveat about its necessity has been entered. I am obliged to conclude that Aslanbeigui and Oakes do not fully appreciate the implications of what Coase said for the economic approach to policy making which he criticized. Coase did not want any economic policy making to proceed without the most detailed analysis of institutional possibility. One has to ask whether economic policy making normally has the capacity to do this, for it normally lacks anything like the necessary legal and social theoretical competence. In the end, I believe that Aslanbeigui and Oakes and I can take such differing views because I believe that when Coase told us he feared that "if I am right, current economic analysis is incapable of handling many of the problems to which it purports to give answers,"[78] it is best to take him completely literally, and Aslanbeigui and Oakes do not. It is with a comment on the consequences of basing policy on such analysis that I shall now conclude this article.

CONCLUSION: THE CURRENT IMPORTANCE OF ALL THIS

Pigou's political views were sympathetic to socialism but, while he seems to have acknowledged the possibility of an ultimately completely planned economy, his views in this respect were so gradualist as to have no real definition, and he was, for all practical purposes, committed to a mixed

---

[76] ASLANBEIGUI & OAKES, *supra* note 27, at 168 (quoting part of the passage from Pigou quoted in the text accompanying *supra* note 33).

[77] ASLANBEIGUI & OAKES , *supra* note 27, at 169.

[78] Coase, *The Firm, the Market and the Law*, *supra* note 27, at 15.

economy.[79] Maintaining that "the working of self-interest is generally beneficent,"[80] his conception of intervention was that it was exceptional, or, perhaps better, "piecemeal" in the sense that Popper distinguished "piecemeal" from "Utopian" social engineering, finding the former to be "methodologically sound."[81] Giving effect to Pigou's intention requires the dynamic determination of the optimal mixture of (forms of) private and public allocation:

> The general problem . . . is to ascertain how far the free play of self-interest, acting under the existing legal system, tends to distribute the country's resources in the way most favourable to the production of a large national dividend, and how far it is feasible for State action to improve upon 'natural' tendencies.[82]

There is an inherent limit to the size of the growth of the welfare state in this conception. It is the principle of the even-handed weighing of alternative governance structures that I have claimed is the core of what is valuable in Coase's approach. Only thoroughly well justified cases for intervention should succeed. In a passage, which reads as if it was written by Coase himself, Pigou even goes so far as to say that:

> The issue about which popular writers argue—the principle of laisser-faire versus the principle of state action—is not an issue at all. There is no principle involved on either side. Each particular case must be considered on its merits in all the detail of its concrete circumstance. High-sounding generalisations on these matters are irrelevant fireworks. They may have a place in political perorations, but they have none in real life. Accumulation of evidence, the balancing of probabilities, judgement of men, by these alone practical problems . . . can be successfully attacked.[83]

Though I do not seek to defend it here,[84] it is my opinion, based on more than twenty-five years of study of the theory and practice of

---

[79] ARTHUR C. PIGOU, SOCIALISM VERSUS CAPITALISM 137-38 (William Pickering 1994).

[80] PIGOU, *supra* note 24, at 128 (quoting EDWIN CANNAN, THE HISTORY OF THE LOCAL RATES IN ENGLAND 176 (2nd ed. 1912)).

[81] KARL R. POPPER, THE OPEN SOCIETY AND ITS ENEMIES 158 (5th ed. 1966).

[82] PIGOU, *supra* note 24, at viii.

[83] PIGOU, *supra* note 35, at 127-28.

[84] Examples from domestic economic policy have been eclipsed by international climate change policy, the first attempt to intervene at a truly global scale and scope in response to a perceived externality. So woeful has been the lack of attention paid to the legal and institutional foundations of this policy that it is based on a legal agreement to allow the major industrialising countries, principally China and India, to give economic growth and poverty eradication priority over emissions reduction, though, given the population, poverty and growth of particularly the former, has been enough to make global emissions reduction completely impossible from the outset. *See generally* David Campbell, Matthias Klaes & Christopher Bignell, *After Cancun: The Impossibility of Carbon Trading*, 29 U. QUEENSLAND L. J. 163 (2010). The essentially unchanged position left by the Paris Conference, which itself did not change, indeed it even worsened the position, is set out in David Campbell, *What does the Paris Agreement Actually Do?*, 28 ENERGY & ENV'T 883 (2016).

regulation in Coase's sense which has identified numerous instances of serious government failure that stem, and could only stem, from a very marked or even complete absence of proper inquiry into whether intervention was practically possible, that Coase's conservative stance has very considerable weight. Against a situation where the principle of piecemeal intervention is now besmirched, one might even say characterized, by interventions based on arguments of this quality, it was Coase's even-handedness that led to his conservatism:

> All solutions have costs, and there is no reason to suppose that government regulation is called for simply because the problem is not well handled by the market or the firm. Satisfactory views on policy can only come from a patient study of how, in practice, the market, firms and governments handle the problem of harmful effects. Economists need to study the work of the broker in bringing parties together, the effectiveness of restrictive covenants, the problems of the large-scale real-estate development company, the operation of government zoning and other regulating activities. It is my belief that economists, and policy-makers generally, have tended to over-estimate the advantages which come from governmental regulation. But this belief, even if justified, does not do more than suggest that government regulation should be curtailed. It does not tell us where the boundary line should be drawn. This, it seems to me, has to come from a detailed investigation of the actual results of handling the problem in different ways.[85]

I have mentioned above that Coase insisted that "[a] call for sensible regulation is not a call for no regulation." He said this in order to try to limit misunderstanding of his statement of his:

> belief that there should be less regulation than we now have is based on the fact that for a long time government regulation has been adopted so uncritically as a method of solving social problems that we must have a good deal of unwise regulation—regulation, that is, which costs more than the benefits it brings.[86]

Debate and policy about climate change have been heavily influenced by a report which Sir Nicholas (now Lord) Stern produced whilst a senior official in the U.K. Treasury at the request of Gordon Brown, then U.K. Chancellor of the Exchequer. *See generally* NICHOLAS STERN, THE ECONOMICS OF CLIMATE CHANGE: THE STERN REVIEW (2007). Its almost 700 pages contain no concrete discussion of the legal foundations of climate change policy at all. As such, *The Stern Review* must be one of the most influential and blatant works of blackboard economics ever written.

The treatment of "international collective action" in Pt. VI, and particularly chapter 26 of *The Stern Review* is a masterpiece of *ceteris paribus* reasoning which has allowed Lord Stern to maintain his commitment to climate change policy despite the inevitable complete failure to reach any agreement on global emissions reductions. Such an agreement, once thought necessary "to save the world" (NICHOLAS STERN, A BLUEPRINT FOR A SAFER PLANET: HOW WE CAN SAVE THE WORLD AND CREATE PROSPERITY (2009)) is, Lord Stern now tells us, not necessary at all: NICHOLAS STERN, WHY ARE WE WAITING? THE LOGIC, URGENCY, AND PROMISE OF TACKLING CLIMATE CHANGE 251 (2015). David Campbell, *What is Climate Change Policy Now Trying to Achieve?* 35 ECON. AFFAIRS 428 (2015).

[85] Coase, *supra* note 13, at 118-19.

[86] Coase, *supra* note 10, at 40.

That piecemeal intervention has gotten so out of hand that this is a sound position is in considerable part traceable to the merely gestural aspect of Pigou's conception of what the principle of even-handedness requires in its application, which is a change in approach so that proper empirical study, not *ceteris paribus* reasoning, be made central to economic policy formulation.

The consequences of this for a welfare state, which *inter alia* for this reason has not known any principled bounds,[87] must ultimately be a disastrous legitimation crisis[88] or a disastrous magnification of a current legitimation crisis. Now, the welfare state has been in one crisis or another for longer than I have been alive; T. H. Marshall gave 1952 as the year of the first crisis of the British welfare state.[89] But, fully recognizing the possibility of crying wolf, I do still wish to state that, if they want the welfare state to survive, its citizens must now clearly establish its limits. Coase has shown us that painstaking institutional analysis and design, which confirms or denies *prima facie* cases, is an essential part of the necessary work. I fear that by rather overestimating the value of what I have called *ceteris paribus* reasoning in Pigou, Backhouse and Medema (and Aslanbeigui and Oakes) detracts from the core of sense in Coase's critique of Pigou. No doubt not entirely to the liking of Coase, who in 1951 left the UK for the USA in part because of "a lack of faith in the future of socialist Britain,"[90] this core of sense is an indispensable resource for those seeking to defend the welfare state, if only they would use it, but this requires recognition that the best way now to defend the welfare state is to very significantly reduce its scope.[91]

---

[87] The fundamental issue is, as Backhouse and Medema acutely conclude, the attitude one takes towards making policy on the basis of "human improvement." *See* Backhouse & Medema, *supra* note 2, at 993. I myself do not hold to the Cambridge view in any direct way, for, though my views on concrete welfare policies are often to the left of any of the figures Backhouse and Medema discuss, I, like Orwell, think the only possible justification of socialism is that it is necessary to "preserve and even enlarge the atmosphere of liberalism." George Orwell, *Inside the Whale*, *in* 12 COMPLETE WORKS: A PATRIOT AFTER ALL 86, 110 (2001). But, however this is, surely it cannot be doubted that ensuring welfare cases for intervention are sound should be a general goal.

[88] JURGEN HABERMAS, LEGITIMATION CRISIS 69 (1976) ("economic crisis has been intercepted and transformed into a systematic overloading of the public budget . . . If governmental crisis management fails, it lags behind programmatic demands *that it has placed on itself*. The penalty for this failure is withdrawal of legitimation.").

[89] T. H. MARSHALL, SOCIAL POLICY 91-92 (1st ed. 1965).

[90] Ronald H. Coase, *Ronald H. Coase*, *in* THE LIVES OF THE LAUREATES 189, 199 (William Breit & Barry T. Hirsch eds., 5th ed. 2009).

[91] Which, as we have seen, Coase had long maintained. *See* Coase, *supra* note 14.

TOWARDS A COHERENT AND WORKABLE ANTITRUST POLICY ON VERTICAL RESTRAINTS

Daniel J. Gifford[*] *& Robert T. Kudrle*[**]

INTRODUCTION

Practitioners and observers of antitrust law commonly say that, since the "antitrust revolution" that brought more rigorous microeconomics in the 1970s, vertical restraints have been governed by a rule of reason. Yet we all know that such statements paint with a broad brush, and that a more precise and accurate picture would identify a range of (sometimes conflicting) criteria that courts use to assess the lawfulness of the principal kinds of vertical restraints. This article considers exclusive-supply arrangements, loyalty and bundled discounts and rebates, and tying arrangements.

In evaluating exclusive-supply arrangements, the courts typically speak of percentage "foreclosure," but their decisions appear to turn on other factors, such as the presence or absence of market power.[1] The criteria applied in loyalty and bundled rebate cases are currently in flux and inconsistent. Recently, some courts have been attempting to apply the *Brooke Group*'s price/cost test[2] outside of the predatory context in which it originated to other restraints such as exclusive-supply contracts and bundled discounts. Courts evaluating tying arrangements apply a nominal per se rule, applicable when the defendant possesses "market power" in the tying product market (which currently is understood to mean something more than a 30% share).[3] This per se treatment is widely recognized as unsatisfactory. Indeed, tying cases have been in need of rational standards for decades.

---

[*] Robins Kaplan Emeritus Professor of Law, University of Minnesota Law School.

[**] Orville & Jane Freeman Professor of International Trade & Investment Policy, Humphrey School of Public Affairs and the Law School, University of Minnesota.

[1] Jonathan M. Jacobson, *Exclusive Dealing, "Foreclosure," and Consumer Harm*, 70 ANTITRUST L.J. 311, 311 (2002).

[2] *See* Brooke Grp. Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209, 222-24 (1993). The *Brooke Group* predation standard requires sales below an appropriate measure of cost and a likelihood of recoupment. It is the first part of the *Brooke Group* test (i.e., relation of price to cost) that has been the focus of contention in non-predatory pricing contexts.

[3] The rule applicable to tying arrangements is called a per se rule by the courts, but it differs from other per se rules in that its application is dependent upon proof of market power, which can involve an extensive economic investigation and proof. This makes its application resemble that of rule of reason. *See* Jefferson Parish Hosp. Dist. No. 2 v. Hyde, 466 U.S. 2, 34-35 (1984) (O'Connor, J., concurring).

The state of affairs over these vertical restraints has stimulated attempts to uncover underlying principles. During the latter years of the Bush administration, the Justice Department proposed a restatement of the application of Section 2 to unilateral conduct, much of which was directed toward vertical restraints.[4] The Antitrust Modernization Commission reviewed antitrust law, made recommendations on a number of issues, including, inter alia, the antitrust evaluation of bundled discounts.[5] The European Commission has proposed a restatement of principles underlying Article 102 (dealing with unilateral conduct).[6] Various commentators have responded to the multiplicity of decisional criteria governing vertical restraints by suggesting new ones. Recently, Joshua Wright has proposed a reformulation of the foreclosure concept.[7] In this same spirit, Daniel Crane and Graciella Miralles have developed a unified approach for evaluating exclusionary vertical restraints.[8] More recently, Thomas Lambert has presented a general approach to exclusionary conduct.[9] The courts have also been moving towards a broader use of the *Brooke Group* predation standard to assess vertical restraints.[10]

Any attempt to unify or reconceptualize vertical restraints confronts a major problem in devising safe harbors that are broad enough to satisfy the urgent need for business firms to contract and otherwise arrange their affairs to carry out their primary business tasks without triggering unforeseen antitrust liabilities and which are not so narrow as to leave society vulnerable to monopoly-like behavior.

Part I will first discuss some conceptual problems that contribute to the confusion over verticals. Part II then evaluates the state of U.S. law governing vertical restraints. Part III will examine the special conceptual challenges for policy towards vertical restrictions with differentiated

---

4 *See generally* U.S. DEP'T OF JUSTICE, COMPETITION AND MONOPOLY: SINGLE-FIRM CONDUCT UNDER SECTION 2 OF THE SHERMAN ACT (2008).

5 ANTITRUST MODERNIZATION COMMISSION, Report and Recommendations 94-100 (2007).

6 *See generally Guidance on the Commission's Enforcement Priorities in Applying Article 82 of the EC Treaty to Abusive Exclusionary Conduct by Dominant Undertakings* (EC), 2009 O.J. (C 45) [hereinafter *EU Guidance*].

7 Joshua Wright, *Moving Beyond Naïve Foreclosure Analysis*, 19 GEO. MASON L. REV. 1163, 1165 (2012).

8 Daniel A. Crane & Graciela Miralles, *Toward a Unified Theory of Exclusionary Vertical Restraints*, 84 S. CAL. L. REV. 605, 607-09 (2011). They stress their focus on exclusionary restraints and ignore restraints that promote cartelization or price discrimination. Exclusionary restraints are directed at rivals, preventing or impeding them from competing. Exclusionary restraints are distinguished from exploitive restraints which generate supra-competitive prices and profits. Exclusionary restraints are the focus of the present paper as well, although price discrimination receives some attention. They consider predatory pricing; we do not. We consider such behavior sufficiently distinct to warrant separate analysis.

9 *See generally* Thomas Lambert, *Defining Unreasonably Exclusionary Conduct: The 'Exclusion of Competitive Rival' Approach*, 92 N.C. L. REV. 1175 (2014).

10 *See* Jacobson, *supra* note 1, at 311-12.

product competition, and Part IV will explore several attempts by other writers to simplify and unify the law and policy concerning vertical restraints.[11] Part V concludes with our own preferred approach.

I. PERSISTENT AMBIGUITIES

The different U.S. treatment of vertical restrictions over time can be partially explained by the emergence of the Chicago School antitrust paradigm in the 1970s. But after that paradigm was fully accepted, the courts continued to apply different analyses to the several recognized types of vertical restraints: a straightforward rule-of-reason analysis was applied to territorial and customer restrictions[12] and (after 2007) to vertical price maintenance agreements,[13] but a "substantial share" version of the rule of reason was being applied to exclusive-supply contracts,[14] and a per se rule was applied to tying contracts.[15] Widespread attention to the possibly anticompetitive impact of discounting has been afforded only since the Third Circuit's 2003 decision in *LePage's*.[16]

Tying arrangements have long presented a puzzle to antitrust observers. For many years the courts viewed them with deep hostility, repeatedly asserting that they "serve hardly any purpose beyond the suppression of competition."[17] Even today, when the courts have retreated from that extreme hostility, tying arrangements remain subject to a per se rule. Under that rule, ties entered into by a firm with market power in the market for the tying product are per se illegal. But modern analysis has established that tying contracts generally have much more to contribute to the welfare of society than has often been recognized, even during recent

---

[11] We find many insights in Sean P. Gates, *Antitrust by Analogy: Developing Rules for Loyalty Rebates and Bundled Discounts*, 79 ANTITRUST L.J. 99 (2013). A weakness in our view, however, is its copious use of the term "anticompetitive" without definition. That term is used dozens of times, while "consumer surplus" appears only once in a footnote. Our approach begins with an assumption that an increase in consumer (or total) surplus (or perhaps some hybrid) should be the sole aim of competition policy and that the failure to cleave closely to such a goal invites the time-worn confusion between protecting competition and protecting competitors. In particular, we take the view that price cutting is prima facie competitive behavior.

[12] *See generally* Cont'l T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36 (1977).

[13] *See generally* Leegin Creative Leather Prods., Inc. v. PSKS, Inc., 551 U.S. 877 (2007).

[14] Allied Orthopedic Appliances Inc. v. Tuco Health Care Grp. L.P., 592 F.3d 991, 996 (9th Cir. 2010); Omega Envtl., Inc., v. Gilbarco, Inc., 127 F.3d 57, 1162 (9th Cir. 1997).

[15] Town Sound & Custom Tops, Inc. v. Chrysler Motors Corp., 959 F.2d 468, 479 (3d Cir. 1992); Cascade Health Sols., v. PeaceHealth 515 F.3d 913 (9th Cir. 2008).

[16] LePage's Inc. v. 3M, 324 F.3d 141, 160 n.14 (3d Cir. 2003).

[17] N. Pac. Ry. Co. v. United States, 356 U.S. 1, 6 (1958) (citing Standard Oil Co. v. United States, 337 U.S. 293, 305 (1949)). In recent years the Court has moved away from such a purely negative view of tying arrangements. *See* Illinois Tool Works Inc. v. Indep. Ink, Inc., 547 U.S. 28, 36 (2006).

years.[18] Contrary to the traditional approach to tying, which has looked for "leveraging" the power of a seller over the tying product, thus disadvantaging consumers of the tied product, these analysts have shown how a tie could increase the welfare of both sellers and buyers.[19]

Because vertical practices raise antitrust issues only when they affect the structure of the entire market, we will carefully consider the confusing term "foreclosure," and the elusive notion of the "equally efficient firm" as they relate to the evaluation of market structure. These concepts are critical to modern analyses of vertical restraints; sometimes, however, they raise more questions than they answer.

Industrial organization economics developed through both its Harvard and Chicago phases largely without benefit of the "foreclosure" concept. Indeed, Richard Caves, in his influential introduction to the economics of industrial organization, observes: "foreclosure defines the event and does nothing to indicate its significance."[20] Justice Breyer has also questioned the value of the term: "virtually *every* contract to buy 'forecloses' or 'excludes' alternative sellers from *some* portion of the market, namely the portion consisting of what was bought."[21] "Foreclosure" finds great currency in legal writing, although Jonathan Jacobson devotes his influential article to the use of the term without offering a definition.[22] Instead, he argues persuasively that the term has been unfortunately aimed at market shares subject to some vertical restriction when attention should have been focused on market power.[23] In his widely-discussed article *Tying, Foreclosure, and Exclusion*, Michael Whinston also leaves the term undefined.[24] The EU Guidance to Article 82 TFEU (now Article 102) does not really offer a definition either but instead observes that "'anti-competitive foreclosure' is used to describe a situation where effective access of actual or potential competitors to supplies or markets is hampered or eliminated as a result of the conduct of the dominant undertaking whereby the dominant undertaking is likely to be in a position to profitably

---

[18] Among the important contributions to a revamped antitrust approach to tying are: Patrick Greenlee, David Reitman & David S. Sibley, *An Antitrust Analysis of Bundled Discounts*, 26 INT. J. IND. ORGAN. 1132 (2008) (their analysis of bundling can be applied to certain tying arrangements), Erik Hovenkamp & Herbert Hovenkamp, *Tying Arrangements and Antitrust Harm*, 52 ARIZ. L. REV. 925 (2010), and Dennis Carlton & Michael Waldman, *Upgrades, Switching Costs and the Leverage Theory of Tying*, 122 ECON. J. 675 (2011).

[19] *See, e.g.*, Greenlee et al., *supra* note 18; Hovenkamp & Hovenkamp, *supra* note 18.

[20] RICHARD E. CAVES, AMERICAN INDUSTRY: STRUCTURE, CONDUCT, PERFORMANCE 95 (7th ed. 1993).

[21] Barry Wright Corp. v. ITT Grinnell Corp., 724 F.2d 227, 236 (1st Cir. 1983). Judge Breyer's approach is cited and discussed in Crane & Miralles, *supra* note 8, at 633.

[22] Jacobson, *supra* note 1.

[23] *Id.* at 312-13.

[24] *See* Michael D. Whinston, *Tying, Foreclosure, and Exclusion,* 80 AM. ECON. REV. 837 (1990).

increase prices to the detriment of consumers."[25] This is useful, but it still leaves open the broader meaning of foreclosure.

Our reading of the literature suggests three broad categories of the use of "foreclosure." First, there is "trivial" foreclosure, which, at the limit, means merely that any sale by one firm prevents that sale from going to another. By extension, trivial foreclosure can also refer to the use of such practices as exclusive dealing (or loyalty discounting with similar results) in highly competitive downstream markets. The upstream seller may indeed have market power, but its vertical practices may make a negligible contribution to the increase or maintenance of that power. A second important use of the term will be called "limiting" foreclosure, a characteristic of market structure that, with various degrees of obduracy, limits entry or expansion of challenging firms into some or all of a market. This is a situation in which resellers sometimes make a significant complementary contribution to the market power of the incumbent seller or sellers and for that reason may exercise varying degrees of market power of their own.[26] Limiting foreclosure is cumulative: when several firms employ exclusive supply contracts or other similar devices, the aggregate effect may bar entry and reinforce an effective supplier oligopoly. But, the mere extent of exclusive dealing in a market is typically called "foreclosure," whether it is trivial or limiting. Finally, there is what we will call "strategic" foreclosure, which seems to be what the EU Guidance attempts to capture. This is a situation resulting from specific purposeful conduct by a dominant firm (or firms) against its—usually easily identified—would-be or struggling rivals. Michael Whinston uses "strategic foreclosure" in his influential demonstration of the use of tying by a single dominant firm to prevent profitable activity by a challenger.[27] Indeed, many of the "post-Chicago" economic models showing the possible harm of various types of vertical arrangements posit a game between one dominant firm and a challenger.[28] In contrast, limiting foreclosure may develop across an industry over time as a result of successful attempts by major market participants to increase market penetration at each other's expense with the prevention of entry or the elimination of weak players as only collateral

---

25 *EU Guidance*, *supra* note 6, at ¶ 19.

26 The economics literature deals extensively with buyers or coalitions of buyers that share in the market power of an incumbent dominant supplier. *See generally, e.g.*, Robert Innes & Richard J. Sexton, *Strategic Buyers and Exclusionary Contracts*, 84 AM. ECON. REV. 566 (1994). Dominant suppliers theoretically lack power to impose exclusive agreements on coordinated buyers, but collective-action problems vitiate buyers' resistance. *See* NicSand, Inc. v. 3M Co., 457 F.3d 534, 550 (6th Cir. 2006).

27 Whinston, *supra* note 24, at 840. We considered the use of the term "monopoloid" instead of "strategic" to stress that the distinguishing feature was unity of purpose towards challengers rather than the number of incumbents, but we decided on the more intuitive term.

28 *See, e.g.*, Eric B. Rasmusen, J. Mark Ramseyer & John Shepard Wley Jr., *Naked Exclusion*, 81 AM. ECON. REV. 1137 (1991).

results.[29] Limiting foreclosure may nonetheless increase barriers to entry and impede effective competition while not aiming specifically at entrants.

The legal literature on vertical restrictions appears to overemphasize one entry barrier, economies of scale, to the neglect of others, particularly the attraction of "brand" in differentiated-product markets. Barriers to entry were first systematically explored by Joe Bain in the 1940s and 1950s.[30] Bain wrote of economies of scale, absolute cost, and product differentiation barriers that allow incumbents to charge prices persistently in excess of their costs without attracting new firms to the industry.[31] Some of Bain's conceptualizations have been very effectively criticized, especially by George Stigler, who defined entry barriers as costs incurred by entrants that were not incurred by incumbents.[32] It is not necessary to pursue that theoretical dispute here. A widely neglected insight from the barriers literature, however, with strong implications for vertical restraint policy, is that buyer brand preferences in differentiated-product markets constitute advantages to a dominant incumbent that may—or may not—take considerable time or expense to erode. This, in turn, raises uncertainty and the cost of capital for the entrant. Accordingly, our analysis will take into account brand preferences and differentiated-product competition where an examination of these factors is necessary to a proper evaluation of vertical restraints. Typically, both judicial and scholarly approaches to vertical restraints simply acknowledge these preferences without adequately examining how they complicate the analysis. The almost exclusive emphasis in the literature on economies of scale as a barrier obscures the pervasive competitive importance of product differentiation barriers based on entrenched purchaser brand preferences. In particular, because the equally-efficient firm approach smooths over or ignores brand preferences, this preoccupation with scale misleadingly suggests that the equally-efficient firm approach has far broader applicability than is warranted.

---

[29] The foreclosures considered by the Court in Standard Stations, both the 6.7% represented by Standard's exclusive contracts and the exclusive contracts by Standard's rivals, probably constituted limited foreclosure. Since neither Standard's nor its rivals individually controlled a sufficient share of distribution to threaten any rival supplier, any foreclosure could not have constituted "strategic" foreclosure. *See* Standard Oil Co. v. United States, 337 U.S. 293, 295 (1949).

[30] *See* JOE S. BAIN, BARRIERS TO NEW COMPETITION (1956).

[31] *Id.*

[32] GEORGE J. STIGLER, THE ORGANIZATION OF INDUSTRY 70 (1968). For some of the main criticisms of Bain and alternative definitions of entry barriers, *see* R. Preston McAfee, Hugo Mialon, & Michael Williams, *What is a Barrier to Entry?*, 92 AM. ECON. REV. 461-62 (2004).

II. THE STATE OF U.S. LAW

Most current vertical-restraint controversies involve alleged "foreclosures" imposed by exclusive-supply arrangements, loyalty or bundled rebates or ties. These kinds of restraints are widely employed competitive marketing mechanisms, often having the effect of both reducing the buyer's price and reducing the seller's cost. Antitrust issues arise only in unusual situations where firms with market power use them to generate exclusionary effects that eclipse their pro-competitive impact.

The case law under Section 1 of the Sherman Act and Section 3 of the Clayton Act recognizes pro-competitive effects of exclusive-supply contracts by providing an effective safe harbor to exclusive-supply contracts that involve 40% or less of the market, making such shares "insubstantial."[33] As market share grows larger, however, a supplier becomes increasingly vulnerable to a charge of unlawful foreclosure. The presence of factors, such as the market power of the supplier (emphasized by Jacobson),[34] the length of the supply contract, substitute sources of supply, and barriers to the entry of competing suppliers then become relevant. At even higher levels of alleged foreclosure, a plaintiff may invoke the monopolization and attempted monopolization clauses of Section 2 of the Sherman Act[35], where the antitrust evaluation continues to be governed by these same factors. The ultimate issue in an exclusive-supply case is the extent to which supplier competition remains open and viable. Some cases have indicated that foreclosure will become unlawful when, *inter alia*, the foreclosure denies entrants or other rivals the minimum viable scale required to operate in a concentrated market.[36]

A recurring issue in exclusive-supply cases concerns the lawfulness of a dominant supplier capturing the most efficient distribution systems for itself. Does this raise antitrust concerns? Many courts have answered this question negatively, but the opinion of the D.C. Circuit in the *Microsoft* antitrust case, in which the firm pressured service providers to favor its internet browser, supports the contrary view.[37] Where a dominant supplier has captured the most efficient distribution system, the supplier gains a cost

---

33 McWane, Inc. v. FTC, 783 F.3d 814, 837 (11th Cir. 2015).

34 *See generally,* Jacobson, *supra* note 1, at 326-28.

35 15 U.S.C. § 2 (2012).

36 ZF Meritor, LLC v. Eaton Corp., 696 F.3d 254, 271 (3d Cir. 2012); Theme Promotions, Inc. v. News Am. Mktg. FSI, 546 F.3d 991, 1003 (9th Cir. 2008); *In re* Hypodermic Prod. Antitrust Litig., 2007 U.S. Dist. LEXIS 47437 (D.N.J. 2007).

37 U.S. v. Microsoft Corp., 253 F.3d 34, 64, 70-71 (D.C. Cir. 2001) (recognizing unlawfulness of exclusion carried out by relegating rival to less efficient distribution system and implicitly raising its costs). *See also* LePage's v. 3M Corp., 324 F.3d 141, 160 n.14 (3d Cir. 2003); Natchitoches Par. Hosp. Serv. Dist, v. Tyco Int'l, Ltd., No. 1:-5-CV-12024-PBS, 2009 WL 4061631, at *7-8 (D. Mass. Nov. 20, 2009).

advantage from the distribution contracts that may not be due to its own efficiencies, but that arises from the efficiencies of the distributors with whom it has contracted. Analogously, when a supplier's product is strongly favored over competing substitutes, the supplier can further impede the distribution of competing products by requiring exclusivity from its own dealers, as was done in *Dentsply*[38] and *McWane*.[39] This technique uses the attractiveness of the supplier's product to its dealers as a means of denying rivals access to those dealers who want to carry additional products.

Although exclusivity prevents free-riding and ensures the alignment of dealer incentives with those of the supplier, at some point these pro-competitive effects may be offset by the boost to supplier-level competition that would result from the elimination of exclusivity. This analysis fits the raising rivals' costs paradigm outlined by Thomas Krattenmaker and Steven Salop in the 1980s and has since been absorbed into mainstream analysis.[40] Such focus on the characteristics of the distribution systems over which the defendant has obtained exclusive rights also fits nicely within the framework of the existing law which has long been concerned with scarcity of distribution as an entry barrier. If distribution systems were unavailable to an entrant, the entrant or potential entrant would have to enter on both the supplier and distribution levels, thus incurring significant extra costs. *Microsoft* incorporates this same concern (of raising rivals' costs), applying it to the exclusive-supply context.[41]

The competition-thwarting issue raised by loyalty and bundled discount cases is the same as in exclusive-supply cases. Indeed, a loyalty discount, one that is tailored to the situation of the individual buyer, produces effects almost identical to exclusive-supply contracts, except for the incentives on buyers to adhere to the contract terms. The supplier in both situations typically reduces selling price in return for an exclusivity commitment. In both cases, the exact terms of the agreement typically turn on aspirations of the seller to maximize profits through its ability to price discriminate across buyers interacting with whatever bargaining power the buyer might have.

In an ordinary exclusive-supply contract, the buyer's penalty for breaching the contract diminishes over the course of the contract because a breach will expose it to liability for the seller's lost profits, a measure that shrinks as the completed deliveries (and accompanying payments) under the contract increase. As the seller is paid for each delivery, its potential for losing profits diminishes. By contrast, in a loyalty rebate arrangement, the rebate earned by the buyer over the entire contract period is paid at the end,

38 United States v. Dentsply, Int'l, Inc., 399 F.3d 181, 184-86 (3d Cir. 2005).

39 McWane, Inc. v. FTC, 783 F.3d 814, 819 (11th Cir. 2015).

40 Thomas G. Krattenmaker & Steven C. Salop, *Anticompetitive Exclusion: Raising Rivals' Costs to Achieve Power Over Price*, 96 YALE L.J. 209 (1986).

41 *Microsoft*, 253 F.3d at 51, 55-56.

when the buyer has fully performed. A buyer's breach forfeits its rebates on purchases already made, so that the penalty for a breach grows larger over the length of the contract period. One way of describing this effect is to say that the seller is extending its market power from the (contingent) discount on sales already made into the buyer's growing incentive to continue buying from the same source. [42] Indeed, the contingency governing loyalty rebates is referred to in Europe as a "suction effect" because of the heavy pressure on the buyer near the end of the contract period to continue purchasing from the original supplier.[43] The mechanics can generate negative prices that no rational buyer would refuse.[44] Because of their analytical similarities, loyalty discounts ought to be treated similarly to exclusive-supply contracts except where the length of the contract is a critical factor. In those cases, analytical clarity would be served by judicial recognition that the continually declining penalty for breach in an exclusive-supply contract makes it ever easier for the buyer to shop elsewhere, while the opposite is true under a loyalty rebate arrangement.

Bundled discounts can generate exclusionary effects by extending market power in one product to others, sometimes similar to sales expansion generated by single product loyalty discounts, just discussed. In *LePage's Inc. v. 3M*, the Third Circuit saw 3M's bundle pricing as a mechanism for excluding LePage's transparent tape.[45] As the court saw it, there was no way that LePage's could reduce the price of its single product to offset 3M's reduced prices on the bundle.[46] Thus 3M could achieve the results of predatory pricing without selling any item below cost.[47] In the court's view, the bundle pricing was a means of maintaining 3M's monopoly on Scotch tape.[48] But the attractiveness of the bundle could also stem not so much from the seller's monopoly on a product included in the bundle as simply from the attractiveness of the discount, especially as magnified by the size of the bundle. In this case, the sales increase would arise from the desirability of the discount to buyers, inducing them to take a larger bundle or a bundle including less desirable products, in order to qualify for the discount on the products that the buyer wants. Of course, if

---

42 See the discussion of the suction effect in DANIEL J. GIFFORD & ROBERT T. KUDRLE, THE ATLANTIC DIVIDE IN ANTITRUST: AN EXAMINATION OF US AND EU COMPETITION POLICY 123-25 (2015).

43 EUROPEAN COMMISSION, DG DISCUSSION PAPER ON THE APPLICATION OF ARTICLE 82 OF THE TREATY TO EXCLUSIONARY ABUSES 44-45 (2005).

44 *Id.*

45 LePage's Inc. v. 3M, 324 F.3d 141, 161 (3d Cir. 2003).

46 *Id.* at 161.

47 *Id.* at 160 n.14. *See also* Ortho Diagnostic Sys, Inc. v. Abbott Labs., Inc., 920 F. Supp. 455, 467 n.14 (S.D.N.Y. 1996) (describing the use of bundled discounts to achieve the results of predatory pricing without selling below cost).

48 *LePage's Inc*, 324 F.3d at 161.

the buyer wants all the products in the bundle, the bundle pricing reduces to a simple price reduction (and thus to behavior that is transparently competitive).

The courts and official bodies have been troubled by bundled discounts. In *LePage's*, the Third Circuit condemned 3M's bundle as monopolization without providing any underlying theory[49]. The Antitrust Monopolization Commission (AMC) then created a safe harbor when the discount on the entire bundle was attributed to a product in which bundler was competing with a plaintiff rival supplier and, under the restructured discount, the defendant was still pricing above cost.[50] Later, the Ninth Circuit in its *PeaceHealth* decision, turned the AMC's safe harbor into a test governing legality.[51]

Traditional judicial analysis has also recently come under challenge by a test based on the concept of an equally-efficient competitor. Do the contracts at issue prevent an equally-efficient rival supplier from competing? This criterion, first developed in predatory pricing analysis and endorsed in *Brooke Group*, was expanded in the 1990s beyond predatory pricing. In 1996, it was applied to bundled discounts,[52] and recently, the Sixth Circuit employed this analysis in *NicSand*, an exclusive-supply case.[53] There, the court thought that NicSand, which was excluded as a supplier by exclusive-supply contracts between 3M and large retailers, had the opportunity to underbid 3M and thereby gain some or all of the exclusive contracts for itself.[54] Under this approach, the courts would be taking the criterion of an equally-efficient competitor from predatory-pricing analysis and using it to evaluate other restraints that have traditionally been treated as involving non-price factors. The intuition underlying this approach is that some non-price restraints can be analytically recast as issues of price competition. In such cases, the equally-efficient competitor approach could serve to evaluate a range of restraints.

Unfortunately, the equally-efficient competitor standard is ambiguous for non-homogenous products, even in predatory-pricing contexts. The judicial operationalization of the concept is in *Brooke Group*, a predatory-

---

49 *LePage's Inc*, 324 F.3d at 169.

50 ANTITRUST MODERNIZATION COMMISSION, *supra* note 5, at 99-100.

51 Cascade Health Sols. v. PeaceHealth, 515 F.3d 883, 910 (9th Cir. 2008). The AMC used the discount attribution rule as a safe harbor. If, after attributing the entire discount on the bundle to the competitive product, the defendant still sold the competitive product above its costs, then the case would be dismissed. If, after the attribution of the discount, the defendant's price were below its cost, then the case would proceed to examine the recoupment and market effects of the practice. The Ninth Circuit, however, made the application of the discount attribution rule a substantive part of the offense. *Id.* at 910.

52 Ortho Diagnostic Sys., Inc. v. Abbott Labs., Inc., 920 F. Supp. 455, 469 (S.D.N.Y. 1996).

53 NicSand, Inc. v. 3M Co., 507 F.3d 442 (6th Cir. 2007).

54 *Id.* at 457.

pricing case.[55] It allows an established firm's own price-cost relation to gauge whether a comparable firm would need to price below cost to match the dominant firm's pricing.[56] But if the established firm can sell a similar (perhaps functionally indistinguishable) product for a higher price-cost margin than a competitor because of greater market acceptance, then it may be able to price above some measure of its own cost while still forcing its rival to price below its own (the rival's) costs.

The very meaning of "equally efficient" is ambiguous in this context. If the criterion of efficiency is the generation of willingness to pay relative to resource use, then a higher cost firm that can sell its product for a greater excess of price over marginal or average cost (including contemporaneous promotion expenses) could still be regarded as more efficient. While antitrust law cannot easily incorporate such complexities, it can recognize them through skepticism about the sufficiency of the equally efficient firm criterion, narrowing the application of that criterion for gauging exclusionary conduct where appropriate.

Finally, a word about tying. Tying is the only vertical restraint that is at least nominally subject to a per se rule.[57] It has been a subject of antitrust attention since at least 1914, when Congress called for special scrutiny of tying arrangements in Section 3 of the Clayton Act. Decades ago, tying was understood as a form of "leverage," in which the seller used a product over which it exercised power to force buyers to purchase its brand of another product (which, in the absence of the tie, would be trading competitively).[58] Despite the discrediting of this leveraging theory in many contexts,[59] ties have remained subject to a per se rule, albeit with some defenses.

III. THE POLICY CHALLENGE OF DIFFERENTIATED PRODUCT COMPETITION

The legal literature dealing with vertical restraints tends to focus upon a minimum efficient scale and generally assumes competition in a

---

55 Brooke Grp. Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209 (1993).

56 *Id.* at 227.

57 U.S. DEP'T OF JUSTICE & FED. TRADE COMM'N, ANTITRUST ENFORCEMENT AND INTELLECTUAL PROPERTY RIGHTS: PROMOTING INNOVATION AND COMPETITION 103-04 (2007). *See also* Jefferson Parish Hosp. Dist. No. 2 v. Hyd*e*, 466 U.S. 2, 9, 16-18 (1984); Illinois Tool Works Inc. v. Indep. Ink, Inc., 126 S. Ct. 1281, 1291-92 (2006).

58 *See* Ward S. Bowman, Jr., *Tying Arrangements and the Leverage Problem*, 67 YALE L.J. 19, 25 (1957).

59 *See Id.* Bowman's article is widely believed to have begun the discrediting of leveraging theory.

homogeneous product or products that are perceived as very similar.[60] Thus that literature does not adequately emphasize the overriding importance of entry barriers generated by product differentiation in assessing the competitive consequences generated by vertical restrictions. The American antitrust cases have not recently explored the role of brand loyalty in differentiated-product competition.[61] The EU, however, has attempted to address the issue, particularly through the European Commission and its Guidance on abuses of a dominant position.[62] The European approach posits that buyers sometimes cannot switch all of their purchases from a dominant firm to a rival supplier[63] because of the demand of their own customers (i.e., the customers of the purchasers from the dominant firm). These customers want the dominant firm's "brand."[64] In European parlance, this constriction of the choice of a buyer (from a dominant firm) is described as the buyer's "non-contestable" or "required" share because it is effectively reserved to the original supplier.[65] The amount of a buyer's purchases that is open to alternative suppliers is referred to as its "contestable" share.[66] As noted, the suction effect continuously reduces the effective price of new purchases.[67] When brand preference generates a non-contestable share, this suction effect is enhanced because all of the customer's purchases (including its non-contestable purchases) generate contingent discounts. Thus, when the customer is free from brand preference considerations to buy the remaining amounts that it needs, it will also have acquired conditional discounts that may further effectively bind it to its original supplier. The European Commission approaches this problem by first determining whether the targeted discounting practiced by the dominant firm yields an effective price[68] below its own costs at a sales

---

60 *See, e.g.*, Crane & Miralles, *supra* note 8, at 608-09; Willard K. Tom, David A. Balto & Neil W. Averitt, *Anticompetitive Aspects of Market-Share Discounts and Other Incentives to Exclusive Dealing*, 67 ANTITRUST L.J. 615, 623-24 (2000); Wright, *supra* note 7, at 1166, 1185: Lambert, *supra* note 9, at 1211.

61 In the 1970s and 1980s, the FTC was exploring brand preferences and differentiated-product competition. *See* Borden Inc. v. FTC, 674 F.2d 498 (6th Cir. 1982); Kellogg Co., [1970-73 Transfer Binder] Trade Reg. Rep. (CCH) ¶ 19,898; Ethyl Corp., [1976-80 Transfer Binder] Trade Reg. Rep. (CCH) ¶ 21,579. *See* Richard Schmalensee, *Entry Deterrence in the Ready-to-Eat Cereal Industry*, 9 BELL J. ECON. 305 (1978).

62 *EU Guidance*, *supra* note 6.

63 The text treats market competition as between one incumbent and one challenger, but with additional complexity—and greater uncertainly for all concerned—the argument applies to more than one incumbent and two or more challengers.

64 *EU Guidance*, *supra* note 6, at ¶ 36.

65 *See* EUROPEAN COMMISSION, *supra* note 43, at ¶¶ 155-56 (defining "required share").

66 *See EU Guidance*, *supra* note 6, at ¶ 42 (defining "contestable share").

67 EUROPEAN COMMISSION, *supra* note 43, at 44-45.

68 An effective price in this circumstance is the list price less the discounts to which the buyer is entitled.

volume within the contestable range of total sales to a given buyer.[69] This effectively is the equally-efficient competitor approach modified to accommodate a differentiated-product market. If the Commission determines the effective price of the sales to be below the seller's costs, it will deem rival firms to have been foreclosed and will assess the legal significance of that strategic or limiting foreclosure under its standards relating to consumer harm.[70]

This analytical framework needs a closer look. The European Commission analysis implies that the supposed problem of loyalty discounts affecting single-product competition may actually involve (at least) two products, if products are defined as suggested in the U.S. merger guidelines: the ability to raise prices above cost by some substantial percentage for a sustained period.[71] Part of the dominant firm's demand is quite inelastic because of its "must have" attraction, which makes that part of its sales largely immune to competition from substitutes; another part is more elastic where it competes in a segment of the market where customers lack a strong brand preference. So viewed, the price/cost test, which is often a powerful analytical tool, can produce misleading analysis. Because the incumbent in a differentiated product market will often enjoy a product acceptance advantage over a challenger's product, using its own discounted price to calculate an effective price relative to its costs may not accurately reflect the disadvantage of a firm with similar costs but weaker demand.[72] Moreover, differentiated products may be produced at dissimilar costs. Another source of uncertainty lies in the estimate of what is "contestable." If the incumbent makes a generous assumption about the competitive potential of the challenging firm, it can justify a higher cumulative rebate percentage than would otherwise be the case even if its price-cost relation at any given scale is the same as its competitor. When these ambiguities are combined, they may challenge the persuasiveness of the as-efficient-competitor approach, despite its innovative character.

We conclude that the European Commission has developed the antitrust significance of differentiated-product competition and brand preference in ways that go beyond current U.S. analysis. But the European

69 *EU Guidance*, *supra* note 6, at ¶¶ 41-44.

70 *EU Guidance*, *supra* note 6, at ¶ 20.

71 DANIEL GIFFORD & ROBERT KUDRLE, THE ATLANTIC DIVIDE IN ANTITRUST 126-27 (2015).

72 For an insightful article anticipating some of the Commission's analysis, *see* Willard K. Tom, David A. Balto, & Neil W. Averitt, *Anticompetitive Aspects of Market-Share Discounts and Other Incentives to Exclusive Dealing*, 67 ANTITRUST L.J. 615, 628 (2000). For a discussion of price premia of branded grocery goods, *see* Johan Anselmsson, Ulf Johansson, Antonio Maranon, & Niklas Persson, *The Penetration of Retailer Brands and the impact on Consumer Prices—A Study Based on Household Expenditures for 35 Grocery Categories,* 15 J. RETAILING & CONSUMER SERVS. 42 (2008). For a discussion focusing on computer components, *see* R. Venkatesh & Vijay Mahajan, *Products with Branded Components: An Approach for Premium Pricing and Partner Selection,* 16 MARKETING SCI. 146 (1997).

Commission so far has not pursued the critical related questions that affect the applicability of the approach: How do the price-cost margins of the incumbent firm compare with those of the challenger? What are the attractions of the "non-contestable" share and how obdurate are they? What is the minimum efficient (or viable) scale of the challenger?

While the European antitrust stance on loyalty discounts differs from the usual U.S. treatment, there is reason to believe that the Federal Trade Commission (FTC) has been influenced by the European analysis, particularly in its proceeding against Intel's loyalty discounts.[73] While the FTC's Intel case was settled, these events suggest that some version of the European approach may ultimately be litigated in U.S. courts. In differentiated product markets where some portion of a buyer's purchases are "must haves," the European Commission—and perhaps the FTC—views that situation as the supplier using its power over a "must have" brand (part of the customer's "required share" in European parlance) to its advantage in the "contestable" part of its demand, where in the absence of the rebates, alternative suppliers could compete for the customer's patronage. Underlying this analysis is the fact that the European Commission—and perhaps the FTC—looks at loyalty rebates as analogous to ties, whereby the dominant supplier is able to use its power over the customer's required share to pressure the customer to divert purchases in the contestable range from alternative suppliers to that same dominant supplier.[74]

The case for applying the "as efficient competitor test" outside of the predatory pricing context depends both on acceptance of the abstract criterion and the feasibility of applying it to various other restraints. The issues raised by first point are often ignored: they concern tradeoffs between consumer welfare and efficiency. Preserving high-cost suppliers can act as a check on the pricing of a dominant supplier and thus help to preserve consumer welfare but at the cost of a resource misallocation that reduces aggregate welfare. We recognize these effects but reject the social waste inherent in such a strategy as deviating from the efficiency goals of antitrust. Current law reflects the logic that a firm need not price above

---

[73] *In re* Intel, Corp., Docket No. 9341, (FTC Oct. 29, 2010). The original FTC complaint attacked Intel for not charging prices sufficient to contribute to the recovery of its sunk costs. The consent decree seems to have abandoned this position, although a version of long run incremental cost as a minimum price standard has been suggested by some economists. *See* William J. Baumol, *Predation and the Logic of the Average Variable Cost Test*, 39 J. L. & ECON. 49 (1996); Patrick Bolton, Joseph F. Brodley, & Michael H. Riordan. *Predatory Pricing: Strategic Theory and Legal Policy*, 88 GEO. L. J. 2239 (2000).

[74] *See EU Guidance*, *supra* note 6, at ¶ 39 ("A conditional rebate granted by a dominant undertaking may enable it to use the 'non-contestable' portion of the demand of each customer (that is to say, the amount that would be purchased by the customer from the dominant undertaking in any event) as leverage to decrease the price to be paid for the 'contestable' portion of demand (that is to say, the amount for which there are price-attractive substitutes).").

some measure of its own costs.[75] Although some commentators object,[76] we think that the efficiency standard embodied in current law is sound and should be maintained.[77]

The use of an analytic framework for bundled discount analysis based on that of predatory pricing is evident in *Ortho*, a case litigated in the Southern District of New York in the mid-1990s, where the court was able to dismiss that bundled discount case because all products were being sold above cost and the plaintiff remained profitable.[78] Another attempt to employ a logic similar to that of *Brooke Group* can be seen in the Antitrust Modernization Commission's proposal for dealing with bundled discounts: a safe harbor is provided when all of a bundled discount is applied to the competitive product and that product is still sold above cost by the defendant.[79] Many others have devised schemes of varying permissiveness towards the bundled sales by a firm with one or more monopolized products.[80]

There are classes of cases where the "as efficient competitor test" does not work even when prices and costs between a dominant firm and a challenger are sufficiently similar to make a *Brooke Group*-like test possible. *Dentsply*, *Eaton*, and *McWane* are recent examples.[81]

The as-efficient-competitor approach asks whether a rival could avoid foreclosure by offering a more attractive price than the incumbent. If the rival is as efficient, it can match or undercut the incumbent. In the cited cases, this was impossible for reasons unrelated to the relative efficiencies of the incumbent and its rivals. Dentsply was the dominant supplier with 67% of the market for artificial teeth.[82] Dentsply's popularity with dental labs and its consequent sales volume enhanced its attractiveness to dealers, who understood that handling the Dentsply line helped to ensure their own high volume of sales.[83] As a consequence of the high volume of Dentsply dealers, those dealers were likely to attain a higher level of efficiency than

---

[75] The same factors that can lead to a permissive price-cost test for established sellers in rebate analysis also applies in predatory pricing situations.

[76] *See* Aaron S. Edlin, *Stopping Above-Cost Predatory Pricing*, 111 YALE L.J. 941 (2002).

[77] In fact, it suggests a *de facto* partial recognition of a total surplus standard. For a more general discussion of the way the two standards apply to U.S. antitrust, *see* Joseph Farrell & Michael L. Katz, *The Economics of Welfare Standards in Antitrust*, 2 COMPETITION POLICY INT'L 2 (2006).

[78] Ortho Diagnostic Sys, Inc. v. Abbott Labs,, Inc., 920 F.Supp. 455, 469 (S.D.N.Y. 1996).

[79] ANTITRUST MODERNIZATION COMMISSION, *supra* note 5, at 99-100.

[80] *See, e.g.,* Thomas A. Lambert**,** *Appropriate Liability Rules for Tying and Bundled Discounting*, 72 OHIO STATE L. J. 1 (2011); Daniel A. Crane, *Mixed Bundling, Profit Sacrifice, and Consumer Welfare*, 55 EMORY L. J. 423 (2006).

[81] United States v. Dentsply, Int'l, Inc., 399 F.3d 181 (3d Cir. 2005); ZF Meritor, LLC v. Eaton Corp., 696 F.3d 254 (3d Cir. 2012); McWane, Inc. v. FTC, 783 F.3d 814 (11th Cir. 2015).

[82] Dentsply's share was 67% on a unit basis and 75-80% on a revenue basis. *Dentsply,* 399 F.3d at 184.

[83] *Dentsply,* 399 F.3d at 185.

dealers in competing brands.[84] Thus, a rival supplier could not just underbid Dentsply for a dealer's patronage. The dealer would also have to be compensated for losing access to a high-selling product and the efficiency loss that would accompany its defection from Dentsply. *Dentsply* is thus a case where the demand for the supplier's product probably made the price-cost test unworkable, as there may have been no feasible price at which a rival supplier could have offered its product to dealers that would have been more attractive to a Dentsply dealer than the combination of price and product volume (and consequent dealer profit) offered by Dentsply.

Notice that the central determining factor is that Denstply's product was favored over the products of other sellers by both dealers and dental labs.[85] Dentsply thus involves differentiated-product competition. Although rival manufacturers might have made inroads over time, restrictions within the distribution system made such a gradual encroachment on Dentsply far more difficult by allowing the combination of final purchaser brand acceptance, established distributor appeal, and distributor economies of scale to reinforce each other. Any price-cost analysis done on the basis of Dentsply's own price and cost data would be essentially worthless.

Eaton is similar to *Dentsply* in that the price-cost test does not work there either, albeit for different reasons. The *Eaton* case involved supply contracts between Eaton, the principal North American supplier of heavy-duty truck transmissions, and all four of the North American manufacturers that were direct-purchasers of these transmissions.[86] Most of the supply contracts at issue ran for five-year terms, although one ran for seven years.[87] These contracts contained rebate provisions conditioned upon the buyer's purchases from Eaton reaching target percentages of its (the buyer's) entire requirements.[88] These percentages were generally around 90% except in the case of Volvo, which manufactured some its own transmissions.[89] The agreements with two manufacturers (Freightliner and Volvo) also provided that Eaton could withdraw from the arrangement if the targets were not met.[90] Eaton had entered into these contracts during a period in which the market for heavy-duty trucks was contracting and in which its monopoly over heavy-duty truck transmissions was being challenged by the plaintiff, ZF Meritor, a joint venture between Meritor and ZF AG, a European truck-

---

84 *Id.* at 192.
85 *Id.*
86 ZF Meritor, LLC v. Eaton Corp., 696 F.3d 254, 265 (3d Cir. 2012).
87 *Id.* at 286-87.
88 *Id.* at 265.
89 *Id.* at 284-85.
90 *Id.* at 265.

transmission manufacturer.[91] Due to the supply contracts, the market share of the joint venture dropped from 17% in 1999 to 8% in 2003, below the joint venture's understanding of its minimum efficient scale, and the joint venture withdrew from the market and dissolved.[92] Two years later, the market share of Meritor, which stayed in the market after the dissolution of the joint venture, fell to 4% and it also left the market.[93]

ZF Meritor brought a monopolization case against Eaton, relying partially on the Sixth Circuit's *NicSand* decision.[94] Eaton contended that its behavior was lawful since it had not sold transmissions below cost.[95] The Third Circuit in *Eaton* treated these contracts as *de facto* exclusive-supply contracts and applied the law that it believed applied to exclusive contracts.[96] Because, in the court's view, price was not the predominant method of exclusion, it refused to apply the price/cost test.[97] Rather, the court ruled that the lawfulness of these contracts was governed by the rule of reason which, as applied to exclusive-supply contracts, asked whether the contracts foreclosed more than a substantial share of the market to the seller's rivals.[98]

The Third Circuit was correct in 2012 to reject the price/cost test in *Eaton*. When Eaton announced that it might withhold supplies from those manufacturers who purchased from Eaton's European rival, there was no price that the rival could have offered to any of Eaton's customers that would have offset their loss by switching.[99] Because the North American heavy truck manufacturers needed Eaton's product (heavy-duty transmissions) in most of their own production, they had to accept the terms on which Eaton conditioned its continuing sales.[100] *Eaton*, like *Dentsply* and *McWane*, is an example of an exclusive relationship between a supplier and its customers that is so much in the interest of the participating customers that these customers cannot be wooed away by feasible supply-price adjustments. Even though the transmissions were being purchased by professional buyers, no truck manufacturer could risk a complete abandonment of a supplier with an established record in favor of a new source just because it was cheaper. Hence, all of the truck-makers were

---

91 *Eaton*, 696 F.3d at 264-65.
92 *Id.* at 264, 267.
93 *Id.* at 267.
94 *Id.*
95 *Id.* at 278.
96 *Id.*
97 *Eaton*, 696 F.3d, at 279-80.
98 *Id.* at 281.
99 *Id.* at 278.
100 *Id.*

faced with all or nothing offers that they could not refuse, and the entrant could not encroach on the incumbent.[101]

Finally, in the most recent exclusive-supply case, *McWane*, a firm that dominates approximately 90% of the market for domestic pipe fittings together with two other major suppliers, imposed exclusivity on all of its distributors.[102] According to the Federal Trade Commission (FTC), McWane's purpose was to impede the growth of a rival, Star Pipe Products, that had entered the domestic pipe fittings market.[103] The FTC saw McWane's imposition of exclusivity on its dealers as monopoly maintenance.[104] Although fittings are considered commodities in the industry,[105] McWane's distributors accepted the exclusivity that McWane imposed on them, even rejecting a larger rebate from Star, apparently because of concerns about the adequacy of Star's inventory, its quality, and the timeliness of delivery.[106] Thus, McWane's power in the domestic fittings market was the result, not of the characteristics of its product, but of the confidence of buyers in the attributes of McWane, their supplier. Like *Dentsply* and *Eaton*, *McWane* is a case where the price/cost test does not work.

Overall, the "as efficient competitor" standard cannot provide a sufficient basis for a unified theory on vertical restraints because, as we have shown, that test does not work in a range of cases (like *Dentsply*, *Eaton*, and *McWane*) where customers are unwilling to switch suppliers for feasible compensation. Where the test works (such as in circumstances like *NicSand*), however, it can play a role in an overall approach to rationalizing vertical restraint analysis.

IV. PROPOSALS FOR REVISING THE STANDARDS GOVERNING VERTICAL RESTRAINTS

We now consider several other options that aim towards a general approach to evaluating vertical restrictions: (1) the approach of the European Commission embodied in its Guidance governing the application of what is now Article 102 of Treaty on the Functioning of European Union; (2) a recent proposal by Crane and Miralles; (3) the approach

---

[101] *Id.* at 335-36 (Stating that under the distribution contracts, Eaton offered a discount to each of the buyers which was paid in the form a rebate conditioned upon the buyer meeting a purchasing target. This cast the arrangement in a loyalty-contract mode. The court, however, chose to focus on the contractual provisions that permitted Eaton to terminate its supplier role if a purchaser failed to meet its purchasing target.).

[102] McWane, Inc. v. FTC, 783 F.3d 814, 820 (11th Cir. 2015).

[103] *Id.* at 821.

[104] *Id.* at 823, 827.

[105] *Id.* at 819.

[106] *Id.* at 821.

suggested by FTC Commissioner Wright; (4) a proposal by Thomas Lambert; and (5) our approach. Although this article is primarily focused on American antitrust law, we have included a consideration of the European Commission's Guidance because (1) the European approach to antitrust issues in this area has been highly elaborated, (2) the Crane and Miralles proposal draws heavily from the Guidance, and (3) there is ground for believing that the FTC may well be disposed towards incorporating some of the EU approach.

A. *The European Commission Guidance*

The EU Guidance (the "Guidance") sets forth the standards that the Commission intends to use in enforcing the Article 102 prohibition of abuses of dominant position.[107] It deals with both price-based and non-price based exclusionary conduct and devotes several pages to specific forms of abuse, including all of the practices considered in this article.[108] When dealing with price-based exclusionary conduct, the Guidance uses the "as efficient competitor" standard.[109] Under that approach, the Guidance asks whether an as-efficient competitor would be able to match a dominant firm's offers without selling below cost.[110] Since the rival is required to be a competitor who is "as efficient" as the dominant firm, the question posed is whether the dominant firm is selling below its own costs.[111] If the dominant firm is not selling below its costs, the Commission concludes that there is no adverse impact on competition.[112] The Commission employs this approach when dealing with predation, loyalty rebates, and bundling.[113] In evaluating non-price based exclusionary conduct, the Commission reverts to inquiring whether "effective access of actual or potential competitors to supplies or markets is hampered or eliminated," enabling the dominant firm to increase prices.[114]

The European Commission has used its "as efficient competitor" standard in its approach to differentiated-product competition despite the fact that the branded offerings of various firms may command widely

107 *EU Guidance*, *supra* note 6.

108 *Id*. at ¶¶ 24, 27, 32-90.

109 *Id.* at ¶ 67.

110 *Id.* at ¶ 25.

111 *Id.* at ¶ 23.

112 *Id.* at ¶ 27.

113 *EU Guidance*, *supra* note 6, at ¶¶ 32, 47, 63. *See also EU Guidance*, *supra* note 6, at ¶ 14 (stating that exclusive dealing by dominant firms, which may involve a market share as little as 40%, is generally illegal).

114 *Id.* at ¶ 19.

different shares and may sell at varying prices.[115] Moreover, the Commission's approach simply posits a need to determine non-contestability without an exploration of its causes or its vulnerability to erosion. Nevertheless, the Commission's development of the "suction effect" and its recognition of both brand-specific and more general demands by the same buyer is challenging U.S. (and other) enforcement agencies to develop their analyses of differentiated-product competition.[116]

Much of the Commission's Guidance is directed to the economic incentives affecting a particular customer's purchasing decisions rather than to the market effects of the practice, and this may seem problematic. But this is the whole point of the suction effect: it tells us about the incentives of a particular customer to keep dealing with the same supplier. If all customers are similar and make similar purchasing decisions, then we could generalize from the impact of rebates on a particular buyer to their impact on the entire market. If the buyer in question is a highly representative buyer, this may be possible, but the generalizability of the Commission's analysis of a particular buyer to the overall market impact is not always apparent.

Despite its greater attention to analyzing the impact of practices, such as loyalty rebates, on particular customers, the Guidance contains a second step where the effects of identified restraints on the general market are assessed. The criteria that it employs in this second step resemble those traditionally employed under U.S. law to evaluate exclusive-supply contracts: "the higher the percentage of total sales in the relevant market affected by the conduct, the longer its duration, and the more regularly it has been applied, the greater is the likely foreclosure effect."[117]

The Commission most likely would have agreed with the Sixth Circuit's decision in the *NicSand* case because *NicSand* permits an easy application of the "as efficient competitor" test.[118] *Eaton*, however, can also be considered using the EU approach. There, the U.S. heavy truck producers were buying transmissions from both Eaton, their principal supplier, and ZF Meritor, a new supplier.[119] When Eaton demanded

---

115 The European Commission's Discussion Paper at ¶ 33 describes the characteristics of differentiated-product competition. The competition between Intel and AMD illustrates a circumstance in which one company (Intel) possesses greater market acceptance than its rival (AMD). Intel apparently generally sold at higher prices than AMD. *See* Crane & Miralles, *supra* note 8, at 648.

116 EUROPEAN COMMISSION, *supra* note 43, 44-45.

117 *EU Guidance*, *supra* note 6, at ¶ 20. *Compare* Tampa Elec. Co. v. Nashville Coal Co., 365 U.S. 320, 329 (1961) ("To determine substantiality in a given case, it is necessary to weigh the probable effect of the contract on the relevant area of effective competition, taking into account the relative strength of the parties, the proportionate volume of commerce involved in relation to the total volume of commerce in the relevant market area, and the probable immediate and future effects which pre-emption of that share of the market might have on effective competition therein.").

118 NicSand, Inc. v. 3M Co., 507 F.3d 442 (6th Cir. 2007).

119 ZF Meritor, LLC v. Eaton Corp., 696 F.3d 254, 305 (3d Cir. 2012).

exclusivity, the truck producers complied because they needed a substantial portion of their supplies from Eaton.[120] The Third Circuit ruled that this coerced exclusivity constituted monopolization.[121] The EU Commission would have reached the same result. Under the Commission's analysis, that would constitute the noncontestable part of each buyer's demand.[122] Eaton's use of rebates in this market might be seen by the Commission as extending its power over the noncontestable part of each buyer's demand into the contestable part, thus reducing the contestable demand available to its rival, ZF Meritor. Stage one of the Commission's Guidance analysis would thus find foreclosure of Meritor from each of the customers. Stage two would address market impact. Since the foreclosure would have left ZF Meritor with a market share that apparently was not adequate for it to attain scale economies, and ZF Meritor was at that point Eaton's only rival, the requisite anticompetitive effect of the market would be established.

Overall, the innovative EU approach disappoints in two major dimensions. First, it fails to consider the complexities that necessarily attend the application of price-cost testing to the very differentiated products that seem to concern it the most, and, at the level of the entire market, it offers little beyond the time-tested "substantial share" standard. Secondly, the EU approach does not articulate standards for limiting foreclosure, such as aggregate market foreclosure brought about by the independent (non-collusive) actions of several suppliers.

B. *The Crane-Miralles Proposal*

Crane and Miralles argue that: "[i]n every exclusionary vertical restraints case, the ultimate question should be whether the loyalty-inducing provision poses an unacceptable risk of harming consumer welfare by denying to rivals a reasonable opportunity to participate efficiently in the market and whether it does so without a sufficient efficiency justification."[123] The Crane-Miralles proposal is a reformulation of the test from the European Commission's Guidance just discussed and is an impressive attempt to simplify and unify a vertical restraints framework. First, its grounding in the conceptual framework employed by the European Union maximizes the chance that European authorities would accept it. Second, it reformulates the European Commission's concepts into language familiar to American courts, thereby increasing the chances of the proposal being adopted in the U.S. Third, the Crane-Miralles proposal nicely combines the two prevalent U.S. judicial tests governing exclusive

---

120 *Id.* at 278.

121 *Id.* at 285.

122 *EU Guidance*, *supra* note 6, at ¶ 39.

123 Crane & Miralles, *supra* note 8, at 607.

distribution: the substantial share test derived from *Standard Stations* and *Tampa Electric* that traditionally has governed exclusive supply contracts and the price-cost test derived from *Brooke Group*.[124] Finally, although the proposal incorporates two of the current U.S. judicial tests, it preferences the *Brooke Group* price-cost test by making proof of below-cost sales a sine qua non of liability where that test applies.[125] That is, if a rival supplier is able to match an incumbent supplier's price and other relevant terms to a buyer, the rival has a reasonable sales opportunity. Nevertheless, like the EU approach, their proposal glosses over some central difficulties and adds some of its own.

The Crane-Miralles proposal, like the Guidance (and the U.S. courts), employs a two-step process to determine when conduct violates the antitrust laws. The writers first develop an idiosyncratic use of the term "foreclosure."[126] Under their usage, foreclosure is present whenever an existing arrangement prevents a seller's rivals from having a "reasonable sales opportunity" to make a sale.[127] For example, in *NicSand*, the plaintiff, NicSand, had a reasonable opportunity to make the sales in question. It could have undercut 3M without selling below cost, but did not. Thus, there was no "foreclosure" because little stood in the way of NicSand displacing essentially any amount of 3M's sales, so NicSand failed step one of the Crane-Miralles test. Under that test, the case would be dismissed. Without employing the Crane-Miralles language, the Sixth Circuit in *NicSand* applied essentially their reasoning to rule in favor of 3M.

If a plaintiff fails to establish "foreclosure" (in the Crane-Miralles sense) by the defendant, he fails the first step in their analysis, and the case is over; there is no need to proceed to the second step. However, if the plaintiff succeeds in establishing foreclosure, the case moves to the second step, where the plaintiff must prove the "substantiality" of the aggregate foreclosure on the entire market.[128] Although Crane and Miralles took the "substantiality" term from the U.S. exclusive-supply cases where it originated and from the European Commission which has adopted it as a standard of assessing the market effects of a restraint, they acknowledge the term's indefiniteness and argue that it "should be given a functional,

---

124 The price-cost test is part of the Crane-Miralles "reasonable sales opportunity" component of their overall test. *See* Crane & Miralles *supra* note 8, at 634. That "reasonable sales opportunity" is then combined with the substantial share component taken from *Standard Stations* and *Tampa Guidance*.

125 Crane & Miralles, *supra* note 8, at 634-35.

126 *Id*. at 607-09.

127 *Id.* at 634 ("[A] contract or contractual provision should be deemed to foreclose some share of the market only when it prevents an equally efficient competitor from profitably offering its own set of contractual terms that the customer reasonably might choose in lieu of the defendant's terms for some increment of the market's output.").

128 *Id.* at 633.

economic definition."[129] They then translate this definition as according the plaintiff a "reasonable survival opportunity."[130] Under the reasonable survival opportunity test:

> [F]oreclosure is not problematic unless an equally efficient rival would lack a reasonable opportunity to obtain a sufficient share of the nonforeclosed portion of the market to reach minimum viable scale.[131]

Crane and Miralles largely ignore the difficulties arising from differentiated-product competition outlined above where the very concept of an equally efficient firm is brought into question when differentiated products of varying costs sell at varying prices and margins. Moreover, in the case of some vertical restraints, such as targeted loyalty discounts, that part of the market open to contestation may be hard to estimate.

While Crane and Miralles stress the relevance of the non-foreclosed part of the market growing or shrinking for minimum viable scale, they treat that crucial market feature as largely independent of the efforts of the entrant. They do not explicitly consider the possible erosion of the initially "foreclosed" part resulting from targeted expenditures by the challenger despite their explicit recognition that the acceptance of new offerings varies greatly across industries and sometimes that acceptance grows quite rapidly.[132] Thus, in their approach, the incumbent bears the burden of disproving the strength and permanency of its foreclosure. Foreclosure, however, is defined in part by its effects on rivals.

Crane and Miralles explicitly reject prevailing market share minima for foreclosure in both the U.S. and the EU. We agree with their observation that "market share numbers, picked from the air [such as 30 or 40%], are utterly arbitrary from an economic perspective. Whether foreclosure is substantial in an economic sense depends on whether the quantity of the foreclosure prevents rivals from functioning efficiently in the market."[133] In our terms, Crane and Miralles here are describing limiting foreclosure. Limiting foreclosure, of course, is necessarily measured over the entire market and (whether or not it is accompanied by strategic foreclosure) its effects can matter greatly.

After assessing its strength and effectiveness, Crane and Miralles ultimately apply a quantitative meaning to unlawful foreclosure: "[a]s a general rule, we propose that foreclosure should not be deemed substantial if the minimum viable scale, expressed in units or revenues, is less than the units or revenues in the nonforeclosed segment of the market divided by the

---

129 *Id.* at 638.

130 *Id.* at 639.

131 *Id.*

132 Crane & Mirealles*, supra* note 8, at 642.

133 *Id.* at 639.

number of competitors."[134] Their explanation suggests an assumption that each competitor, including the foreclosing incumbent or incumbents, has an equal shot at a pro-rata share of the non-foreclosed part of the market but not at the foreclosed part. They explicitly assume that "over time, the new entrant has an equal chance of winning business as every other competitor."[135] However, for Crane and Miralles, "over time" is not a very long time. Without explanation, they assume competitive parity between an entrant and an incumbent in market acceptance over a period of one year. On the other hand, they neglect that "over time" some of the foreclosed part of the market may also be vulnerable. The advantages of the foreclosing incumbent or incumbents remain unexplained parameters of the problem, although there is some recognition of the formal characteristics of the foreclosing devices themselves, such as length of exclusivity.

Crane and Miralles do not explore the implications of their quantitative standard. The larger the number of actual or potential suppliers the smaller must be minimum viable scale for the arrangement to be lawful. For example, assume there are two firms, the dominant incumbent and the potential entrant. Assume further that the former has foreclosed 60% of the market and minimum viable scale (MVS) in this market is equal to a 10% market share. The foreclosure is lawful because the non-foreclosed market (40%) divided by the two companies (the incumbent and the entrant) is 20%, which is greater than the 10% MVS. But, if there are four incumbents plus the entrant jostling for a share, then the standard deems the foreclosure illegal (40 divided by 5 = 8% < 10%). We need an explanation of why the market is deemed uncompetitive with five participants but competitive with two. The case involving five firms would appear, all else equal, to present a potentially less coordinated response to the entrant. Moreover, the plausibility of the assumption of an equal chance at the unforeclosed market by each player seems greater where the number of players is larger and the shares are more equal. Finally, suppose one of the incumbents merges with another. This appears under the Crane-Miralles proposal to make the market more amenable to new competition, a counterintuitive implication.

FTC Commissioner Joshua Wright has criticized the Crane-Miralles proposal on two main grounds. First, he suggests that the difficulty of estimating minimum viable scale makes the first part of their approach "unadministrable."[136] Wright admits that such estimates are often used in merger proceedings,[137] but we agree that they should be avoided, if possible. Wright is correct that the second prong of their approach, the attempt to determine "reasonable survival opportunity," is pitched at such a high level

---

[134] *Id*. at 643.
[135] *Id*.
[136] Wright, *supra* note 7, at 1185.
[137] *Id.* at 1186.

of abstraction as to be of little use.[138] The concept requires the identification of the factors underlying "a reasonable survival opportunity" to become operational.[139]

The Crane-Miralles proposal can also be viewed through the lens of standing. Doing so helps to identify the particular contribution of their proposal. In the first step of their analysis, the plaintiff must show that it was denied a reasonable sales opportunity. This is a winnowing role analogous to establishing standing in private antitrust actions, where the plaintiff must show (antitrust) injury.[140] The second step uses a version of the substantial-share approach derived from *Standard Stations*[141] and *Tampa*,[142] modified into "a reasonable survival opportunity" and operationalized by asking whether minimum viable scale, expressed in units or revenues, is less than the units or revenues in the non-foreclosed segment of the market divided by the number of competitors. Despite the seeming limitations of their test, Crane's and Miralles's focus on the aggregate market effects of the challenged restraint moves their proposal in the right direction. But much is left unexplored. In particular, the great variety of both products and purchasers in terms of their respective numbers, market shares, their power relationships and the significance of these factors for successful entry go largely undeveloped.

C. *Wright's Approach*

Joshua Wright's approach employs a concept—as he acknowledges, not an original idea[143]—of "But for Foreclosure" (BFF) that involves a counterfactual analysis. BFF, however, while conceptually simple, is difficult to apply. Wright proposes that either cross-section or time series data be used to determine the differences in market share that the vertical restraint appears to generate.[144] To use one of his examples, if a firm without exclusive dealing boasts a 50% share and that rises to 55% with exclusivity, then the BFF is 5%.[145] The data problems surrounding the estimation could be formidable because the competitive environment of the firm could vary far more than its distributional practices.

Consider a situation based on a prominent recent case, already discussed. Suppose that Dentsply sells exclusively through dealers that (except for the grandfathering recognized in its distribution policy) handle

---

138 *Id.* at 1185-86.

139 *Id.* at 1186.

140 Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc., 429 U.S. 477, 489 (1977).

141 Standard Oil Co. v. United States, 337 U.S. 293, 314 (1949).

142 Tampa Electric Co. v. Nashville Coal Co., 365 U.S. 320, 327-28 (1961).

143 Krattenmaker & Salop, *supra* note 40, at 214.

144 *Id.* at 1186.

145 *Id.* at 1187.

only its own brand of artificial teeth. As we understand Wright's BFF analysis, we would undertake a counterfactual analysis to ascertain (as best we could) the effects of its exclusive arrangements on its market share. In this way, we would avoid attributing market effects to the exclusive arrangements that were the result of other causes. Thus, we could try to determine what Dentsply's market share would have been absent its exclusive dealer relations to get a better sense of the market effects of the exclusive arrangements. But, Dentsply had only recently instituted its exclusive policy. The focus of the court's decision was on likely future effects. So the BFF test would show little actual foreclosure and would do little to illuminate the present or future competitive significance of the prevailing distribution system.

Despite these criticisms, it is clear that Wright's inquiry identifies an important piece of what the law should be looking for. Much of the usefulness of his test, however, depends upon whether the judicial system is generally capable of providing adequately precise counterfactual assessments.[146] Counterfactual analysis has its own analytical challenges, especially when dealing with differentiated-product competition and strong brand preferences. For example, a product strongly preferred by many buyers may be sold under an exclusive-supply contract combined with a loyalty-discount program or just under a loyalty-discount program. The cumulative discount earned through "must have" purchases may generate significant foreclosure effects on purchases in that part of their demand where brand preference is muted or non-existent (this is the "suction effect" identified by the European Commission[147]). Under Wright's BFF analysis, these purchasing restraints produce no effect on purchases that are made under the influence of the product's brand attraction, but other purchases are nonetheless influenced by the cumulative discount earned on the strongly preferred purchases. So the BFF can be a useful concept, but it may be only an initial part of addressing a real competition problem.

Even if data permit the BFF to be calculated, its competitive significance remains unclear. For example, let's assume that a firm pioneered the development of a particular product—perhaps based on patents or trade secrets—and established an 80% market share, which is subsequently eroded to 50% by competitors employing their own improvements. The previously dominant firm then introduces exclusive dealing that raises its penetration to 45%. The calculated BFF would be 5%, but the firm might be heavily insulated from future share erosion. This

---

[146] It would also be interesting and relevant to do a similar counterfactual analysis based on an assumption that Dentsply's exclusives accounted only for 40% of market sales—a standard generally recognized as lawful under Sherman Act § 1 and Clayton Act § 3—and that it employed nonexclusive distribution methods on additional sales. But this inquiry would be still more complex and perhaps beyond the capabilities of the judicial system.

[147] EUROPEAN COMMISSION, *supra* note 43, at 44-45.

would be particularly likely if nearly all of the rest of the market were in the hands of two other firms, both of which employ effectively restrictive vertical practices that protect their own market positions. And assume further that the "next big thing" in product improvement lies in the hands of a firm that has not yet entered what is an effectively foreclosed market. Under Wright's BFF approach, the innovator might be unable to enter the market despite what looks like an innocuous practice because BFF focuses on the foreclosure attributable to a particular defendant, regardless of the larger market context. In our terms, this denies the relevance of limiting foreclosure. In fact, Wright argues strongly against attention to the cumulative effects of vertical practices in a market and specifically attacks Einer Elhauge for advancing the contrary position, which Wright deems "not administrable."[148] We disagree. The impact of any exclusive contract or similar relationship is a contextual matter. How much room is left for entry depends, not on a single contract, but on all of the exclusive arrangements in the aggregate. This was apparent even in *Standard Stations* where the Court observed that, in addition to the defendant, "all the other major suppliers have also been using requirements contracts."[149] And the cumulative nature of foreclosure was recognized in the DOJ's vertical restraints guidelines that were in force from 1985 until 1993.[150] Thus, in drawing the bounds of their safe harbor, the guidelines made use of two concepts that were indices of cumulative foreclosure: the Vertical Restraints Index and the coverage ratio. The Vertical Restraints Index squares the market share of each firm that is a party to such a restraint, and adds the result.[151] The coverage ratio is the percent of each market involved in a restraint of the kind in question.[152] We think that only a market-wide look at possibly restrictive practices will suffice to serve as a welfare criterion, and an attempt to consider the entire market lies at the heart of our proposal presented below.

D. *The Lambert Contribution*

Thomas Lambert has recently attempted to unify the treatment of "Unreasonably Exclusionary Conduct."[153] Lambert suggests a general standard for exclusion employing a concept different from, but closely related to, the "equally efficient competitor" standard.[154] Lambert calls his

---

148 Wright, supra note 7, at 1184; Einer Elhauge, *Tying, Bundled Discounts, and the Death of the Single Monopoly Profit Theory*, 123 HARV. L. REV. 397, 475-77 (2009).

149 Standard Oil Co. v. United States, 337 U.S. 293, 309 (1949).

150 DOJ, VERTICAL RESTRAINTS GUIDELINES 4 CCH Trade Reg. Rep. ¶ 13,105 (1985).

151 *Id*. at ¶ 4.1, n.25.

152 *Id*. at ¶ 4.1, n. 26.

153 Lambert, *supra* note 9.

154 *Id*. at 1182.

standard "exclusion of a competitive rival" (ECR) to distinguish it from the "equally efficient competitor" or "equally efficient rival" standard.[155] The ECR standard is based on the excludability of "a rival that is as aggressive a competitor as the defendant and would be capable at minimum efficient scale (MES) . . . of matching or exceeding the defendant's productive efficiency."[156] Unless a defendant's behavior would jeopardize the existence of such a rival, the defendant's behavior would be treated as lawful. The claimed advantage of the ECR standard lies in its ability to offer safe havens for pro-competitive behavior that, because of its novelty, is vulnerable to misperception by the courts as anticompetitive.

Lambert's approach focuses on what we have called strategic foreclosure. He contends that a critical advantage of the ECR standard in making safe harbors available for behavior generating unrecognized pro-competitive effects lies in the ability of a defendant to "take steps to avoid excluding truly competitive rivals."[157] In Lambert's words:

> Consider how a competitive rival would maintain (and, if necessary, grow) its scale in the face of a competitor defendant's foreclosure-causing conduct. If truly *determined*, a rival losing a significant number of customers because of the defendant's conduct would (after exhausting all other reasonably available options for expanding its sales) seek to maintain scale by offering to become a supplier to the defendant, ultimately lowering its price to the level of its incremental cost at the scale it would achieve as a supplier, presumably MES [footnote omitted]. If the rival *could meet or beat the defendant's productive efficiency at that scale*, then any defendant that was pursuing efficiency rather than seeking to enhance its market power by foreclosing rivals would be willing to accept the rival supplier's offer [footnote omitted]. If a rival complaining of exclusion did not seek to become supplier to the defendant or did not lower its price to the level of its incremental cost at MES, then it was not a determined rival.[158]

Lambert's suggestion that a rival become a supplier to a defendant engaging in apparently exclusionary conduct is highly problematic. The defendant would then be marketing not only its own output but also those of its rival. If the rival ceased all production for its own independent sales and converted itself into the defendant's supplier, instead of separate decision-making by the defendant and the rival about quantities and prices, all such decision-making would now be consolidated in the defendant. In the situation described (where the defendant is dominant and the rival is its principal competition), this combination would be a *per se* violation of Section 1 of the Sherman Act.[159] Lambert deals with this issue by claiming

---

155 *Id.* at 1206.

156 *Id.* at 1180.

157 *Id.* at 1215.

158 *Id.* at 1215-16 (emphasis added).

159 In *United States v. Masonite Corp.*, Masonite settled patent litigation by becoming a supplier to its former competitors. United States v. Masonite Corp., 316 U.S. 265, 272-73 (1942). In this arrangement, price and output decisions were made exclusively by Masonite. This arrangement was

that actual instances of rivals becoming suppliers to a defendant would probably be rare and by claiming that collusion between a rival and a defendant (by coordinating their production levels) would be easily detected and prosecuted.[160]

Regardless of legality, Lambert does not persuasively address the absence of effective horizontal competition. The mechanism explained would apparently increase the profits of the dominant firm by lowering its cost while doing nothing directly to increase product competition. While the output of the weaker firm would expand by assumption, experience as a supplier does not equip a firm to become a formidable competitor in differentiated product markets. The antitrust laws should seek to increase welfare through vigorous competition, not to keep firms in business.

Vertical restraint policy aims to prevent arrangements between suppliers and their customers that impede competition. Different rules apply to exclusive-supply agreements, tying arrangements and loyalty and bundled discounts in the U.S., while the EU has developed some innovative approaches to vertical problems, particularly bundled discounts. Some version of an "efficient competitor" test drawn from predatory pricing cases is now widely employed on both sides of the Atlantic to evaluate several vertical restraints. This article argues, however, that most markets involve differentiated products that render the approach problematic. We examine proposals from other antitrust scholars for reforming and possibly unifying the treatment of vertical restraints. All of these approaches are found wanting, and we suggest a simplified version of the rule of reason that is keyed to changes in output and market-structure effects. This focus best equips the proposal to evaluate restraints affecting differentiated-product competition. The proposal also incorporates safe harbors that are designed to protect innovative behavior from condemnation by courts that are unfamiliar with its effects, and the burden of uncertainties falls on the plaintiff.

V. OUR PROPOSAL: THE RULE OF REASON WITH A FOCUS ON THE ULTIMATE ANTITRUST IMPERATIVE OF WELFARE

Our approach to the evaluation of the vertical restraints discussed here (exclusive-supply arrangements, loyalty discounts and rebates, bundled discounts and rebates, and tying arrangements) employs a simple version of

held to be in *per se* violation of the Sherman Act. *Id.* at 282. Lambert reverses the supply arrangement by making the rivals suppliers to the dominant firm, but the centralization of decision-making is similar.

[160] Lambert, *supra* note 9, at 1242-43.

the rule-of-reason. Has output increased or decreased?[161] Our approach uses several safe harbors. The first safe harbor applies when the unencumbered share of the entire market is more than 200% of the share of the smallest viable incumbent. The second safe harbor is the price/cost test derived from *Brooke Group*.[162] When the price/cost test can be applied, passage of that test will immunize the defendant's conduct from challenge. Third, restraints occurring in markets in which the Herfindahl index is less than 2000 fall within a safe harbor. Finally, because our approach is based upon the rule-of-reason, it presumes restraints lawful, imposing the burden of proving otherwise on the party challenging them. This presumption, while not set forth in precise metes and bounds, has the practical effect of protecting a wide swath of behavior. We explain our proposal below.

Except for tying arrangements where the defendant possesses power in the tying product market, antitrust law subjects the restraints discussed here to assessment under a nominal rule of reason standard, albeit with different evaluative criteria. Our proposal shares with each of the critiques of U.S. practice discussed above a basic commitment to rule-of-reason evaluation and extends treatment under the rule-of-reason to tying arrangements. Our version of the rule-of-reason standard, however, is closer to what we will call a modernized Brandeis standard, in that it focuses directly on market impact, eschewing preliminary evaluative steps. And the European Commission has also been pursuing a version what American antitrust observers might call a rule-of-reason in its pursuit of "fact based" evaluation of challenged restraints. But agreement on a rule-of-reason approach by scholars, judges, and other officials at this abstract level does not extend to concrete applications. This is evident in all of the proposals and judicial practices we have reviewed. At these less abstract levels, the differences in the several proposals are quite substantial. In current antitrust usage, these differences could be described as different approaches to a structured rule of reason.[163] We think an application of the rule-of-

---

161 This criterion, of course, is merely shorthand for an increase in consumer surplus. Hence, if a practice results in higher output for one product but less for another, a more careful examination of the demand functions involved would be called for.

162 Brooke Grp. Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209, 222-24 (1993).

163 Shortened or "structured" versions of the rule-of-reason have developed over the years to assess "exploitative" restraints. A so-called "quick-look" version was employed in the *NCAA* and *Indiana Federation of Dentists* cases. FTC v. Ind. Fed. of Dentists, 476 U.S. 447 (1986); Nat'l Collegiate Athletic Ass'n v. Bd. of Regents, 468 U.S. 85 (1984). Use of a "quick look" version, however, was rejected in *California Dental Association* where the Court thought that the subject-matter was too complex and ambiguous for a "quick look." Cal. Dental Ass'n v. FTC, 526 U.S. 756, 781 (1999). Recently, in *Leegin,* the Court suggested that as the lower courts gain experience with a particular restraint, they will be able to "establish the litigation structure to ensure that the rule [of reason] operates to eliminate anticompetitive restraints from the market and to provide more guidance to businesses." Leegin Creative Leather Prods., Inc. v PSKS, Inc., 551 U.S. 877, 879 (2007). According to the *Leegin* opinion, the courts could do so by devising rules for offering proof, or even presumptions, making the

reason should be sought that is most likely to produce the desired result of condemning welfare-reducing behavior while protecting socially-beneficial behavior from judicial condemnation that the courts do not yet fully understand.

Justice Brandeis's comprehensive description of the rule-of-reason in his *Chicago Board of Trade* opinion[164] is a version that is both unwieldy and costly for both the parties and the courts, as critics often point out. But, at base, the rule of reason, even under the Brandeis version, is violated when output in the general market is reduced from the level that it would be in a competitive market. Brandeis understood this and pointed out both that in the case before him there was no adverse impact on prices in the general market and that the challenged restraint expanded the market and created efficiencies that allowed more grain "to arrive."

In some cases, it may be possible to measure aggregate market output to reach a conclusion as to whether the rule has been violated. But the burden of uncertainty rests on the plaintiff and thus avoids false positives. Sometimes the plaintiff's story involves uncompetitive prices throughout and sometimes it suggests a price cut followed by a price rise (as in predatory pricing). In the latter case, output should increase in the short-run but (if the plaintiff is correct) decrease in the long run. The evaluation of all anti-competitive restraints that focus on consumer or total welfare necessarily involves the courts with static or dynamic counterfactuals.[165] In particular, the burden on the courts in cases involving vertical restraints can be challenging when the relevant events may occur a year or more in the future. But such a burden is routinely borne in other realms such as predatory pricing and merger cases.

---

rule of reason "a fair and efficient" way of sorting out the anticompetitive from the procompetitive. *Id.* at 899. The Court most recently reaffirmed this approach in its *Actavis* decision. FTC v. Actavis, Inc., 133 S. Ct. 2223 (2013). There the Court again recognized the desirability of devising shortened versions of the rule of reason while avoiding the use of theories too abbreviated to permit proper analysis. 133 S. Ct. at 2238. The Court, however, left the working out of this litigation structure to the lower courts. *Id.*

164 Bd. of Trade v. United States, 246 U.S. 231, 238 (1918). In that case, Justice Brandeis described the rule of reason as follows:

> The true test of legality is whether the restraint imposed is such as merely regulates and perhaps thereby promotes competition or whether it is such as may suppress or even destroy competition. To determine that question the court must ordinarily consider the facts peculiar to the business to which the restraint is applied; its condition before and after the restraint was imposed; the nature of the restraint and its effect, actual or probable. The history of the restraint, the evil believed to exist, the reason for adopting the particular remedy, the purpose or end sought to be attained, are all relevant facts. This in not because a good intention will save an otherwise objectionable regulation or the reverse; but because knowledge of intent may help the court to interpret facts and to predict consequences.

Id.

165 Restraints can be both "exploitative" (immediate price raising) and "exclusionary" (expelling or repelling rivals), although we stress the latter effect in this article.

Our approach to a rule-of-reason evaluation of vertical restraints can be described as a direct one (a modernized Brandeisian version, as we have argued above) or as a structured rule-of-reason (where the structure focuses on output changes over time). A particular contribution of our proposal is that the plaintiff bears the burden of uncertainty. As it would apply to allegedly exclusionary restraints, the plaintiff would contend that the defendant is expanding its own output at the expense of the plaintiff and is doing so through means that cannot be justified on efficiency grounds.[166] But safe harbors provide critical winnowing.

When either exclusive dealing or loyalty discounting—or both—are pervasive in the market for a well-defined product that is sold under competitive conditions, such devices should not be suspect.[167] Thus we suggest that in markets where the Herfindahl index is less than 2000, there is unlikely to be any serious antitrust concern. In more concentrated markets, we suggest an examination of the market share of the smallest apparently viable seller in the market, whether that firm employs these vertical restrictions or not. We propose a safe harbor if exclusive dealing or targeted or bundled discounting leaves more than 200 percent of the share of the smallest viable seller unencumbered by such restrictions. This is meant to capture a situation in which there is an adequate part of the market *open* to unimpeded contestation, where "adequate" is calculated conservatively relative to the scale of an actual market participant. Where the safe harbor is inapplicable, vertical restrictions may be challenged.

Why do we prefer this safe harbor to the one contained in the now withdrawn Vertical Restraints Guidelines?[168] The Guidelines established a safe harbor if (1) the firm employing the restraint had a market share of 10% or less; or (2) the Vertical Restraints Index was below 1,200 and the coverage ratio[169] was below 60% in the same (supplier or dealer) market; or (3) the Vertical Restraints Index was below 1,200 in both the supplier and dealer markets; or (4) the coverage ratio was below 60% in both supplier and dealer markets.[170] We have no quarrel with this safe harbor. We believe, however, that our proposed safe harbor is superior because it is simpler and less cumbersome, yet it will still exclude truly problematic arrangements.

Under our approach to the rule-of-reason, all arrangements would be presumed lawful, with the plaintiff bearing the burden of proving otherwise.

---

166 As our earlier discussion of the Wright proposal argued, exploring "before" and "after" effects of a restraint is challenging, but we think that demonstrating harm is appropriately borne by the plaintiff.

167 This paper deals with market power by incumbent sellers but allows for aggressively bargaining "power buyers." Competitive suppliers dealing with powerful buyers present problems of monopsony that are beyond the scope of this work.

168 DOJ, *supra* note 150.

169 The coverage ratio is the percent of each (supplier or dealer) market involved in a restraint.

170 DOJ, *supra* note 150, at ¶ 4.1.

In particular, challenges to practices that involve price discounting would have to overcome a presumption that lower prices tend to increase output and generally benefit final purchasers. This presumption would need to be countered with a persuasive case that the apparent increase in output and accompanying price decrease were misleading and that ultimately market-wide sales to final purchasers were likely to decrease and prices were likely to rise, or have risen, as a result of the restrictions.

We propose a focus on aggregate market effects: has the market structure moved significantly in an anticompetitive direction or not? In calling for an inquiry into the probable effects of a defendant's behavior on market structure, we are focusing on the ultimate concern of the antitrust laws: the maintenance of a competitive market structure. To connect the defendant's behavior with probable market effects that occur in the future will be a challenge to the courts, but it is a challenge that they can handle. Courts currently perform similar tasks in predatory pricing cases when they resolve issues of recoupment. Thus in a predatory pricing context, recoupment is possible only if the defendant's behavior produces critical changes in market structure, changes that generate new market power and enable the defendant to recover its earlier losses. The Court in *Brooke Group*, for example, was able to decide for the defendants because the plaintiff had failed to muster evidence showing that the intensely-competitive cigarette market would be transformed into a collusive oligopoly.[171]

Our approach, like that of Crane and Miralles, the EU Guidance and at least some U.S. courts, makes use of the price/cost test used in *Brooke Group*. When the test applies, it serves as a negative filter. A defendant who passes the test is entitled to judgment. This test, however, when taken out of the specific predatory-pricing context in which it originated, either does not work (as illustrated by *Dentsply*, *Eaton*, and *McWane*) or is prone to a number of errors, so its limitations need to be recognized. Thus that test does not work in certain common situations and it ignores the problems inhering in many differentiated product markets. The Sixth Circuit employed that test in *NicSand*, and the European Commission holds itself open to using that test in its Guidance. But the European Guidance recognizes that the test does not always work. And *Dentsply*, *Eaton*, and *McWane* illustrate the limitations of that standard: in those cases the standard did not work because the more entrenched firm enjoyed such superior (and lasting) product acceptance that it was largely immune to undercutting by rivals. That standard can be defended in the abstract, but where products are highly differentiated, inputs of similar cost may be producing outputs of quite differing value as judged by the market.[172] Thus

[171] Brooke Grp. Ltd. v. Brown & Williamson Tobacco Corp., 509 U.S. 209, 238 (1993).

[172] The microprocessor market provides an example: the production costs of Intel and AMD, its main competitor, are similar, but the market acceptance of AMD lags behind that of Intel.

the "as efficient competitor" test will more likely accomplish its intended purpose where firms both produce at similar costs and sell at similar prices.[173] Where that is not the case the test will systematically tend towards type 2 errors[174] even if both the challenger and incumbent evaluate the "contestable" share accurately.

Where price-cost similarities make the price/cost analysis meaningful, we accept Baumol's average avoidable cost as a generally superior standard to average variable cost[175] in predatory and related antitrust analyses. Again, however, average avoidable cost has its limitations. Such a standard simply does not work in many circumstances in which large front end expenditures are made followed by negligible marginal costs—as in software. On the other hand, attempts to oblige firms to build any particular level of total cost recovery into their pricing through some use of Baumol's related LAIC (long run average incremental cost) concept may be unworkable. Only a full examination of the behavior of the dominant firm will suffice.

Our suggested framework addresses the apparent competitive effect of a firm's challenged vertical practice on other firms competing for sales to buyers but places all such issues in the context of their impact on the general market. It also weighs any possible impediments to competition against increased efficiency and thus judges them under a consumer (or total) [176] welfare standard. Finally, we believe that our framework can help resolve issues raised by differentiated-product competition.

Our approach deals better with differentiated-product competition than any of the other approaches reviewed here. The EU Guidance mentions established distribution channels as a barrier to entry, but otherwise product differentiation is treated only as an on-off switch to motivate buyers' purchases of "must have" products. The strength and obduracy, and not just the existence, of a brand advantage is necessary for consideration of the competitive prospects of an entrant. Such considerations as the industry's market share stability and its history of acceptance of new products must be considered directly. It is the cumulative impact of these factors that attests to the practical exclusionary impact on entrants and that, in turn, affects the competitiveness of the market structure. In most cases, strong brand preference will be unproblematic from an antitrust perspective. But the failure of antitrust analysis to develop an approach to differentiated product competition has left a gap that occasionally adversely affects judicial analysis. Thus, in *LePage's*, the Third Circuit treated 3M's "Scotch" brand

---

173 Similar, of course, does not mean identical.

174 Type 2 errors are false negatives.

175 *See generally* Baumol, *supra* note 73.

176 The simple "does output go up or down?" rule that underlies the consumer surplus standard must be broadened for the total surplus standard to apply to the economy as a whole and not just to the product market in question.

sticky tape as a monopoly, even though it was competing with LePage's physically similar product, and the technology involved was easy to master.[177] Again, although Intel's microprocessors are similar to AMD's and the two companies enjoy a duopoly in that industry, the European Commission treated Intel as a dominant firm and ignored AMD's persistent hold of 30% of the microprocessor market.[178] Since our proposal directs attention to market structure, it is uniquely equipped to assess antitrust issues raised by differentiated product competition.

If similar distributors are available to rival suppliers at a competitive price and there are no significant scale economies at the distribution level, the whole issue of exclusive dealing and discounting dissolves. In sharp contrast, if downstream commerce is not distribution but physical incorporation, a fixed constellation of buyers may simply need to be faced. Intel and AMD, for example, had no control over the configuration of the set of computer manufacturers who were their customers. Many situations fall in between. Our approach recognizes that both the choice of vertical restraints and their possible anticompetitive effects depend on the characteristics of both the upstream and downstream markets.

Our proposal avoids the complexities of Crane and Miralles in their explication of the reasonable survival test that involves comparing estimated minimum viable scale with the non-foreclosed part of the market divided by the number of competitors. We also avoid deficiencies in Joshua Wright's "but for foreclosure" that ignores aggregate market foreclosure in favor of foreclosure attributed to a particular defendant. And we avoid the antitrust problems inherent in Thomas Lambert's "exclusion of a competitive rival" test under which he would tolerate (and even encourage) a rival supplier to transform a competitive relationship to a dominant supplier into a supplier/customer relationship when necessary to maintain minimum efficient scale.

Finally, our approach forthrightly recognizes that the law governing tying must be overhauled. As argued earlier, tying is typically innocuous from the standpoint of competition policy. Moreover, it often involves product design in which another market—distribution—is not a critical part of the analysis. Sometimes the distinction is blurred: Microsoft pressured other firms into favoring its browser and media player for use with their products before testing the legality of incorporation.

We would abolish the per se approach to tying arrangements. There should be no recognized tie between perfect complements used in a fixed ratio, because they should be deemed a single product and subject to the

---

[177] LePage's Inc. v. 3M, 324 F.3d 141, 166 (3d Cir. 2003).

[178] European Commission Press Release IP/09/745, Antitrust: Commission imposes fine of €1.06 bn on Intel for abuse of dominant position; orders Intel to cease illegal practices (May 13, 2009).

strictures of the single monopoly theorem.[179] The use of complements in variable proportions should be recognized as typically a device to facilitate output expansion. Such second-degree price discrimination increases seller profits, but, as previously noted, it may often also increase consumer surplus. The practice is generally not an antitrust problem. Moreover, the sale of two "monopoly" goods tied together avoids double marginalization, an important efficiency gain.

Our rule-of-reason approach would inquire as to whether the total value of the dominant product plus the tied product is increased or decreased by the tie or, alternatively whether consumer surplus is increased across the two products. This logic applies to Microsoft's incorporation of new features into its Windows operating system, which—apart from ties that suppress the development of alternatives to its operating system (an issue that underlay the U.S. antitrust case)—should be evaluated similarly. As the *Microsoft* antitrust case has demonstrated, it is also possible that such an incorporation could, in certain unusual circumstances, generate anti-competitive effects. In the *Microsoft* case, the unusual circumstances were the combination of the seller's monopoly over the tying product (platform for software applications) with the potential of the rival to the tied product (the browser) to develop into an alternative platform, thus breaking the tying's product's platform monopoly. Our version of the rule-of-reason, whose primary focus is on the structure of the supplier market, would have identified these critical circumstances.

CONCLUSION

This paper stresses that the welfare effects of each of the practices discussed should, to the extent possible, be determinative of the lawfulness of those practices. The purpose of each of the forms of a structured rule of reason we propose is to assist the judiciary in discovering and upholding practices that increase welfare and to bar practices that decrease welfare. Production and distribution efficiency must necessarily be important criteria in these determinations because the lower the cost of producing and delivering the product, the greater will be the likely welfare effects. All of the practices considered here have been employed in highly competitive markets and such employment suggests a potential both to reduce cost and

---

[179] A clear statement of the theorem and its limitations is presented in Steven C. Salop, *Economic Analysis of Exclusionary Vertical Conduct: Where Chicago Has Overshot the Mark*, in HOW THE CHICAGO SCHOOL OVERSHOT THE MARK: THE EFFECT OF CONSERVATIVE ECONOMIC ANALYSIS ON U.S. 144-46 (Robert Pitofsky, ed. 2008).

to lower prices.[180] Thus a heavy burden of showing a reduction in consumer welfare lies appropriately with the challenger.

[180] The latter effect is less dependable and suggests that, all else equal, positive effects of efficiency will be recognized more under an aggregate welfare standard than they will be under a consumer welfare standard because the latter standard only recognizes benefits accruing to consumers.

ILLUMINATING THE OFF-LABEL FABLE: HOW OFF-LABEL PROMOTION MAY ACTUALLY HELP PATIENTS

*Colleen Conners**

INTRODUCTION

The current framework of the off-label use of pharmaceuticals is as follows: physicians are free to prescribe off-label as they see fit; the Food and Drug Administration ("FDA") acknowledges the value of off-label use (going so far as to say physicians could have an obligation to prescribe off-label in certain circumstances); but, per FDA policy, drug manufacturers are restricted from sharing truthful and non-misleading information about off-label uses.[1] The scheme, on its face, is inconsistent and, as could be expected, raises significant concerns regarding free speech, consumer protection, and public safety.

Off-label medicine is defined as, "the technical term for medicines which have not been approved [by regulatory authorities] for the therapeutic purpose for which they are prescribed."[2] And because approximately one out of every five prescriptions written is for an off-label use, it is alarming that manufacturers are barred from sharing truthful information about a particular drug's off-label uses.[3] The issue has received an increasing amount of attention over the past decade and the argument that the restrictions on off-label promotion are inconsistent with the modern interpretation of the First Amendment has since developed. Furthermore, now on two separate occasions, the first in December 2012 in *U.S. v. Caronia* and the second in August 2015 in *Amarin v. FDA*, courts in the Second Circuit have held that the restrictions on off-label pharmaceutical marketing run afoul of the First Amendment.[4] Despite the growing amount of attention being paid to the issue and the recent case law, the FDA has proceeded slowly and strategically in addressing the constitutional concerns facing the restrictions on off-label pharmaceutical marketing.[5]

In April 2014, sixteen months after the *Caronia* decision, Janet Woodcock, FDA Director of the Center for Drug Evaluation and Research

---

* J.D. Candidate 2017, George Mason University School of Law.

[1] United States v. Caronia, 703 F.3d 149, 153 (2nd Cir. 2012); *see also* Buckman Co. v. Plaintiffs' Legal Comm'n., 531 U.S. 341, 350 (2001).

[2] DAVID CAVALLA, OFF-LABEL PRESCRIBING: JUSTIFYING UNAPPROVED MEDICINE xiii (John Wiley & Sons, Ltd. ed. 2015).

[3] Amarin Pharm, Inc. v. FDA, 119 F. Supp. 3d 196, 200 (S.D.N.Y. 2015).

[4] *Caronia*, 703 F.3d at 168; *Amarin*, 119 F. Supp. 3d at 200.

[5] CAVALLA, *supra* note 2, at 141.

("CDER") first announced, "[w]e are currently evaluating our policies in light of court decisions on First Amendment issues."[6] However, this evaluation remains ongoing. Furthermore, the administrative agency has only continued to defend its restrictions on off-label promotion in litigation and continues to send warning letters to drug manufacturers threatening prosecution for off-label speech.[7] For example, in September 2015, shortly after the *Amarin* decision, Pacira Pharmaceuticals sought injunctive relief in response to one such warning letter.[8] In an amicus brief supporting Pacira Pharmaceuticals, the Medical Information Working Group ("MIWG"), an entity composed of major manufacturers in the healthcare industry, expressed its frustration in saying, "on the heels of *Caronia* and *Amarin*, FDA has yet to proffer a constitutionally permissible interpretation of its authority to regulate off-label speech."[9]

As the FDA continues to evade addressing the First Amendment challenge, support for the freedom to disperse information pertaining to off-label medicine builds. This growing support for off-label use information is of grave importance because First Amendment commercial speech protection has evolved to depend on whether a restriction on speech advances a substantial governmental interest and is no more restrictive than necessary to satisfy the interest.[10] With increasing pressure to ease the restrictions on off-label promotion, the FDA's ability to assert a substantial governmental interest is waning. This Comment highlights the mounting arguments, primarily regarding patient welfare, that call for the free flow of information pertaining to off-label medicine. In addition to the existing case law, these mounting patient welfare arguments show that drug manufacturers should be afforded First Amendment protection in off-label communications and that the FDA is fighting a losing battle in its attempt to prevent this protection.

Part I of this Comment addresses the current regulatory authority that governs off-label promotion by manufacturers and proceeds to examine the evolution of the First Amendment to protect commercial speech. Part I concludes by looking at how the courts in *Caronia* and *Amarin* resolved the conflict between the expansion of the First Amendment and the FDA's restrictions on off-label promotion and the impact of those decisions.[11] Part II of this Comment sheds light on a host of policy considerations that call for the free flow of information concerning off-label medicine. These

---

6 Jill Wechsler, *FDA Takes First Amendment Issues "Seriously*," PHARMEXEC (May 14. 2014), http://www.pharmexec.com/fda-takes-first-amendment-issues-seriously.

7 CAVALLA, *supra* note 2, at 141.

8 Complaint at 8, Pacira Pharm., Inc. v. FDA, No. 1:15-cv-07055 (S.D.N.Y. 2015).

9 Brief for Medical Information Working Group, as Amici Curiae Supporting Petitioners, at i, 14, Pacira v. FDA, No. 1:15-cv-07055 (S.D.N.Y. 2015).

10 *See* Central Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557, 564-66 (1980).

11 United States v. Caronia, 703 F.3d 149 (2nd Cir. 2012); Amarin Pharma, Inc. v. FDA, 119 F. Supp. 3d 196 (S.D.N.Y. 2015).

perspectives primarily concern patient welfare. By highlighting these high-stake issues, Part II shows that the FDA's defense of its restrictions on off-label promotion, which rests upon the existence of a substantial governmental interest, is no match for the position that off-label speech is owed First Amendment protection.

I. BACKGROUND

A. *Current Regulatory Authority Over Off-Label Marketing*

The Food, Drug, and Cosmetic Act ("FDCA") was enacted in 1938 and mandated that manufacturers submit a New Drug Application ("NDA") to the FDA and that the drug subsequently be approved for safety before entering the market.[12] However, this approval process was severely flawed. A drug manufacturer needed only to provide a relatively insignificant amount of evidence regarding a drug's safety and if the FDA refused approval the burden would be on the agency to prove the drug was unsafe.[13] Furthermore, no mandate existed pertaining to a drug's legitimacy and effectiveness.[14] This gap in the regulatory scheme resulted in a tendency for drug manufacturers to mislead in order to gain bigger profits at the expense of public safety.[15] During this period of history, "the curative claims of predatory sham medicine salesmen were limited only by the gullibility of their targets."[16] It was not until the birth of thousands of deformed children in Western Europe due to the distribution of the drug thalidomide that deceptive practices and the drug approval process were re-evaluated.[17]

To account for the aforementioned abuses and public safety concerns, Congress amended the FDCA by enacting the Kefauver-Harris Amendments of 1962.[18] These amendments, "essentially put into place the protectionist system of drug regulation that we have today."[19] The amendments require manufacturers to demonstrate that their drugs are both

---

[12] The Federal Food, Drug and Cosmetic Act of 1938, 21 U.S.C. § 301 (1938).

[13] Katharine A. Van Tassel, *Slaying the Hydra: The History of Quack Medicine, the Obesity Epidemic and the FDA's Battle to Regulate Dietary Supplements Marketed as Weight Loss Aids*, 6 IND. HEALTH L. REV. 203, 224-25 (2009).

[14] *Amarin,* 119 F. Supp. 3d at 199.

[15] *Id.* at 200.

[16] Van Tassel, *supra* note 13, at 216.

[17] *Id.* at 228.

[18] *Id.* at 229. *See* Drug Amendments of 1962, Pub. L. No. 87-781, 76 Stat. 780 (codified as amended at 21 U.S.C. § 355 (Supp. II 1982)) [hereinafter Kefauver-Harris Amendments of 1962].

[19] Katherine A. Helm, *Protecting Public Health from Outside the Physician's Office: A Century of FDA Regulation from Drug Safety Labeling to Off-Label Drug Promotion*, 18 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 117, 129 (2007).

safe and effective for their intended uses before they are approved for distribution.[20] In order to accomplish this, manufacturers must take their drugs through a series of extensive preclinical and clinical trials.[21]

It is under this regulatory scheme that off-label promotion cases have been prosecuted. Given that the FDCA requires all drugs be approved as effective *for their intended use*, the FDA has concluded that when a manufacturer promotes a drug for a use that has not been approved, the manufacturer is guilty of misbranding and as having an intent to defraud or mislead. The agency has articulated its stance in one guidance document as, "[a]n approved drug that is marketed for an unapproved use (whether in labeling or not) is misbranded because the labeling of such a drug does not include 'adequate directions for use.'"[22] Therefore, the FDA requires that before a manufacturer may promote a new use for an already approved drug, the manufacturer must go through another application process—the supplemental NDA ("SNDA")—"in which a sponsor requests the right to amend the drug's labeling and regulatory status so it is deemed 'effective' for a new indication."[23]

The FDA justifies its interpretation of the FDCA and the SNDA process by pointing to instances where prescribing off-label has led to adverse effects. For example, the drug Gabitril was prescribed off-label for psychiatric conditions but instead "caused patients to suffer seizures and status epilepticus."[24] Similarly, the drug quinine was prescribed off-label to treat nighttime leg cramps but "caused adverse reactions, including thrombocytopenia and gastrointestinal bleeding."[25]

Given such examples and the admirable objective of ensuring that drugs are proven effective for all of their intended uses, the restrictions on off-label promotion may appear reasonable and even appropriate. However, the other side of the coin is that the SNDA process creates a significant and costly delay in providing physicians and patients access to information to make the best and most informed medical decisions.[26] Furthermore, critics have identified the FDA's silencing of drug

---

[20] Van Tassel, *supra* note 13, at 228-29. *See also* Kefauver-Harris Amendments of 1962, *supra* note 18.

[21] *Id.*

[22] U.S. FOOD & DRUG ADMIN, GOOD REPRINT PRACTICES FOR THE DISTRIBUTION OF MEDICAL JOURNAL ARTICLES AND MEDICAL OR SCIENTIFIC REFERENCE PUBLICATION ON UNAPPROVED NEW USES OF APPROVED DRUGS AND APPROVED OR CLEARED MEDICAL DEVICES: GUIDANCE FOR INDUSTRY (2009).

[23] DANIEL CARPENTER, REPUTATION AND POWER: ORGANIZATIONAL IMAGE AND PHARMACEUTICAL REGULATION AT THE FDA 534 (Ira Katznelson et al. eds., 2010).

[24] Amarin Pharma, Inc. v. FDA, 119 F. Supp. 3d 196, 204-05 (S.D.N.Y. 2015).

[25] *Id.*

[26] George S. Craft, Jr., *Promoting Off-Label in pursuit of Profit: An Examination of a Fraudulent Business Model*, 8 HOUS. J. HEALTH L. & POL'Y 103, 109 (2007).

manufacturers as unconstitutional given the Supreme Court's recognition that the First Amendment protects commercial speech.[27]

B. *Commercial Speech Doctrine*

The Supreme Court decision in *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council, Inc.*[28] is recognized as having given "birth to the commercial speech doctrine."[29] *Virginia Citizens* debated the constitutionality of a 1976 Virginia law that banned pharmacists from advertising the price of prescription medicines.[30] While the state argued that the law worked to prevent aggressive advertising, which could potentially hurt consumers, a consumer group argued that consumers had a right to the pricing information.[31] In a 7-1 ruling, the Court sided with the consumer group and found that even while drug pricing information is considered "commercial speech," it still falls under the protection of the First Amendment.[32] This was a departure from earlier cases in which the Court held the First Amendment did not extend to communications that are "purely commercial."[33]

The Court found that continuing to distinguish between regular speech and commercial speech was a "simplistic approach."[34] In ruling that the disputed pricing information is constitutionally protected speech, the Court stressed the "consumer's interest[s] in the free flow of commercial information," especially in a "predominantly free enterprise economy."[35] In such an economy, consumers should be able to make "intelligent and well-informed" decisions."[36] The Court concluded by holding that so long as commercial advertisements are truthful and non-misleading, the State may not "completely suppress the dissemination" of such information.[37]

Two years after the *Virginia Citizens* decision, the Supreme Court was tasked with clarifying and determining at what point commercial speech could be silenced in *Central Hudson Gas & Electric Corp. v. Public Service*

---

27 *See* Central Hudson Gas & Electric Corp. v. Pub. Serv. Comm'n, 447 U.S. 557 (1980).

28 Virginia State Bd. of Pharm. v. Virginia Citizens Consumer Council, Inc., 425 U.S. 748 (1976).

29 Elliot Zaret, *Commercial Speech and the Evolution of the First Amendment,* DISTRICT OF COLUMBIA BAR (Sept. 2015), https://www.dcbar.org/bar-resources/publications/washington-lawyer/articles/september-2015-commercial-speech.cfm.

30 *Virginia Citizens*, 425 U.S. at 749-50.

31 *Id.* at 753-54, 768.

32 *Id.* at 762.

33 *Id.* at 755.

34 *Id.* at 759.

35 *Id.* at 763-65.

36 *Virginia Citizens*, 425 U.S. at 765.

37 *Id.* at 773.

Commission.[38] The case involved the constitutionality of a New York law that banned utility companies from promoting their services during the 1970s energy crisis.[39] In its decision the Court laid out a four-part test that must be satisfied in order to regulate or ban commercial speech: (1) the speech concerns lawful activity and is not misleading; (2) the government has a substantial interest; (3) the regulation in question "advances the governmental interest;" and (4) the regulation is "no[] more extensive than is necessary to serve that interest."[40] The Court ultimately concluded that the New York law did not satisfy all four requirements, and was therefore unconstitutional.[41]

The application of the *Central Hudson* test has become the crux of the commercial speech doctrine.[42] In the 1990s, the Washington Legal Foundation filed a series of lawsuits against the FDA, claiming that FDA policies restricting manufacturers from sharing *third-party* information pertaining to off-label uses (e.g., an article previously published in a peer-reviewed professional journal) ran afoul of the commercial speech doctrine.[43] The courts carried out the *Central Hudson* analysis and the aftermath of the suits was that "the FDA did not have the right to prevent the dissemination of truthful, non-misleading scientific and medical information, at least in the form of peer-reviewed journal articles, medical textbooks and sponsorship of continuing medical education programmes."[44]

FDA policies were also found to be in conflict with the commercial speech doctrine in *Thompson v. Western States Med. Ctr.*,[45] which concerned restrictions placed on pharmacists from advertising the practice of compounding drugs.[46] Compounded drugs have traditionally been "designed in response to an individual's unique . . . needs," such as allergies.[47] Given the highly individualized nature of these drugs, they are not automatically subject to the drug approval process.[48] The FDA argued that the restrictions on the advertising of compounded drugs served a substantial governmental interest in preserving the effectiveness and integrity of the drug approval process because advertising should be reserved for only those drugs that go through the approval process and are

---

38 *See* Central Hudson Gas & Elec. Corp. v. Pub. Serv. Comm'n, 447 U.S. 557 (1980).

39 *Id.* at 558-59; Zaret, *supra* note 29.

40 *Central Hudson*, 447 U.S. at 566.

41 *Id.* at 570-71.

42 Zaret, *supra* note 29.

43 Washington Legal Found. v. Friedman, 13 F. Supp. 2d 51 (D.D.C. 1998); Washington Legal Found. v. Henney, 128 F. Supp. 2d 11 (D.D.C. 2000), *dismissed and vacated in part*, 202 F. 3d 331 (D.C. Cir. 2000).

44 CAVALLA, *supra* note 2, at 139.

45 Thompson v. Western States Med. Ctr., 535 U.S. 357 (2002).

46 *Id.* at 360.

47 *Id.* at 360-61, 370.

48 *Id.* at 360-62.

intended for larger patient populations.[49] The Court found that the government's restriction did not satisfy the four parts of the Central Hudson test.[50] Specifically, the Court stressed that the government could not satisfy the fourth prong of the test, which requires that "if the government could achieve its interest in a manner that does not restrict speech, or that restricts less speech, the Government must do so."[51] The Court offered up several alternative ways in which the government could meet its objective in drawing a distinction between compounded drugs and FDA-approved drugs.[52]

One of the more recent decisions in addressing the ultimate question of whether off-label promotion by drug manufacturers is owed First Amendment protection is the 2011 case *Sorrell v. IMS Health*.[53] The case concerned the constitutionality of a Vermont law that banned using pharmacy records for pharmaceutical marketing purposes.[54] The law targeted detailers, pharmaceutical sales representatives that engage directly with prescribing physicians, who used pharmacy records to strengthen their sales pitches. The state argued that the law banning this practice catered to the substantial governmental interest of preventing the undermining of the "doctor-patient relationship by allowing detailers to influence treatment decisions."[55] However, the Court found this interest unpersuasive and, instead, articulated that "[i]f pharmaceutical marketing affects treatment decisions, it does so because doctors find it persuasive," and "the fear that speech might persuade provides no lawful basis for quieting it."[56] By not recognizing government justifications that resemble paternalistic screening of information and holding that "[s]peech in aid of pharmaceutical marketing . . . is a form of expression protected by the Free Speech Clause," the Supreme Court strengthened the merits of future constitutional challenges to restrictions on off-label pharmaceutical marketing.[57]

C. *Decisions on the Protection Afforded to Off-Label Pharmaceutical Marketing*

A case concerning the constitutionality of the restrictions on off-label pharmaceutical marketing was finally brought before the Court of Appeals for the Second Circuit by a pharmaceutical sales representative, Alfred

---

49 *Id.* at 370-71.

50 *Id.* at 373.

51 *Thompson*, 535 U.S. at 371.

52 *Id.* at 372.

53 Sorrell v. IMS Health Inc., 564 U.S. 552 (2011).

54 *Id.* at 557.

55 *Id.* at 575.

56 *Id.* at 576.

57 *Id.* at 557; CAVALLA, *supra* note 2, at 140.

Caronia, who was found guilty of "conspiracy to introduce a misbranded drug into interstate commerce."[58] During a sting operation, Caronia was recorded informing a physician of off-label uses of the drug Xyrem.[59] While Xyrem has only been approved to treat "narcolepsy patients who experience cataplexy, a condition associated with weak or paralyzed muscles" and "narcolepsy patients with excessive daytime sleepiness," it is also prescribed off-label for "fibromyalgia, depression, schizophrenia, chronic fatigue syndrome and severe cluster headaches."[60] Xyrem's manufacturer, Orphan Medical, accepted a $20 million settlement following the sting operation. However, Caronia sought to be found innocent of any wrongdoing for sharing information about the off-label uses of the drug.[61] After appealing a conviction, the Court of Appeals for the Second Circuit agreed with Caronia and ruled that the restrictions on off-label pharmaceutical marketing are "in violation of his right of free speech under the First Amendment."[62]

Unsurprisingly, the FDA decided to read *Caronia* narrowly, as a fact-bound decision, and continued to threaten prosecution for off-label marketing within the Second Circuit.[63] This continued practice by the FDA led to the follow-up decision of *Amarin Pharma, Inc. v. U.S. FDA*.[64] In *Amarin*, the pharmaceutical company Amarin sought injunctive relief that would prevent the FDA from prosecuting the company for disseminating off-label use information of its drug Vascepa.[65] Specifically, Amarin wished to share that Vascepa has proven successful in "reducing the triglyceride levels of persons with persistently high triglyceride."[66] The statements Amarin desired to share were truthful, having been "derived largely from an FDA-approved study."[67] The district court for the Southern District of New York granted the motion for preliminary injunction and found that the prosecution of off-label promotion was inconsistent with the First Amendment.[68]

Both Judge Chin and Judge Englemeyer, in *Caronia* and *Amarin* respectively, arrived at their decisions concerning the constitutionality of off-label promotion by explicitly relying on *Sorrell*, which again held that

---

58 *See* United States v. Caronia, 703 F.3d 149, 152 (2d Cir. 2012).

59 *Id.* at 155.

60 *Id.*; CAVALLA, *supra* note 2, at 140.

61 CAVALLA, *supra* note 2, at 140.

62 *Caronia*, 703 F.3d at 152.

63 *See* Amarin Pharma, Inc. v. FDA, 119 F.Supp.3d 196, 224-25 (S.D.N.Y. 2015).

64 *Id.*

65 *Id.* at 196.

66 *Id.* at 209.

67 *Id.* at 198.

68 *Id.* at 236-37.

the First Amendment protects pharmaceutical marketing.[69] Furthermore, both opinions included a *Central Hudson* analysis and arrived at the same conclusion.[70] The first two parts of the test are satisfied: (1) off-label drug use concerns lawful activity, and (2) the government does have a substantial interest in preserving the effectiveness of the drug approval process.[71] However, the FDA could not satisfy the third and fourth prongs, which require that (3) the policy directly advance the government's interest, and (4) the government's policy be narrowly drawn to further the interest served.[72]

The courts held that the restrictions on off-label promotion by manufacturers do not directly advance the government's interest in preserving the drug approval process because off-label prescribing is a legal and common practice and other groups lawfully promote off-label.[73] In other words, the drug approval process is not threatened by allowing manufacturers to promote off-label because off-label use information is already being released via third parties and used by prescribing physicians. The courts also held in both cases that the restrictions on off-label marketing are not narrowly drawn to further the interest purported to.[74] In *Caronia* the court offered up several alternatives that the government could employ:

> [I]t could guide physicians and patients in differentiating between misleading and false promotion, exaggerations and embellishments, and truthful or non-misleading information . . . could require pharmaceutical manufacturers to list all applicable or intended indications when they first apply for FDA approval, enabling physicians, the government, and patients to track a drugs development . . . could create other limits, including ceilings or caps on off-label prescriptions . . . and even perhaps further regulate, the liability surrounding off-label promotion and treatment decisions.[75]

Accordingly, the courts in *Caronia* and *Amarin* found that the FDA's construction of the FDCA's misbranding provisions prohibiting manufacturer off-label promotion unconstitutionally restricts free speech.[76]

---

[69] *See* Sorrell v. IMS Health Inc., 564 U.S. 552, 552 (2011); United States v. Caronia, 703 F.3d 149, 162; *Amarin*, 119 F. Supp. 3d at 225-26.

[70] *See* Central Hudson Gas & Electric Corp. v. Pub. Serv. Comm'n, 447 U.S. 557, 566 (1980); *Caronia*, 703 F.3d at 165-68; *Amarin*, 119 F. Supp. 3d at 225-26.

[71] *Caronia*, 703 F.3d at 165; *Amarin*, 119 F. Supp. 3d at 225.

[72] *Caronia*, 703 F.3d at 166-68; *Amarin*, 119 F. Supp. 3d at 225-26.

[73] *Caronia*, 703 F.3d at 166; *Amarin*, 119 F. Supp. 3d at 225-26.

[74] *Caronia*, 703 F.3d at 167; *Amarin*, 119 F. Supp. 3d at 225-26.

[75] *Caronia*, 703 F.3d at 168.

[76] *Caronia*, 703 F.3d at 168; *Amarin*, 119 F. Supp. 3d at 226.

D. *Post-Amarin and Looking to the Future*

Since the seminal 2012 *Caronia* decision, the FDA has failed to adequately address First Amendment concerns regarding off-label promotion, despite promises to do so. FDA leadership has assured that there stands a "commitment at the highest levels of the agency . . . to realign FDA's regulatory posture in this area."[77] Furthermore, on June 6, 2014, the agency formally granted two citizen petitions submitted by the MIWG "in 2011 and 2013, which requested clarification of FDA's position on drug and device manufacturer communications concerning unapproved uses of approved products in light of recent First Amendment case law."[78] Unfortunately, however, the FDA has been slow to provide this clarification. Rather, the agency has avoided addressing the issue by interpreting unfavorable decisions narrowly, by not appealing unfavorable decisions to higher courts, and by entering into settlement agreements.

After Amarin was granted preliminary relief, the case was stayed by judicial order while the company and the FDA explored settlement options.[79] At the same time, Pacira Pharmaceuticals followed the lead of Amarin and filed for injunctive relief in response to an FDA warning letter threatening prosecution for off-label promotion.[80] In October 2015, less than a month after Pacira filed, the FDA withdrew the warning letter and the case was settled by year's end.[81] The Pacira settlement exemplifies how "the government repeatedly has taken aggressive enforcement positions regarding promotion of off-label uses, only to retreat when challenged."[82] In March 2016, Amarin and the FDA reached their own settlement.[83] The parties agreed to be bound by the August 2015 judicial declaration. However, the FDA made sure to announce that the settlement "is specific to this particular case and situation, and does not signify a position on the First Amendment and commercial speech."[84]

---

77 Paul E. Kalb, *Off-Label Communications and the Constitution: Will FDA Finally Change its Policies?*, SIDLEY AUSTIN (Nov. 4, 2014), http://www.pharmacongressportal.com/assets/313/resources/313kalb_2.pdf.

78 David L. Rosen, Jason L. Drori & Melissa Y. Lerner, *How Will FDA Regulate Off-Label Communications in the Post-*Facteau *World?*, FOOD AND DRUG LAW INSTITUTE: FOOD AND DRUG POLICY FORUM (Sept. 26, 2016).

79 Jennifer L. Bragg, Amarin *Settlement Erodes Off-Label Promotion Enforcement*, LAW360 (March 11, 2016), http://www.law360.com/articles/770109/amarin-settlement-erodes-off-label-promotion-enforcement.

80 Rosen, et al., *supra* note 78.

81 *Id.*

82 *Id.*

83 Bragg, *supra* note 79.

84 Eric Palmer, *With FDA Settlement, Tiny* Amarin *Creates Opening for Pharma in Off-Label Marketing*, FIERCEPHARMA (Mar. 9, 2016), http://www.fiercepharma.com/pharma/fda-settlement-tiny-amarin-creates-opening-for-pharma-off-label-marketing.

Other cases concerning off-label promotion were also argued in court in 2016 and were decided using the rationales of *Caronia* and *Amarin*.[85] These persisting First Amendment victories sparked a great deal of attention to the FDA's stalling tactics in addressing its off-label speech policies. In May 2016, Rep. Fred Upton, Chairman of the House Committee on Energy and Commerce, wrote a letter to Department of Health and Human Services Secretary Sylvia Burwell expressing his "perplex[ity] by the agency's unwillingness or inability to publicly clarify its current thinking on these issues in a coherent manner."[86] Upton continued that, "[i]f FDA continues to remain silent . . . settlement agreements will be the only means by which policy is formulated–and it will be in an ad hoc manner lacking any semblance of consistency and cohesiveness."[87] Approximately three months later, it was announced in the Federal Registrar that the FDA would hold a two-day meeting on off-label communications beginning on November 9, 2016.[88]

After the November meeting and after receiving written comments, the FDA released a sixty page memorandum in January 2017, which "provides additional background on the issues FDA is considering as part of its comprehensive review" of off-label promotion and First Amendment concerns.[89] To the frustration of many, the memorandum plainly seeks to reaffirm and bolster the FDA's interest to regulate off-label promotion. The FDA pinpoints numerous instances of off-label promotion causing patient harm and explains why the alternatives offered in *Caronia* are unsatisfactory. The memorandum makes clear that the FDA makes no concessions on its constitutional authority despite the case law.

However, the FDA also released in January, draft guidance that could be an important first step in easing restrictions on off-label promotion.[90] The guidance allows manufacturers to share certain off-label information, specifically health care economic information (HCEI) and information about investigational products, with payors and other similar entities.[91] Payors' coverage determinations have significant influence over

---

85 Rosen, et al., *supra* note 78.

86 *Id.*

87 *Id.*

88 Manufacturer Communications Regarding Unapproved Uses of Approved or Cleared Medical Products, 81 Fed. Reg. 60,299 (Sept. 1, 2016) (to be codified at 21 C.F.R. pt. 15).

89 Manufacturer Communications Regarding Unapproved Uses of Approved or Cleared Medical Products; Availability of Memorandum; Reopening of the Comment Period, 82 Fed. Red. 12,6367 (Jan. 19, 2017) [hereinafter Manufacturer Communications]. *See also* Memorandum from the Food & Drug Admin. on Public Health Interests and First Amendment Considerations Related to Manufacturer Communications Regarding Unapproved Uses of Approved or Cleared Medical Products (Jan. 6, 2017).

90 U.S. FOOD & DRUG ADMIN, DRUG AND DEVICE MANUFACTURER COMMUNICATIONS WITH PAYORS, FORMULARY COMMITTEES, AND SIMILAR ENTITIES – QUESTIONS AND ANSWERS: GUIDANCE FOR INDUSTRY (2017).

91 *Id.*

physicians' prescribing decisions, and payors presumably have more time and resources to evaluate the merits of off-label information. Public comments on the memorandum and draft guidance were due on April 19, 2017.[92]

Nonetheless, despite these developments, the fate of the constitutionality of off-label promotion remains uncertain. And while the FDA, the pharmaceutical industry, and stakeholders continue to haggle over what communications are and are not permissible, medical advances continue. The free flow of information is increasingly recognized as critical in leveraging these advances for the benefit of patient welfare. Therefore, the FDA's ability to defend its paternalistic position will decline as time marches on.

II. ANALYSIS – PATIENT WELFARE ARGUMENTS

To the average patient, the thought of being prescribed a drug for a use which it has not been approved may seem concerning. The FDA's approval of a new drug has evolved, "into the international 'gold standard' on product safety and effectiveness."[93] However, upon deeper examination, the off-label use of drugs can be viewed as innovative and sensible, just as easily as ill-supported and dangerous.

The FDA asserts that if the restrictions on off-label promotion were lifted, the integrity of the drug approval process would be compromised and the risk of patient exposure to unsafe and ineffective drugs would increase.[94] In other words, if the restrictions were not in place, manufacturers would be deterred from filing a SNDA, and thus, deterred from proving the safety and effectiveness of the supplemental use. However, in this assertion, the FDA fails to acknowledge that the drug approval process, as it stands today, is not set up to accommodate the efficient approval of all off-label uses, even those with life-saving capabilities.

Most Americans are inclined to trust their government and the agencies that the government defers to on matters of public welfare. Furthermore, "the FDA has consistently been named or identified as one of the most popular and well-respected agencies in government."[95] However, it is important to recognize that the FDA can succumb to reputational concerns, politics, and a lack of resources, all of which can impact drug approval decisions

---

92 Manufacturer Communications, *supra* note 89.

93 James T. O'Reilly, *Losing Deference in the FDA's Second Century: Judicial Review, Politics, And A Diminished Legacy of Expertise*, 93 CORNELL L. REV. 939, 949 (2008).

94 United States v. Caronia, 703 F.3d 149, 166 (2nd Cir. 2012).

95 CARPENTER, *supra* note 23, at 12.

In the early 2000s, there was an "unprecedented surge in industry lobbying at senior levels of Health and Human Services and the White House."[96] It did not go unnoticed that this increase in outside pressure was impacting decision-making and policies within the FDA. On September 2005, the New England Journal of Medicine published an article acknowledging that "recent actions of the FDA leadership have made a mockery of the process of evaluating scientific evidence . . . squandered the public trust, and tarnished the agency's image."[97] Also during that time, dozens of high-ranking officials from the FDA came forward as whistleblowers and echoed the sentiments of FDA scientist Dr. David Graham, who "testified before Congress that FDA drug safety managers felt pressure to approve certain drugs, despite their substantial risks."[98] These reports, coupled with recalls of FDA-approved drugs, such as Merck's Vioxx in 2004, led many to believe that the FDA enjoys "a too-cozy relationship with the pharmaceutical industry."[99] Likely in efforts to combat the "too-cozy" perception, drug approvals in 2005 dropped by approximately 45%.[100] The lower approval rates continued for the next several years and have been referenced to as the "painful trough of 2005-2010."[101]

In additional to reputational concerns, politics also has the ability to upset the efficiency of the drug approval process. One easy-to-identify example of this is the 2013 government shut down. Due to Congress's inability to reach a consensus on the budget, the government was shut down for sixteen days beginning on October 1, 2013. The shutdown reduced staff and "cut[] off funding for health services, research laboratories, and other efforts that depend on federal support."[102] Many research studies were delayed and even cancelled.[103] Studies forced to restart required significantly more time and money.[104] Federal laboratories had to euthanize animals, and it is speculated that "specialized and extremely costly mice [used] in research on such disorders as Alzheimer's disease and cancer . . . die[d] in the absence of constant monitoring by scientists who were furloughed."[105] It is therefore not surprising that "the shutdown likely did

---

96 O'Reilly, *supra* note 93, at 962.

97 *Id.* at 963.

98 *Id.* at 964.

99 *Id.* at 963.

100 U.S. FOOD & DRUG ADMIN. CDER, NOVEL NEW DRUGS 2013 SUMMARY 3 (2014), http://www.fda.gov/downloads/drugs/developmentapproval process/druginnovation/ucm381803.pdf.

101 Bernard Munos, *2014 New Drug Approvals Hit 18-Year High*, FORBES (Jan. 2, 2014), http://www.forbes.com/sites/bernardmunos/2015/01/02/the-fda-approvals-of-2014/#30fa214452ac.

102 Bridget M. Kuehn, *Shutdown Underscored Vulnerability of US Public Health and Biomedical Research to Political Wrangling*, 310 J. AM. MED. ASS'N. 1907, 1907 (2013).

103 *Id.*

104 *Id.* at 1908.

105 *Id.* at 1909.

lasting damage to the US public health and biomedical research enterprise, setting back research and public health programs by years."[106] That the threat of a shutdown occurs nearly every year underscores the vulnerability of the drug approval process.

One of the greatest inefficiencies in the drug approval process is a lack of resources at the FDA. Limited staff and resources to review applications can cause significant delays. As science advances at a rapid pace and cutting-edge therapeutics are studied, the agency moves slowly and cautiously as they lack the internal expertise to evaluate submissions. Dr. Margaret Hamburg, who served as FDA commissioner from May 2009 to April 2015, explained the problem as: "[t]oday, FDA is relying on 20th century regulatory science to evaluate 21st century medical products. Regulatory science is needed to provide better tools, standards, and pathways to evaluate products under development."[107] The FDA's struggle to keep pace with science greatly hinders the accessibility of modern-day medicine.

Because of reputational concerns, national politics, and limited resources at the FDA, new drug development and approval, on average, costs manufacturers ten to fifteen years and more than $1.3 billion.[108] Furthermore, the FDA approval process often "requires companies to seek approval for a narrow initial indication."[109] Given the amount of resources it takes manufacturers to gain initial FDA approval, manufacturers are reluctant, and often times unable, to exert additional "time, labor, expenses, and resources needed to conduct clinical trials" for supplemental use applications.[110] For example, while an off-label use may have life-saving value, a small patient population may not justify the filing of an application.[111]

Despite the very real obstacles for manufacturers in getting secondary uses approved, physicians are very aware that secondary uses exist and are necessary for providing the best and most personalized care they can. In a study by George Mason University that included five hundred physicians, doctors overwhelmingly supported off-label medicine.[112] This is the case

---

106 *Id.* at 1907.

107 CENTER FOR HEALTH TRANSFORMATION, CREATING A 21ST CENTURY FOOD & DRUG ADMINISTRATION 3 (2012).

108 Andrew E. Podgomy, *Supporting the Rationale Behind the Hatch-Waxman Act and Patent Law: How Reverse Payment Settlements Under* FTC v. Actavis*, Inc. can be Procompetitive*, 12 IND. HEALTH L. REV. 423, 448 (2015).

109 Santosh V. Coutinho, *License to Promote, or Just What the Doctor Ordered? The New FDA Guidance on Dissemination of Off-Label Reprints by Pharmaceutical Companies,* 28 TEMP. J. SCI. TECH. & ENVTL. L. 279, 301-02 (2009).

110 *Id.* at 302.

111 Craft, *supra* note 26, at 105.

112 Daniel B. Klein & Alexander Tabarrok, *Who Certifies Off-Label?*, HEARTLAND (Summer 2004), https://www.heartland.org/_template-assets/documents/publications/15757.pdf.

despite the FDA's "reminder to physicians that they may face greater tort liability when they prescribe a drug for any indication other than those expressly approved in the new drug application process."[113]

In *Caronia*, the court recognized physicians' interest in having access to information pertaining to off-label medicine, stating that "information can save lives."[114] The following subsections elaborate upon why the healthcare community needs greater access to off-label use information, especially as technological and medical capabilities continue to advance. These arguments further the Second Circuit's position that the FDA's interest in preserving the effectiveness of the drug approval process is not served by the current restrictions on off-label promotion.

A. *The Promotion of Disparities in Public Health*

Barriers to information are most detrimental to vulnerable populations. This is primarily because more educated and wealthier populations are better able to maneuver around such barriers and as a result make more informed decisions. Justice Blackmun in *Virginia Citizens* articulated this notion, in response to a law that forbade the advertisements of prescription drug prices:

> Those whom the suppression of prescription drug price information hits the hardest are the poor, the sick, and particularly the aged . . . Advertising, however tasteless and excessive it sometimes may seem, is nonetheless dissemination of information as to who is producing and selling what product, for what reason, and at what price. So long as we preserve a predominantly free enterprise economy, the allocation of our resources in large measure will be made through numerous private economic decisions. It is a matter of public interest that those decisions, in the aggregate, be intelligent and well informed. To this end, the free flow of commercial information is indispensable.[115]

This policy is extremely applicable to the FDA's current restrictions on off-label promotion. Because pharmaceutical companies and manufacturers are restricted from sharing off-label use information, the information is not proportionately disseminated. Physicians in more prestigious hospitals with bigger research facilities and better ties to the larger scientific community will have greater access to off-label use information than physicians practicing in less connected areas. Therefore, patients that have access to the more prestigious physicians will potentially receive more innovative and informed treatments.

---

113 CARPENTER, *supra* note 23, at 618.

114 United States v. Caronia, 703 F.3d 149, 167 (2nd Cir. 2012).

115 Virginia State Bd. of Pharm. v. Virginia Citizens Consumer Council, Inc., 425 U.S. 748, 765 (1976).

This rationale may be supported by one study done on trends of off-label prescribing to children.[116] The study revealed that specialists were significantly more likely to prescribe off-label than general pediatricians (68% as opposed to 59% for general pediatricians).[117] Because a specialist holds a higher level of expertise and prestige in a particular area of medicine it is not surprising that a specialist would have greater access to off-label use information relevant to that specialty, and in turn feel more comfortable prescribing off-label. Furthermore, individuals with a higher socioeconomic status are more likely to visit a specialist.[118] It is suggested that members in the middle-income and high-income classes are 30% more likely to visit a specialist.[119] This is the case even while "low income, minority status, and lesser educational attainment are all associated with a higher prevalence of chronic conditions and poorer health, and, therefore, are presumably linked with a greater clinical need for specialist services."[120] Because those with a higher socioeconomic status are more likely to frequent a specialist, it follows that they are also more likely to receive an off-label prescription.

One explanation for the disparity in specialist visits is that "people with higher educational attainment may be more adept in obtaining preauthorization for desired services or seeking referrals to specialists."[121] While lifting restrictions on off-label promotion is unlikely to have any effect on this paradigm, it may allow a general physician to provide more comparable care to that of a specialist. The free flow of off-label use information would allow for general physicians to have easier access to information that many specialists already have and use. In other words, off-label promotion may close a knowledge gap between specialists and general practitioners.

B. *Billions Taken Away from New Drug Development*

Another problem with the restrictions on off-label promotion is that nearly all manufacturers continue to engage in off-label marketing regardless. This may be because manufacturers are unsure on the parameters of the restrictions on off-label promotion, provided that FDA's authority "to regulate truthful, non-misleading manufacturer speech [is] not

---

[116] Alicia T.F. Bazzano et al., *Off-Label Prescribing to Children in the United States Outpatient Setting*, 9 ACAD. PEDIATRICS 81 (2009).

[117] *Id.* at 86.

[118] Jeannette Rogowski et al., *Socioeconomic Disparities in Medical Provider Visits among Medicare Managed Care Enrollees*, 45 INQUIRY 112, 121 (2008).

[119] *Id.*

[120] Jan Blustein & Linda J. Weiss, *Visits to Specialists under Medicare: Socioeconomic Advantage and Access to Care*, 9 J. HEALTH CARE FOR POOR & UNDERSERVED 153, 165 (1998).

[121] Rogowski, *supra* note 118, at 114.

from a clear and precise statement of law, but from a labyrinth of statutes and regulations, as well as non-binding guidance documents."[122] The pharmaceutical industry complains that "regulatory compliance is often reduced to guesswork."[123] Manufacturers may also opt to continue marketing off-label because of the potential for enormous profits. While it is likely due to a combination of unclear FDA guidance and monetary incentives, off-label promotion persists even at the risk of billion dollar fines.

The weight and significance of fines on off-label marketing can be made clear through looking at the experiences of major pharmaceutical manufacturer Pfizer. Since 2004, Pfizer has paid over 3.2 billion dollars in fines for off-label marketing.[124] In 2009, Pfizer accepted one settlement offer for 2.3 billion dollars.[125] However, these numbers are offset by the profits that stand to be made. From 2001 to 2008, Pfizer made $16.8 billion from the very drugs it was fined for.[126] It therefore follows that:

> While these fines may seem staggering in isolation, their deterrent effect pales in comparison with the huge financial profits stemming from off-label marketing. As an amoral economic actor, pharma could conclude that such fines are a 'cost of doing business' and not substantially change their illegal business practices.[127]

Marketing teams have reported that, "if we don't get at least one warning letter a year from the FDA we aren't really doing our job."[128]

The continuous amount of energy and money being spent in monitoring, litigation, settlement negotiations, and fines for off-label promotion is highly concerning given that such resources could be spent on new drug development. The billions of dollars being spent every year debating the legality of a practice that actually assists physicians in making informed treatment decisions is preventing the discovery and approval of life-saving cures.

Furthermore, because of the profitable nature of off-label marketing, many drug manufacturers attempt to avoid fines by engaging in inconspicuous, and often times ethically questionable, tactics of off-label promotion. One way in which manufacturers attempt to secure plausible

---

[122] Brief for Medical Information Working Group, as Amici Curiae Supporting Petitioners, at 14, Pacira v. FDA, No. 1:15-cv-07055 (S.D.N.Y. 2015).

[123] *Id.* at 17.

[124] CAVALLA, *supra* note 2, at 131.

[125] *Id.*

[126] Fazal Khan & Justin Holloway, *Verify, then Trust: How to Legalize Off-Label Drug Marketing*, 117 PA. ST. L. REV. 407, 410 (2012).

[127] *Id.* at 410-11.

[128] John Osborn, *Feds Have Beaten Pharma Into Submission Over Off-Label Drug Use, But At What Cost?*, FORBES (November 7, 2013), http://www.forbes.com/sites/johnosborn/2013/11/07/ho-hum-another-multi-billion-dollar-drug-company-fraud-settlement/#23447c653bb5.

deniability is through ghost writing. Ghost writing is the "process by which a pharmaceutical company contracts with or hires a medical education and communications company [] to draft articles about new uses for FDA-approved drugs or medical devices."[129] This type of promotion does not hold the manufacturer accountable if the information is false or misleading.[130] Another tactic of questionable character that is employed by manufacturers is the wooing of doctors through lavish gifts and monetary incentives.[131] It is highly probable that costly and ethically questionable marketing strategies would decrease if manufacturers were able to freely disperse off-label information.

C. *The Rise of Personalized Medicine & Developing Scientific Capabilities*

In the day of rising personalized medicine, where knowledge in molecular genetics and cell biology is accelerating, the promotion of off-label uses is integral. Regulatory approaches need to, "fully adapt to a different paradigm where treatment is highly specific to individual patients."[132]

Personalized medicine is greatly at odds with the restrictions on off-label pharmaceutical marketing. For example, under the current regulatory scheme, if all approved drugs for a particular use were proven to be ineffective in treating a particular patient, a physician would be hindered from learning about viable off-label options. The drugs that have been able to gain FDA approval will not satisfy the needs of all Americans, given each individual's unique genetic makeup. In keeping pace with science's growing understanding of genetic disparity, FDA policies should take account of the fact that,"[n]either Pfizer nor Washington can ever stuff health itself into a one-price, uniform, one-America box."[133]

Today, "genomic analysis to help guide personalized treatment for cancer" specifically, is generating significant attention.[134] Such analysis could be the future of cancer treatment, being that "targeting drivers of uncontrolled growth present in a given tumor in a timely manner will be highly effective with reduced side effects."[135] However, Richard Schilsky, who was appointed Chief Medical Officer of the American Society of

---

129 Khan & Holloway, *supra* note 126, at 418.

130 *Id.*

131 *Id.* at 421.

132 KEWAL K. JAIN, TEXTBOOK OF PERSONALIZED MEDICINE 353 (Springer 2009).

133 PETER W. HUBER, THE CURE IN THE CODE: HOW 20TH CENTURY LAW IS UNDERMINING 21ST CENTURY MEDICINE 226 (Basic Books 2013).

134 MICHAEL SNYDER, GENOMICS AND PERSONALIZED MEDICINE: WHAT EVERYONE NEEDS TO KNOW 36 (Oxford University Press 2016).

135 *Id.*

Clinical Oncology ("ASCO") in 2013, has articulated that "[o]ne of the major challenges to implementing personalized medicine is the lack of information about the risks and benefits of targeted drugs that are used off label to treat patients whose tumor harbors a genomic abnormality."[136]

Following a genomic analysis that reveals a driver mutation not typical of a given type of cancer, a physician may identify a commercially available drug that targets the driver but is only approved to treat another form of cancer. That physician could then treat the cancer through off-label medicine. Unfortunately, however, the restrictions on off-label promotion make it difficult for physicians to identify and assess off-label uses. Therefore, it is critical that information on off-label uses be available to take advantage of our growing understanding of genetics and cellular biology.

Furthermore, just as "Americans are biochemically diverse," the diseases that we face are also "biochemically complex."[137] Scientists and physicians having greater access to information concerning drugs already on the market will only help them in understanding and combatting the advanced diseases we see today. The position that off-label use information should be readily accessible is furthered in that, "[a]lmost all of the new information that is gleaned from the clinical trials of a fundamentally new drug is needed to ascertain, directly or indirectly, how human bodies operate at the molecular level."[138] The availability and compilation of such information, coupled with decoding technology, has the potential to "expose the architectures and dynamics of countless molecular chain reactions and networks that make human bodies function well or badly."[139]

It is estimated that "the power of the technology now engaged in decoding life is doubling every year, if not faster."[140] However, advancing technology and science can only be taken full advantage of with the free flow of information. Having unfettered access to off-label use information has the potential to not only save the lives of individuals who do not respond to drugs already on the market, but also to assist in finding cures for the most challenging diseases and viruses in the world.

---

[136] Margot J. Fromer, *Friends of Cancer Research Holds Annual Conference on Clinical Cancer Research*, ASCO POST (Feb. 10, 2015), http://www.ascopost.com/issues/february-10-2015/friends-of-cancer-research-holds-annual-conference-on-clinical-cancer-research/.

[137] HUBER, *supra* note 133, at 233.

[138] *Id.* at xiv.

[139] *Id.*

[140] *Id.* at xv.

D. *Scaring Away Investment*

The FDA's interest to preserve the integrity of the drug approval process through stringent restrictions on off-label marketing is severely eroded by the very tactics being used to prosecute off-label marketing. In fact, it appears that the questionable tactics being employed by the FDA are actually undermining the drug approval process by discouraging investment and participation in the drug discovery industry.

First, it is important to acknowledge the amount of power that the FDA holds. The United States pharmaceutical industry is by far the world's largest; "the American market accounted for $216 billion in spending on prescription drugs in 2006."[141] Given the profits that stand to be made, entry into the U.S. market is what manufacturers aspire to, and "the FDA's veto power over entry into the American health-care system translates into global economic and scientific reach."[142] The FDA's power often has the effect of "secur[ing] 'voluntary' compliance with whatever the agency demands."[143] This is most likely why manufacturers have been so quick to accept settlement offers for instances of off-label promotion. Furthermore, the likelihood of acceptance is increased when a manufacturer has a new drug pending approval. It has been poetically put that "[t]he FDA is standing there with a machine gun against the pharmaceutical industry, so you better be their friend rather than their enemy. They are the boss. If you're a pharmaceutical firm, they own your body and soul."[144]

This magnitude of control has led to what many would consider an abuse of power when it comes to the agency's rulemaking process. Agency rules and guidance are overly broad and pharmaceutical companies and manufacturers find it difficult to know whether they are doing something illegal.

In *Amarin*, Amarin moved primarily under the First Amendment, but alternatively, under the due process clause on the ground that the FDA's regulations as to misbranding were vague and did not "fairly notify Amarin of what off-label promotion is permitted and what is forbidden."[145] In response to the complaint, the FDA sought a resolution and set out its position in what has been called the "Woodcock Letter."[146] Dr. Janet Woodcock, director of the FDA's CDER, wrote the letter on June 5, 2015.[147] The letter informed Amarin that "some [of the proposed

---

141 CARPENTER, *supra* note 23, at 1.

142 *Id.*

143 Lars Noah, *Governance by the Backdoor: Administrative Law(lessness?) at the FDA*, 93 NEB. L. REV. 89, 90 (2014).

144 CARPENTER, *supra* note 23, at 7.

145 Amarin Pharma, Inc. v. FDA, 119 F.Supp.3d 196, 215-16 (S.D.N.Y. 2015).

146 *Id.*

147 *Id.*

communications] 'fall within the scope' of existing FDA guidance allowing manufacturers to disseminate to doctors 'truthful and non-misleading scientific or medical publications on unapproved new uses.'"[148] The Woodcock Letter also proposed defined conditions under which Amarin could communicate some of the information in question to doctors without the threat of prosecution.[149] The Woodcock Letter is a prime example of FDA's case-by-case decision tactics and broad discretion power.

Such tactics have been suggested to have long-term consequences when it comes to agency credibility. Furthermore, such tactics have also been suggested to have implications in the marketplace:

> For too long now, the FDA has not behaved in a consistent manner, in accordance with its own rules and guidelines. And we are already seeing the effects: inventors, entrepreneurs, and venture capital investors have begun to accept this as the new norm . . . . Because of that, their investment of time, energy, and money in companies developing truly medically innovative products has diminished and they have shifted focus toward opportunities that do not rely on FDA approval.[150]

One study that surveyed 150 venture capital firms has confirmed as much, identifying "[FDA] regulatory challenges as being the most significant factor driving away investment from startup companies."[151] The study further indicates that venture capitalist are re-focusing their sights and looking towards Europe and Asia.[152]

The FDA, in driving away capital "from lifesaving and life-sustaining products and into areas less regulated by the FDA as well as into other countries,"[153] is severely undermining the effectiveness and integrity of the drug approval process. The FDA approval process is in place to bring as many safe and effective healthcare solutions to market as possible. However, with less capital, medical innovations will begin to decrease and patient welfare will not progress at the rate at which it could.

CONCLUSION

The FDA drug approval process is unable to accommodate the efficient approval of all safe and effective uses of drugs. Furthermore, the

---

[148] *Id.*

[149] *Id.*

[150] JOSEPH V. GULFO, INNOVATION BREAKDOWN: HOW THE FDA AND WALL STREET CRIPPLE MEDICAL ADVANCES 232 (Post Hill 2014).

[151] *Report Confirms FDA is Scaring off Venture Capital Investments in Medtech*, ORTHOSTREAMS, http://orthostreams.com/2011/10/report-confirms-fda-is-scaring-off-venture-capital-investments-in-medtech/.

[152] *Id.*

[153] *Id.*

FDA's policies that restrict off-label promotion are keeping scientists and physicians alike from doing their jobs to the best of their ability. FDA's prosecution of off-label marketing and the tactics it employs in doing so are driving billions of dollars away from new drug discovery through costs associated with litigation and by deterring capital investment in the healthcare industry. The restrictions on off-label speech are severely limiting the treatment options available to patients, particularly our most vulnerable, impoverished populations. Furthermore, these arguments in favor of the free flow of information will only become more persuasive as our scientific and technological capabilities increase.

The FDA's stringent restrictions on off-label promotion are not only inconsistent with patient welfare but also with case law that has expanded the First Amendment to protect commercial speech. These blatant inconsistencies show that the government is fighting a losing battle in its attempt to defend the restrictions on off-label promotion. The FDA should loosen the reins over off-label speech and spend resources on developing more appropriate policies and practices that protect both the manufacturers' right to inform and the patients' right to protection.

REDEFINING LINCOLN'S LAW: HOW TO SHAPE THE THEORY OF IMPLIED CERTIFICATIONS POST-*ESCOBAR*

Doan Phan[*]

INTRODUCTION

The False Claims Act ("FCA") creates a civil cause of action for the Attorney General or a *qui tam*[1] relator against people who commit certain fraudulent acts against the U.S. Federal Government.[2] In recent years, the FCA has been used to "deter and punish government-contracting fraud across a number of industries, including defense, health care, for-profit higher education, and mortgage lending and financial services."[3] What contributed to the popular use of the FCA is its imposition of civil liability on a person who knowingly submits a "false or fraudulent claim for payment or approval" to the government.[4] Initially, the courts applied this provision of the Act to claims that contain an *express* false certification of compliance.[5] However, in 1994, the U.S. Court of Federal Claims became the first court to recognize the theory of *implied* false certification.[6] The theory of implied false certification is similar to the theory of express false certification, except that a defendant has not signed an express certification of compliance with a statute, regulation, or contract clause.[7] Rather, the theory of implied certification provides that, by submitting a claim, the contractor implicitly certifies that he has complied with all applicable statutes, regulations, and contract clauses.[8] Thus, according to the implied

---

[*] J.D. Candidate 2017, George Mason University School of Law.

[1] *Qui Tam* under the False Claims Act allows a private individual to bring a suit on the government's behalf. The private plaintiff in the case is referred to as a "relator." 31 U.S.C. §§ 3730(b)(1), (c)(3), (d) (2010). *See also* United States *ex rel.* Batty v. Amerigroup Ill., Inc., 528 F. Supp. 2d 861, 871 (N.D. Ill. 2007).

[2] *See* 31 U.S.C. §§ 3729-3733 (2009).

[3] Christopher L. Martin, Jr., *Reining in Lincoln's Law: A Call to Limit the Implied Certification Theory of Liability Under the False Claims Act*, 101 CAL. L. REV. 227, 229 (2013).

[4] *Id.* at 230.

[5] United States *ex rel.* Wilkins v. United Health Grp, Inc., 659 F.3d 295, 305 (3d Cir. 2011) (explaining that the "express false certification theory" of liability applies when an entity is liable under the Act "for falsely certifying that it is in compliance with regulations which are prerequisites to Government payment" concerning payment of federal funds).

[6] *See* Ab-Tech Constr., Inc. v. United States, 31 Fed. Cl. 429, 434 (1994) ("The payment vouchers represented an implied certification by Ab-Tech of its continuing adherence to the requirements for participation in the 8(a) program.").

[7] United States *ex. rel*. Mikes v. Straus, 274 F.3d 687, 699 (2d Cir. 2001).

[8] *Id.*

certification theory, if the contractor submits a claim knowing that he has not complied with all laws, regulations, and contract clauses, the contractor has submitted, by implication, a false claim.[9] The theory of implied false certification broadens the scope of the FCA, which can lead to unfair results, especially given the heavy penalties that often face contractors who run afoul of the FCA.[10] The results remained confusing due to the circuit split on how to treat implied false certification.[11] The courts' treatment of implied false certification have varied from overly broad to completely doing away with the theory.[12] However, in a recent case, *Universal Health Services v. United States ex rel. Escobar,* the United States Supreme Court laid the circuit split to rest while also redefining the theory of implied false certification.[13]

This comment maintains that while the heightened materiality standard under *Escobar* is an effective means of cabining the reach of the theory of implied false certification, the Supreme Court's lack of guidelines in how to find materiality has the potential of leading to a circuit split on what claims would be material. As a way to help redefine with specificity what materiality is, this comment looks to the outcome determinative test under the Eighth Circuit in *Costner v. URS Consultants, Inc.*[14] However, since the test was formulated before the Supreme Court's decision in *Escobar* by two decades, this comment will look to modify the outcome determinative test. In so doing, the new test that this comment advocates is one that focuses on economic detriment to the government and places the burden on the plaintiff to show that the government would have acted differently if it had known of the omission. However, to comply with *Escobar*, the test will be fact-intensive, balancing the multiple factors while placing the most weight on the government's past payment practices. In so doing, the test aims to ensure uniformity by creating a starting point for all lower court analyses.

Part I of this comment will document the history of the FCA, as well as outline its provisions. Part II will discuss the emergence of the implied theory of certification and how it diverged from the primary express theory of certification. Part III will discuss the ensuing circuit splits as they try to grapple with whether or not implied false certification is even a liability under the FCA. Part IV will detail the Supreme Court's ruling in the *Escobar* case, detailing the two major holdings, as well as its impact on future litigation. Finally, Part V will explore the benefits of the heightened

---

[9] *Id.*

[10] 31 U.S.C. § 3729(a)(1)-(3) (2009) (Damages include treble damages as well any other costs of civil actions).

[11] *See infra* Part III.

[12] *See infra* Part III.

[13] Universal Health Servs. v. U.S. *ex. rel.* Escobar, 136 S. Ct. 1989 (2016).

[14] Costner v. Urs Consultants, 153 F.3d 667, 677 (8th Cir. 1998). *See also* Megan L. Hoffman, *The Substantial Weight Test: A Proposal to Resolve the Circuits' Disparate Interpretations of Materiality Under the False Claims Act*, 58 KAN. L. REV. 181, 198-99 (2009).

materiality standard as written by the Supreme Court in *Escobar* and then improvements to that standard to ensure uniformity in the interpretation of what is material.

I. BACKGROUND: THE HISTORY AND PROVISIONS OF THE FALSE CLAIMS ACT

A. *A Brief History of the False Claims Act*

The FCA, also known as the "Lincoln Law," was enacted in 1863 amid concerns that a contractor supplying goods to the Union Army was defrauding them.[15] The manager of the bill in the Senate echoed these concerns as he stated the objective of the FCA: to punish and prevent fraud.[16] Originally, the FCA provided double damages for the government against violators of the statute, plus a penalty of $2,000 for each false claim.[17] There were several later amendments to the FCA[18], with the more significant changes happening in 1986. The 1986 Amendments were passed in response to an increase in fraud from government contractors overcharging the government and the government's inability to effectively retrieve those funds.[19] The 1986 Amendments allowed for an expansion of the role of relators and reduced the difficulties in bringing FCA claims.[20] Additionally, and most importantly to those who have FCA claims brought against them, the double damages were increased to triple damages[21] and the penalties were raised from $2,000 to a range of $5,000 to $10,000.[22]

---

15 U.S. DEP'T OF JUSTICE, THE FALSE CLAIMS ACT: A PRIMER, http://www.justice.gov/sites/default/files/civil/legacy/2011/04/22/C-FRAUDS_FCA_Primer.pdf.

16 CONG. GLOBE, 37th Cong., 3rd Sess. 952 (1863) ("The country . . . has been full of complaints respecting the frauds and corruptions practices in obtaining pay from the Government during the present war . . . From the attention I have been able to give the subject, I am satisfied that more stringent provisions are required for the purpose of punishing and preventing these frauds; and with a view to apply a more speedy and vigorous remedy in cases of this kind the present bill has been prepared.").

17 U.S. DEP'T OF JUSTICE, *supra* note 15.

18 The civil version of the FCA is codified at 31 U.S.C §§ 3720-3722 (2009), while the criminal version of the FCA is codified at 18 U.S.C. § 287 (2015).

19 *Overview of False Claims and Fraud Legislation: Hearing on Legislation to Combat the Growth of Fraud Against the Federal Government Through the Filing of False Claims by Government Contractors Before the S. Comm. On the Judiciary,* 99th Cong. 13-14 (1986) (statements of Senator Orrin G. Hatch and Senator Charles E. Grassley).

20 False Claims Act, Pub. L. No. 99-562, 100 Stat. 3153.

21 Treble damages, as defined by Black's Law Dictionary, are "[d]amages that, by statute, are three times the amount of actual damages that the fact-finder determines is owed." BLACK'S LAW DICTIONARY (10th ed. 2014). Treble damages, along with double damages and other specified 'civil penalties,' are used to provide liquidated damages for "actual losses that cannot be proved or that are otherwise unrecognized by the law." DAN B. DOBBS, LAW OF REMEDIES 359 (2d ed. 1993).

22 U.S. DEP'T OF JUSTICE, *supra* note 15.

B. *Provisions of the False Claims Act*

The main purpose of the FCA is to impose money damages on a person who conducts himself fraudulently when dealing with the United States Government.[23] There are five sections to the statute: (I) section 3729 lists the liability for seven specific acts as well as possible ways for the contractor to reduce their damages for violating the Act; (II) section 3730 authorizes either the Attorney General to bring a civil action under the Act or a *qui tam* relator to bring a suit under the Act on behalf of the government; (III) section 3731 sets forth procedural requirements to bring a claim under the Act, including both the statute of limitations and the standard for deciding the claim; (IV) section 3732 establishes the federal jurisdiction; and (V) section 3733 establishes the procedures for civil investigative demands by the government.[24]

Section 3729(a)(1)(A)-(G) lists the seven specific acts that can give rise to a claim under the Act.[25] Section 3729(a)(1)(A) is central to the issue debated in many cases brought under the Act.[26] The importance of subsection (a)(1)(A) is that it establishes a knowledge standard, imposing liability on a person who "knowingly presents, or causes to be presented, a false or fraudulent claim for payment or approval."[27] The Act defines the terms "knowing" and "knowingly" as meaning a person who: "(i) has actual knowledge of the information; (ii) acts in deliberate ignorance of the truth or falsity of the information; or (iii) acts in reckless disregard of the truth or falsity of the information."[28] Furthermore, the Act specifically states that it "require[s] no proof of specific intent to defraud."[29] Because this subsection requires a pleading based on a false claim, the party bringing the FCA claim must plead each element in accordance with Rule 9(b) of the Federal Rules of Civil Procedure.[30]

---

[23] 31 U.S.C. §§ 3729-3733 (2009). *See also* Martin, *supra* note 3, at 233.

[24] *Id.*

[25] 31 U.S.C. §3729(a)(1)(A)-(G)(2009).

[26] For example, it was a central issue in *United States v. Sanford-Brown, Ltd.*, 788 F.3d 696 (7th Cir. 2015). Additionally, it was noted as being the most litigated Act by Christopher Martin Jr. in his publication. Martin, *supra* note 3, 233-34.

[27] 31 U.S.C. § 3729(a)(1)(A)(2009).

[28] 31 U.S.C. § 3729(b)(1)(A)(i)-(iii)(2009).

[29] 31 U.S.C. § 3729(b)(1)(B)(2009).

[30] *See* Ebeid *ex rel.* United States v. Lungwitz, 616 F.3d 993, 1001 (9th Cir. 2010) (holding that because Ebeid failed to plead with the particularity required of Rule 9(b) of the FRCP, despite articulating the framework of a claim for implied false certification, the case will be dismissed); United States *ex rel.* DeKort v. Integrated Coast Guard Sys., 705 F. Supp. 2d 519, 544 (N.D. Tex. 2010) (explaining that even if the relator cannot allege the details of an actually submitted false claim, it can survive when the scheme to submit false claims is paired with a reliable indication that such claims were submitted) (quoting United States *ex rel.* Grubbs v. Kanneganti, 565 F.3d 180, 190 (5th Cir. 2009)).

The Act also provides monetary incentives for *qui tam* relators, allowing them to "receive at least 15 percent but no more than 25 percent of the proceeds of the action or settlement of the claim."[31] While this award depends on how much the person contributed to the case, a *qui tam* relator could still receive up to 10 percent of the action or settlement of the claim even if the FCA claim was based on "disclosures of specific information (other than information provided by the person bringing the action)."[32]

However, most noticeable about the statute is the subsections on damages. Section 3729(a), after listing the seven acts that can give rise to a claim under the FCA, states that not only will the violator be responsible for a civil penalty between $5,000 and $10,000, they will also be liable for treble damages: "[Three] times the amount of damages which the Government sustains because of the act of that person."[33] The statute does provide for reduced damages, which is for violators who admit to their violation, cooperate with the Government during the investigation, and have another FCA claim pending.[34] However, the reduction in damages is not that significant. The statute still states that the court may still charge the violator "not less than [two] times the amount of damages which the Government sustains because of the act of that person."[35] So, despite cooperating with the government, violators of the FCA still face immense penalties.

II. THE DEVELOPMENT OF THE THEORY OF IMPLIED FALSE CERTIFICATIONS

To better understand what the doctrine of implied false certification is, and why it is so controversial, it is best to start at how it was derived. The FCA was initially read very narrowly, restricted only to claims that were, on their face, factually false.[36] This reading gave way to a broader interpretation, labeled as "legally false," to allow claims that were false because they failed to comply with laws and regulations that were certified as conditions to payment.[37] This broader interpretation was then further divided into implied false certifications and express false certifications.[38] The two sections of the FCA that are most important to understanding the

---

[31] 31 U.S.C. § 3730(d)(1)(2010).

[32] *Id.*

[33] 31 U.S.C. § 3729(a)(1)(2009).

[34] 31 U.S.C. § 3729(a)(2)(2009).

[35] *Id.*

[36] Robert Fabrikant & Glenn E. Solomon, *Application of the Federal False Claims Act to Regulatory Compliance Issues in the Health Care Industry*, 51 ALA. L. REV. 105, 112 (1999).

[37] United States *ex rel.* Connor v. Salina Reg'l Health Ctr., Inc., 543 F.3d 1211, 1217 (10th Cir. 2008) (quoting Mikes v. Straus, 274 F.3d 687, 698 (2d Cir. 2001)).

[38] *Id.*

split between the implied and express false certifications are sections 3729(a)(1)(A) and (a)(1)(B). These two sections provides that a person violates the FCA when he:

> (A) knowingly presents, or causes to be presented, a false or fraudulent claim for payment or approval; [or]
>
> (B) knowingly makes, uses, or causes to be made or used, a false record or statement material to a false or fraudulent claim.[39]

The evolution of restricting the FCA to only factually false cases to include legally false cases hints at the courts' increasing trend to expand the scope of the FCA. Viewed this way, the implied false certification theory can be seen as another attempt to expand the scope of the FCA, sparking controversy as to whether it expands the FCA too far.

A. *Factually False v. Legally False*

The dominant interpretation of the FCA in the past was that claims for payment were considered false or fraudulent only if they were "factually false."[40] A "factually false" claim requires that the relator show the contractor had submitted "an incorrect description of goods or services provided or a request for reimbursement for goods or services never provided."[41] Thus, for example, in *Shaw v. AAA Engineering & Drafting, Inc.*, the contractor was required to prepare orders under the contract.[42] The numbers recorded on the orders were tallied monthly and printed in production reports that the contractor was required to submit to the government under the contract.[43] This was important because the work orders assisted the government employees who monitored the contractor's performance.[44] Before trial, it was discovered that some of the numbers recorded in the work orders had been visibly altered and that they may have been falsely inflated.[45] The court held that there was enough evidence to give rise to the "inference that the work orders had been deliberately or recklessly altered for the purpose of causing the government to pay additional sums in the form of equitable adjustment."[46] Thus, in this case,

---

[39] 31 U.S.C. § 3729(a)(1)(A)-(B)(2009).

[40] Fabrikant & Solomon, *supra* note 36, at 112.

[41] *Salina Reg'l Health Ctr., Inc.*, 543 F.3d at 1217 (citation omitted).

[42] Shaw v. AAA Eng'g & Drafting, Inc., 213 F.3d 519, 523 (10th Cir. 2000).

[43] *Id.* at 524.

[44] *Id.*

[45] *Id.* at 526.

[46] *Id.* at 530.

the court found that there was a factually false claim since the work orders were on their faces false statements or descriptions.

However, this narrow construction of the FCA gave way to a broader view. In *United States v. Neifert-White Co.*, the court held that the FCA should not be given a narrow reading.[47] Rather, it should be read so that the statute reaches "to all fraudulent attempts to cause the Government to pay out sums of money."[48] This rationale—that the FCA should be read broadly to include any instance of defrauding the government—led to the recognition of "legally false" claims.[49] Unlike factually false claims, a claim that is "legally false" must be shown to have "'certifie[d] compliance with a statute or regulation as a condition to government payment,' yet knowingly failed to comply with such statute or regulation."[50] The legally false certification claims were further split into two theories: express false certifications and implied false certifications.[51]

B. *Express False Certifications and Implied False Certifications*

An express false certification applies when a contractor "falsely certifies compliance with a particular statute, regulation or contractual term, where compliance is a prerequisite to payment."[52] It does not matter how the statement is made, so long as it is a false statement that relates to a claim.[53] Conversely, the implied false certification theory does not look at actual statements.[54] Rather, it looks at the "underlying contracts, statutes, or regulations themselves to ascertain whether they make compliance a prerequisite to the government's payout."[55] Thus, by merely submitting a claim, the contractor has implied that he would be in compliance with all relevant contract, statutes, and regulations.[56]

The theory of implied false certifications first arose in *Ab-Tech Constrc. Inc. v. United States*, which was decided in 1994 by the U.S. Court of Federal Claims.[57] In this case, Ab-Tech had secured a construction contract with the government.[58] An investigation held by the government

---

47 United States v. Neifert-White Co., 390 U.S. 228, 233 (1986).
48 *Id.*
49 *See e.g.,* United States *ex rel.* Connor v. Salina Reg'l Health Ctr., Inc., 543 F.3d 1211 (10th Cir. 2008).
50 *Id.* at 1217 (citation omitted).
51 *Id.*
52 *Id.*
53 *Id.* at 1217-18.
54 *Id.* at 1218.
55 *Salina Reg'l Health Ctr., Inc.,* 543 F.3d at 1218.
56 *Id.*
57 Ab-Tech Constrc., Inc. v. United States, 31 Fed. Cl. 429 (1994).
58 *Id.* at 430-31.

into Ab-Tech's business affairs led to a criminal indictment of Ab-Tech's president on two counts of making false statements to the Government about Ab-Tech's relationship with one of its principal contractors.[59] The false statements by Ab-Tech's president hid the true relationship between the two companies, which would not have been approved by the government if it had been submitted for review prior to execution in accordance with Section 8(a) of the Small Business Act.[60] After the criminal proceedings against Ab-Tech's president, the Government asserted counterclaims seeking damages and civil penalties pursuant to the FCA.[61]

The Court of Federal Claims held that the payment vouchers that Ab-Tech submitted to the government constituted false claims because they "represented an implied certification by Ab-Tech of its continuing adherence to the requirements for participation in the [minority contractor] program."[62] The court further concluded that the withholding of information on the indemnification clause, which was critical to the decision to pay, was essentially a false claim.[63] A judgment was entered for the government, and the theory of implied false certification was created.

III. THE ENSUING CIRCUIT SPLITS ON THE THEORY OF IMPLIED FALSE CERTIFICATION

Since the *Ab-Tech* decision, the circuit courts have reached varying decisions on whether or not they would recognize the implied false certification theory, and if it were accepted, how far the scope of the theory would reach.

There were three main positions posited by the circuit courts: (I) that the implied certification theory should be recognized only when there is an express condition-of-payment requirement, which was best articulated by the Second Circuit in *Mikes v. Straus* and also adopted by the Third, Sixth, Tenth, and Fourth Circuits;[64] (II) that an implied certification theory should be recognized broadly, which was adopted by the First Circuit in *United States ex rel. Hutcheson v. Blackstone Medical, Inc.* and supported by the D.C. Circuit and the Federal Circuit;[65] and (III) that the implied certification theory should not be recognized, which was articulated by the Seventh Circuit.[66] The Ninth,[67] Eleventh,[68] Fifth,[69] and Eighth[70] Circuits did not fall

59 *Id.* at 431.

60 *Id.* at 434.

61 *Id.* at 431.

62 *Id*. at 434.

63 *Ab-Tech Construc., Inc.*, 31 Fed. Cl. at 434.

64 United States *ex. rel.* Mikes v. Straus, 274 F.3d 687, 697 (2d Cir. 2001).

65 United States *ex. rel.* Hutcheson v. Blackstone Med., Inc., 647 F. 3d 377, 388 (1st Cir. 2011).

66 United States *ex rel.* Steury v. Cardinal Health, Inc., 625 F.3d 262, 270 (5th Cir. 2010); United States v. Sanford-Brown, Ltd., 788 F.3d 696, 711-12 (7th Cir. 2015).

under any of the three main positions. The Ninth and Eleventh Circuits recognized that there was an implied certification theory, but they had not adopted either the *Mikes* or *Blackstone* views. The Fifth and Eighth Circuits had not addressed whether or not to adopt the implied certification theory in some form or to reject it.

A. *Position One: The Implied False Certification Theory is Recognized under an Express Condition of Payment Requirement*

The Second, Third,[71] Fourth,[72] Sixth,[73] and Tenth[74] Circuit had all taken the view that the implied certification should be recognized with the

---

67 The Ninth Circuit had gone back and forth on which theory was most persuasive. In *United States ex rel. Hopper v. Anton*, the court seemed to require an express certification of compliance. 91 F.3d 1261, 1266-67 (9th Cir. 1996). However, in *United States ex rel. Hendow v. Univ. of Phx*., the court held that "[a]n explicit statement, however, is not necessary to make a statutory requirement a condition of payment, and we have never held as much." 461 F.3d 1166, 1177 (9th Cir. 2006). More recently, the court noted that the Second Circuit's analysis in *Mikes* was persuasive and consistent with their precedent but declined to decide whether or not to adopt the theory. Ebeid *ex rel.* United States v. Lungqitz, 616 F.3d 993, 998 (9th Cir. 2010).

68 *See* McNutt *ex rel*. United States v. Haleyville Med. Supplies, Inc., 423 F.3d 1256, 1259 (11th Cir. 2005) (holding that a party submits false claims when it persists in presenting claims to the government that it knows the government does not owe).

69 The Fifth Circuit has addressed the implied certification theory in several cases, although it has not formally recognized the theory. *See* United States *ex rel.* Steury v. Cardinal Health, Inc., 625 F.3d 262, 268 (5th Cir. 2010), *aff'd*, 735 F.3d 202 (5th Cir. 2013); United States *ex rel.* Marcy v. Rowan Cos., 520 F.3d 384, 389 (5th Cir. 2008); United States *ex rel.* Stebner v. Steward & Stephenson Servs., Inc., 144 F. App'x 389, 394 (5th Cir. 2005); United States *ex rel.* Willard v. Humana Health Plan of Tex., Inc. 336 F.3d 375, 382 (5th Cir. 2003).

70 The Eighth Circuit is the only circuit court that has not addressed the implied false certification theory. *See* United States v. Fairview Health Sys., 2004 WL 1638252, at *3 (D. Minn. July 22, 2004) (noting that the Eighth Circuit had not addressed the issue if implied false certification). However, several district courts in the circuit court's jurisdiction have ruled on the theory, although it is unclear whether they were adopting the *Mikes* view or the *Blackstone* view. *See e.g.*, United States v. R.J. Zavoral & Sons, Inc., 894 F. Supp. 2d 1118, 1126 (D. Minn. 2012) (adopting the theory of implied certification from the D.C. Circuit in United States v. Sci. Applications Int'l Corp.); United States *ex rel*. Bryant v. Williams Bldg. Corp., 158 F. Supp. 2d 1001, 1009-10 (D.S.D. 2001) (adopting a broad theory of implied certification, stating that it was in line with the False Claim Act's legislative history of an expansive application of the FCA).

71 The Third Circuit first recognized the implied false certification theory in *United States ex rel. Wilkins v. United Health Grp*. 659 F.3d 295 (3d Cir. 2011). In that case, the relator filed a *qui tam* action alleging that appellees' sales representatives violated the Act by offering physicians illegal kickbacks and violating Medicare marketing rules while simultaneously accepting payments from government funded health insurance programs. *Id.* The court rejected the relator's argument, finding that compliance with Medicare marketing regulations was not conditional to payment and thus, agreeing with the *Mikes* case. *Id.* at 309-10 (holding that the plaintiff must show that compliance with the regulation which the defendant allegedly violated was a condition of payment from the Government when bringing a claim under the theory of implied false certification).

limitation that it was an express-condition of payment requirement. The leading case for this position is *Mikes v. Straus*, decided by the Second Circuit in 2001.[75]

In *Mikes*, the plaintiff's *qui tam* suit claimed that the submission of Medicare reimbursement claims for spirometry[76] services were fraudulent because the procedures were not performed in accordance to the standard of care provided by ATS Guidelines.[77] The court held that the plaintiff had no merit under a theory of express false certification because there was no set standard of care dictated by the forms.[78] Rather, the forms only required procedures to be "medically necessary," indicating the level of service, not the quality of service.[79] The defendants did not expressly comply with a specific standard of care.

Without a case under the express false certification theory, the plaintiff instead focused on the argument that the defendants' submissions to the government for payment were implied false certifications.[80] While the court accepted the theory of implied certification, it did not agree with the plaintiff, finding that the defendants did not submit implicitly false claims because the Medicare statute did not expressly condition payment on

---

72 The Fourth Circuit had voiced skepticism on the implied false certification theory in the past. *See e.g.*, Harrison v. Westinghouse Savannah River Co., 176 F.3d 776, 787 n.8 (4th Cir. 1999) (calling implied certification "questionable"); United States *ex rel.* Herrera v. Danka Office Imaging Co., 91 F. App'x 862, 864 n.3 (4th Cir. 2004) (not deciding whether the implied theory was viable under the Act). However, the Fourth Circuit explicitly adopted the implied false certification theory in *United States v. Triple Canopy Inc.*, and it recognized claims could be false when a party implied compliance with an express condition of payment. United States v. Triple Canopy, Inc., 775 F.3d 628, 636 (4th Cir. 2015).

73 *See* United States *ex rel.* Hobbs v. MedQuest Ass'ns, Inc., 711 F.3d 707, 714 (6th Cir. 2013). In the *Hobbs* case, the Sixth Circuit held that the "FCA does not impose liability for providers' failure to anticipate needs of the program that have not been promulgated in regulations conditioning payment on compliance . . ." *Id.* at 718. *See also* United States *ex rel.* Williams v. Renal Care Grp., Inc., 696 F.3d 518, 532 (6th Cir. 2012) ("The False Claims Act is not a vehicle to police technical compliance with complex federal regulations.").

74 *See* United States *ex rel.* Lemmon v. Envirocare of Utah, Inc., 614 F.3d 1163, 1168-69 (10th Cir. 2010). *See also* United States *ex rel.* Conner v. Salina Reg'l Health Ctr., Inc., 543 F.3d 1211, 1218 (10th Cir. 2008) ("Under an implied false certification theory . . . courts do not look to the contractor's actual statements; rather, the analysis focuses on the underlying contracts, statutes, or regulations themselves to ascertain whether they make compliance a prerequisite to the government's payment."); United States *ex rel.* Lacy v. New Horizons, Inc., 348 F. App'x. 421, 428 (10th Cir. 2009) (affirming the district court's order requiring an express condition of payment clause).

75 *See generally* United States *ex rel.* Mikes v. Straus, 274 F.3d 687 (2d Cir. 2001).

76 The court defined spirometry as a test used by doctors to detect obstructive and restrictive lung diseases. The test uses spirometers, and defendants specifically used the type that allows the tool to measure pressure when patients blow into a mouthpiece. *Mikes*, 274 F.3d at 694.

77 *Id.* at 696.

78 *Id.* at 698.

79 *Id.*

80 *Id.* at 699.

compliance with its terms.[81] Rather, the statute only details the condition of participation in the Medicare program, not a condition of payment.[82] The court reached this reading by looking at the structure of the statute, which establishes conditions of participation rather than prerequisites for payment, and at the sanctions, which ensured that the statute is "directed at the provider's continued eligibility in the Medicare program, rather than any individual incident of noncompliance."[83]

In applying the implied false certification theory, the Second Circuit declined to follow the Federal Circuit in *Ab-Tech*[84] and asserted that they did not wish to read the "theory expansively and out of context.[85] Instead, the court limited the application only to times when "the underlying statute or regulation upon which the plaintiff relies *expressly* states the provider must comply in order to be paid."[86] Thus, under the interpretation that the *Mikes* court developed, a person would only be liable under the Act if he submits a claim for payment to the government and fails to disclose a knowing violation of a statute or regulatory provision the government has expressly conditioned payment upon.[87]

B. *Position Two: A Broad Recognition of the Implied Certification Theory*

The First, D.C., and Federal Circuits adopted the broader recognition of the implied certification theory, which occurs when a contractor submits a claim for payment and fails to disclose any knowing breach of the provisions of their contract.[88] The decision in *Ab-Tech* was the first circuit court opinion to adopt the theory of implied certifications.[89]

---

81 *Mikes*, 274 F.3d at 702.

82 *Id.*

83 *Id.* at 701-02.

84 *See generally* Ab-Tech Constr., Inc. v. United States, 31 Fed. Cl. 429 (1994).

85 *Mikes*, 274 F.3d at 699.

86 *Id.* at 700.

87 Martin, *supra* note 3, at 243. *See also* Michael Holt & Gregory Klass, *Implied Certification Under the False Claims Act*, 41 PUB. CONT. L.J. 1, 37-38 (2011) ("[A] contractor could avoid FCA liability by expressly informing the Government of any material breach or violation when requesting payment.").

88 *See, e.g.,* United States *ex rel.* Hutcheson v. Blackstone Med., Inc., 647 F.3d 377, 386-88 (1st Cir. 2011).

89 *Ab-Tech Constr., Inc.,* 31 Fed. Cl. at 434, *aff'd mem.*,57 F.3d 1084 (Fed. Cir. 1995) (unpublished table decision). As discussed earlier, the *Ab-Tech* court held that the payment vouchers that Ab-Tech submitted to the government constituted false claims because they "represented an implied certification by Ab-Tech of its continuing adherence to the requirements for participation in the [minority contractor] program." *Id.* at 434.

The First Circuit in *Blackstone* further articulated the same principles echoed in *Ab-Tech*.[90] In *Blackstone*, the relator alleged that her former employer paid kickbacks to the hospitals and doctors so that they would use the company's products in spinal surgeries for Medicare and Medicaid patients.[91] The statute that the relator alleged was violated was the Anti-Kickback Statute,[92] causing the hospitals and doctors to submit false claims because compliance with the Anti-Kickback Statute was required to receive payments from healthcare programs, including Medicare, Medicaid, and TRICARE.[93] The district court dismissed the case because there was no express condition of payment requiring that hospitals had to comply with the Anti-Kickback statute.[94] In essence, the district court adopted the *Mikes* court narrow interpretation of the theory of implied false certifications.[95] The First Circuit reversed, stating that the rule "that only express statements in statutes and regulations can establish preconditions of payment is not set forth in the text of the FCA."[96] The court reaches this conclusion by looking at the text of the FCA and also at other circuits that have ruled on the theory of implied false certifications.[97]

The First Circuit thought that the district court's adoption of the *Mikes*[98] court was too narrow, and the text of the FCA did not intend to limit liability in this way.[99] The First Circuit feared that the narrow view would lead to under-inclusion, foreclosing situations that the FCA was created to handle.[100] The court also rejected the concern posited by the *Mikes* court: that a broad view would lead to federalization of what have been private party tort actions.[101] Instead, the First Circuit believed that the knowledge requirement of the FCA and the requirement that the defect be material were enough to rein in the scope of the implied certification theory.[102]

The D.C. Circuit has also adopted a more expansive view of the implied false certifications theory in recent years, going against what it had

---

[90] *See* United States *ex rel*. Hutcheson v. Blackstone Med., Inc., 647 F.3d 377, 386-88 (1st Cir. 2011).

[91] *Id.* at 378.

[92] 42 U.S.C. § 1320a-7b (2006).

[93] *Blackstone,* 647 F.3d at 381.

[94] *Id.* at 383.

[95] *Id. at* 386.

[96] *Id.* at 388.

[97] *Id.* at 387.

[98] United States *ex rel*. Mikes v. Straus*,* 274 F.3d 687 (2nd Cir. 2010).

[99] *Blackstone,* 647 F.3d at 387.

[100] *Id.*

[101] *Id*. at 388.

[102] *Id.*

ruled in the past.[103] In *United States v. TDC Mgmt. Corp.*,[104] the D.C. Circuit departed from its opinion in *Siewick*, opting to follow *Ab-Tech's* more liberal standard. Following the decision in *TDC*, the D.C. Circuit has now clearly shown it will follow the approach set by the Federal and First Circuits.[105]

C. *Position Three: The Implied Certification Theory is Rejected*

The Seventh Circuit was the last circuit court to weigh in on the implied certification theory before the Supreme Court issued its decision, opting to refuse to recognize the implied certification theory in their decision of *Sanford-Brown*.[106] This ruling is consistent with the Seventh Circuit's trend in past decisions to chip away at the broad scope of the FCA.[107]

In this case, *qui tam* relator Brent Nelson initiated a suit under the FCA after he resigned, alleging that "the college's recruiting and retention practices resulted in the transmission of thousands of false claims to the government, potentially subjecting the college and its corporate parent to hundreds of millions of dollars in liability."[108] Under Title IV of the Higher Education Act (HEA), an institution must enter into a Program Participation Agreement (PPA) with the Secretary of Education in order to receive federal education subsidies.[109] Sanford-Brown entered into two PPAs with

---

[103] *See* United States *ex rel.* Siewick v. Jamieson Sci. & Eng'g, Inc., 214 F.3d 1372, 1376 (D.C. Cir. 2000) (holding that "a false certification of compliance with a statute or regulation cannot serve as the basis for a *qui tam* action under the [False Claims Act] unless payment is conditioned on that certification").

[104] United States v. TDC Mgmt. Corp., 288 F.3d 421, 426 (D.C. Cir. 2002).

[105] *See, e.g.,* United States v. Sci. Applications Int'l Corp., 626 F.3d 1257 (D.C. Cir. 2010). The court held in this case held that the plaintiff need only show "that the contractor withheld information about its noncompliance with material contractual requirements." *Id.* at 1269. Even though there was there was no express condition of payment, the district court allowed the government to argue on a theory of implied false certification. *Id.* at 1264. The D.C. Circuit affirmed, holding that an express condition of payment was not necessary to show liability. *Id.* at 1269. As an example of how devastating the damages under the FCA can be, the treble damages and penalties under the FCA exceeded $6 million, despite a jury award of *only* $78 on the breach of contract claim. *Id.* at 1264.

[106] United States v. Sanford-Brown, Ltd., 788 F.3d 696 (7th Cir. 2015).

[107] *See* United States *ex rel.* Absher v. Momence Meadows Nursing Center, Inc., 764 F.3d 699, 712 (stating that it would lead to absurd results if "even a single regulatory violation would be a condition of any and all payments subsequently received . . . "); United States *ex rel.* Yannacopoulos v. Gen. Dynamics, 652 F.3d 818, 824 n.4 (7th Cir. 2011) (holding that a statute must be a certified compliance for a violation of that statue to support a claim under the FCA); United States *ex rel.* Main v. Oakland City Univ., 426 F.3d 914, 017 (7th Cir. 2005) ("Tripping up on a regulatory complexity does not entail a knowingly false representation."). *See also* United States *ex rel.* Kennedy v. Aventis Pharm., 610 F. Supp. 3d 938, 946 (N.D. Ill. 2009).

[108] *Sanford-Brown, Ltd.,* 788 F.3d at 700.

[109] *Id.* at 701.

the U.S. Secretary of Education for their campus in Fenton, Missouri and in Jacksonville, Florida.[110]

Nelson and the government, as *amicus curiae*, argued, based on Nelson's 31 U.S.C. § 3729(a)(1)(A) False Presentment Theory,[111] that Sanford-Brown must comply with all of the conditions in the PPA in order to remain lawfully eligible for the federal subsidies because adherence to the Title IV Restrictions is a part of the "conditions of payment."[112] The court refused to accept Nelson's argument, citing a prior case,[113] which cautioned against a blanket theory of FCA liability that allowed the regulators to terminate for practically any deficiency, no matter how small.[114] The court concluded that it "would be equally unreasonable for us to hold that an institution's continued compliance with the thousands of pages of federal statues and regulations incorporated by reference into the PPC are conditions of payment for purposes of liability under the FCA.[115] Furthermore, the court held that under the FCA, whether or not an institution has violated conditions of participation after good-faith entry into a PPA is for agencies, not the courts, to evaluate and adjudicate.[116]

IV. THE SUPREME COURT'S RULING ON THE THEORY OF IMPLIED CERTIFICATION

With a circuit split so diverse, leading to uncertainty and the potential problem of forum shopping to reach the decisions that would best suit each party, the U.S. Supreme Court has finally resolved the circuit split in a unanimous decision in *Universal Health Servs. v. United States ex rel. Escobar*.[117]

110 *Id.* at 707. The Court of Appeals noted that the PPA for the Fenton campus incorporated thousands of pages of other statues and regulations aside from the HEA program, and even more laws and regulations for the Jacksonville agreement. *Id.* at 709.

111 Nelson has two theories in this case, False Presentment and False Record. The court does not agree with his False Record theory. For the purposes of this comment, the False Record theory will not be discussed because it does not give rise to the implied false certification doctrine.

112 *Sanford-Brown, Ltd.*, 788 F.3d at 709.

113 United States *ex rel.* Absher v. Momence Meadows Nursing Ctr., Inc., 764 F. 3d 699, 712 (7th Cir. 2014) ("[E]ven a single regulatory violation would be a condition of any and all payments subsequently received by the facility inasmuch as the regulators could terminate the facility for practically any deficiency. Such a result would be absurd.").

114 *Sanford-Brown, Ltd.*, 788 F.3d at 711.

115 *Id.*

116 *Id.* at 714.

117 Universal Health Servs., Inc. v. U.S. *ex rel.* Escobar, 136 S. Ct. 1989 (2016).

A. *Factual Background and Procedural Posture of Escobar*

In *Escobar*, the FCA claim arose from misrepresentations in Medicaid reimbursements.[118] Yarushka Rivera, "a teenage beneficiary of the Massachusetts Medicaid program," received counseling services from a mental health facility operated by a subsidiary of Universal Health Services ("Universal").[119] Tragically, Rivera died after an adverse reaction to medication prescribed for her recently diagnosed bipolar disorder.[120] A counselor that worked at the facility revealed to Rivera's parents that there were few employees at the facility that "were actually licensed to provide mental health counseling and that supervision of them was minimal."[121] The psychologist who had diagnosed Rivera as bipolar received a degree from an "unaccredited Internet college" and was not actually licensed to be a psychologist in Massachusetts.[122] This type of misrepresentation was widespread throughout the organization, and it continued in their dealings with the government.[123] Staff members misrepresented their qualifications and licensing status to the government in order "to obtain individual National Provider Identification numbers, which are submitted in connection with Medicaid reimbursement claims."[124] The FCA claim alleged that Universal "submitted reimbursement claims [for] . . . specific services provided by specific types of professionals" but, in violation of the requirements to receive reimbursement, had staff that were "unqualified, unlicensed, and unsupervised."[125] Then, the FCA claim was brought under the implied false certification theory.[126]

The District Court granted the motion to dismiss because "none of the regulations that [the facility] violated was a condition of payment."[127] However, the Court of Appeals reversed, opting to broaden the scope of the implied false certification theory by holding that "a statutory, regulatory, or

---

118 *Id.* at 1993.

119 *Id.* at 1997.

120 *Id.*

121 *Id.*

122 *Escobar*, 126 S. Ct. at 1997. Additionally, the practitioner that prescribed the medication to Rivera represented herself as a psychiatrist but was actually a nurse who was legally unable to prescribe medication absent supervision.

123 *Id.*

124 *Id.* National Provider Identification numbers are submitted in connection with Medicaid claim and corresponding to specific job titles. *Id.*

125 *Id.* at 1997-98. This was problematic as the Massachusetts Medicaid program required satellite facilities, which this facility was, to have specific types of professionals, licensing requirements, and supervision requirements.

126 *Id.* at 1997.

127 *Id.* at 1998.

contractual requirement can be a condition of payment either by expressly . . . or by implication."[128]

B. *The Supreme Court's First Holding: Implied False Certification May Sometimes Be a Basis for Liability*

In its decision, the Supreme Court held that in certain cases, the implied false certification theory could be a basis for liability.[129] The Court's support for accepting the implied false certification theory lies with the language of the statute.[130] The Supreme Court focused on the statutory language of § 3729(a)(1)(A), specifically that the FCA imposes civil liability for "a false or fraudulent claim."[131] As Congress did not define "false" or "fraudulent" in the context of the FCA, the Supreme Court relied on the statutory rule of construction to conclude that "Congress intends to incorporate the well-settled meaning of the common-law terms it uses" absent other indications.[132] Thus, since "fraudulent" is a term that incorporates the common-law meaning of the term fraud, that is the meaning that Congress intended it to have.[133] Furthermore, since the common-law definition of fraud includes misrepresentations by omission, the Supreme Court emphasized that "false or fraudulent" claims under the FCA should likewise include both express misrepresentations and misrepresentations by omission.[134] Thus, according to the Court, there is both a statutory basis and a common-law basis for implied false certifications under the FCA.[135]

However, the Supreme Court, declining to follow the First Circuit's version of the implied theory of certification, stated two preconditions as a means of limitation: (1) the contractor submits claims for payment, "mak[ing] specific representations about the goods or services provided" and (2) the contractor "knowingly fails to disclose [its] noncompliance with a statutory, regulatory, or contractual requirements makes those representations misleading half-truths."[136]

---

[128] *Escobar*, 136 S. Ct. at 1998. Although this should not be surprising, *Escobar* came up on appeal through the First Circuit, home to the *Blackstone* view, an extremely broad standard for implied certification liability.

[129] *Id*. at 1999.

[130] *Id.*

[131] *Id.*

[132] *Id* (quoting Sekhar v. United States, 133 S. Ct. 2720, 2724 (2013)).

[133] *Escobar*, 136 S. Ct. at 1999 (citing Neder v. United States, 527 U.S. 1, 22 (1999)).

[134] *Id.*

[135] *Id.*

[136] *Id.* at 2001. The preconditions were summarized earlier in the opinion as when "the defendant submits a claim for payment that makes specific representations about the goods or services provided, but knowingly fails to disclose the defendant's noncompliance with a statutory, regulatory, or

C. *The Supreme Court's Second Holding: Liability for Not Disclosing Violations Does Not Turn On Whether Those Requirements Were Expressly Designated As Conditions of Payments*

The Court further held that the government did not have to designate a provision in a contract as being a condition of payment to face FCA liability.[137] This view rejects the interpretation of the implied false certification theory posited by the Second Circuit in *Mikes*, which required there to be an express condition.[138] However, conversely, not every undisclosed violation of an express condition will trigger liability in and of itself.[139] Then, what triggers liability?

The Supreme Court points to the materiality inquiry as a means for determining whether liability should lie.[140] The Court noted that the FCA already had a section defining materiality, § 3729(a)(1)(A), which states that a "misrepresentation must be material to the other party's course of action."[141] The FCA further defines "materiality" as "having a natural tendency to influence, or be capable of influencing, the payment or receipt of money or property."[142] In addition to the materiality definition already contained within the FCA, the Court also noted that the definition of materiality is also derived from the common law.[143] However, the Court declined to determine whether or not the FCA's definition or the common-law definition of materiality reigns.[144] Rather, the Court determined that regardless of which concept of materiality is used, materiality looks for the effect "on the likely or actual behavior of the recipient of the alleged misrepresentation."[145] In parsing the definition this way, the Court makes it clear that materiality is not some minor or throwaway test to determine liability. Instead, the Court acknowledges that the FCA was not intended to catch all sorts of fraud, nor was it meant to punish minor or miniscule breaches or regulatory violations.[146] Thus, there must be some test to limit

---

contractual requirement. In these circumstances, liability may attach if the omission renders those representations misleading." *Id.* at 1995.

137 *Id.* at 2001.

138 *See* United States *ex rel.* Mikes v. Straus, 274 F.3d 687, 702 (2d Cir. 2001), abrogated by *Escobar*, 136 S. Ct. 1989 (2016).

139 *Escobar*, 136 S. Ct. at 2001.

140 *Id.*

141 *Id.*

142 31 U.S.C.A. § 3729(b)(4) (2009).

143 The Court emphasized the need for materiality, "the common law could not have conceived of 'fraud' without proof of materiality." Neder v. United States, 527 U.S. 1, 22 (1999). Specifically, the Court looks to tort and contact law to define materiality, noting that materiality either "induces" and or causes a person to "attach importance" to take a particular action. *Escobar*, 136 S. Ct. at 2003.

144 *Id.* at 2002.

145 *Id* at 2003 (quoting RICHARD A. LORD, WILLISON ON CONTRACTS §69:12 (4th ed. 2003)).

146 *Escobar*, 136 S. Ct. at 2003.

the potentially broad reach of the Act under the theory of implied false certification.

However, for such a demanding standard, the Supreme Court provides sparse guidelines on what would be considered material. One factor that might be indicative of materiality is the government's decision to expressly label something as a condition of payment, but even that would not be automatically dispositive.[147] Another factor could be the government's continual refusal to pay because one of the terms was not being upheld.[148] The Court seems to find it easier to define what would not be material, rather than what would be.[149] Examples of scenarios not dispositive of materiality could be the government paying a claim despite actual knowledge of requirements being violated[150] or the government having an option to decline payment if they knew about the noncompliance or situations where the noncompliance is minor.[151]

Apart from these scenarios, the Supreme Court has left it to the lower courts to determine the boundaries of materiality.

D. *Cases Post-Escobar*

At a glance, the heightened materiality standard, coupled with the required pleadings under the Federal Rules of Civil Procedure,[152] works in favor of defendants. Already, there are several cases that have resulted in dismissal due to the government or relator being unable to plead the case as required by the *Escobar* materiality standard.[153]

In *U.S. ex rel. Lee v. Northern Adult Daily Health Care Center*, one of the first cases to substantively apply *Escobar*, the court granted the motion

---

147 *Id.*

148 *Id.*

149 *Id.*

150 This shows a lack of materiality because the government did not see those terms as important enough to enforce despite knowing that they were a part of the contract and despite knowing that the contractor was not abiding by those terms. *Id.* at 2003-04.

151 *Id.* at 2003.

152 In the Court's justification for the heightened materiality standard, and how pleadings at the summary judgment and motion to dismiss stages can still occur, the Court emphasizes that under the Federal Rules of Civil Procedure 8 and 9(b), plaintiffs must "plead their claims with plausibility and particularity . . . pleading facts to support allegations of materiality." *Escobar*, 136 S. Ct. at 2004 n.6.

153 There are several other cases decided post-*Escobar* that discuss the *Escobar* materiality standard, as well as cases that have been remanded. For cases that have been dismissed due to failing to plead sufficiently, *see e.g.,* United States *ex rel.* Voss v. Monaco Enters., Inc., No. 2:12-CV-00460LRS, 2016 WL 3647872 (E.D. Wash. July 1, 2016); United States *ex rel.* Creighton v. Beauty Basics, Inc., No. 2:13-CV-1989-WEH, 2016 WL 3519365 (N.D. Ala. June 28, 2016). For cases that have been remanded, *see e.g.,* Triple Canopy, Inc. v. U.S. *ex rel.* Badr, 136 S. Ct. 2504 (Mem.), No. 14-1440 (Jun. 27, 2016); Weston Educ., Inc. v. U.S. *ex rel.* Miller, 136 S. Ct. 2505 (Mem.), No. 15-404 (June 27, 2016).

to dismiss the case.[154] By doing so, the court acknowledged that there was a difference between the prior standard and the new standard: "While Relators' argument may have sufficed to support an implied false certification claim under the standard in *Mikes*, it no longer suffices under the standard in [*Escobar*]."[155] Specifically, the court stated that the relators failed to allege how "noncompliance with Title VI and the DOH regulations . . . would have influenced the government's decision to reimburse Northern Adult."[156] Interestingly, the court's opinion also pointed to the common law definition of materiality that *Escobar* provides as a way to define materiality, but not to the definition of materiality under the FCA.[157] However, given that the amended complaint in this case was drafted prior to the *Escobar* decision, the court granted the relators leave to amend the amended complaint.[158]

Similarly, the Seventh Circuit revisited *United States v. Sanford-Brown, Ltd.* to, once again, deny FCA liability.[159] The court explained materiality is no longer "that the Government would have the option to decline to pay" but rather looks at the action's effect "on the *likely* or *actual* behavior of the recipient of the alleged misrepresentation."[160] Differing from the court in *Northern Adult,* the Seventh Circuit cited to the definition of materiality from the FCA, not from the section detailing the common law definition. This difference does not seem to affect the reasoning for dismissal.[161] Similar to the court's reasoning in *Northern Adult*, the Seventh Circuit stated that the relator failed to show that the government's decision to pay would likely or actually have been different if it knew about the noncompliance.[162]

However, just as there are cases that have been dismissed due to their failure to meet the *Escobar* materiality standard, there are cases that have survived summary judgments and will continue to trials.

In *Rose v. Stephens Institute*, the District Court for California denied summary judgment to the defendant.[163] First, the court stated that the two conditions established by the *Escobar* court to determine whether there is a

---

[154] United States *ex rel.* Lee v. N. Adult Daily Health Care Ctr., 13-CV-4933, 2016 WL 4703653, at *12 (E.D.N.Y. Sept. 7, 2016).

[155] *Id.*

[156] *Id.*

[157] *Id.* at *11 (citing *Escobar*, 136 S. Ct. at 2003-04).

[158] *Id.* at *12.

[159] United States v. Sanford-Brown, Ltd., 840 F.3d 445, 448 (7th Cir. 2016).

[160] *Id.* at 447 (citing *Escobar*, 136 S. Ct. at 2002-03).

[161] However, whether the courts consciously chose to defer to the definition of materiality as seen in the Act or the definition at common law, by choosing different sources for the definition, this sets the stage for a potential future circuit split post-*Escobar* on what exactly is materiality.

[162] 840 F.3d at 447.

[163] Rose v. Stephens Inst., No. 09-5966, 2016 U.S. Dist. LEXIS 128269 at *20 (N.D. Cal. Jun. 23, 2016).

basis for liability did not have to be met in every circumstance.[164] This goes against the defendant's argument that the two-part test had to apply in every single implied false certification claim for there to be liability.[165] However, more importantly, the court determined that the relators had fulfilled the materiality requirement under *Escobar*.[166] In reaching this conclusion, the district court determined that the government's decisions to not take action against the defendants "despite its awareness of the allegations in this case is not terribly relevant to materiality."[167] Additionally, the court also did not determine that the government's uneven enforcement of its policies indicated a lack materiality, stating that it "does not prove that [the government considered the defendants'] violations immaterial or unimportant to the Title IV bargain."[168] Rather, the court seemed to sympathize more with the relators, going so far as to excuse the government's past enforcement issue because there had been policy changes.[169] Thus, the court determined that the question of materiality was a triable issue and that summary judgment was inappropriate.[170]

These three cases are only the tip of the iceberg, but they highlight the potential problems of the *Escobar* decision.

IV. IMPLICATIONS OF THE NEW IMPLIED FALSE CERTIFICATION THEORY

With its decision in *Escobar*, the Supreme Court has implemented a new standard for implied false certification claims and, in doing so, effectively rejected all the circuit courts' positions. While the Supreme Court has clarified that implied false certification is a viable theory, its

---

164 *Id.* at *15,

165 *Id.* at *13. It is interesting that the district court interpreted "at least where two conditions are satisfied" as a statement saying the conditions were optional. *Escobar*, 136 S. Ct. at 2001. Rather, the district court's reasoning was that the Supreme Court used "at least" to indicate that it would not determine whether the implied false certification theory was viable in all cases, and the Supreme Court used "at least" in the same context when describing the two conditions. *Rose*, 2016 U.S. Dist. LEXIS 128269 at *14. Thus, the two conditions were not an absolute requirement. *Id.*

It is true that the Supreme Court did decline to determine whether the implied certification theory was viable in all circumstances, but there does not seem to be any tie between the use of "at least" in that circumstance and with its usage in the statement at debate in front of the district court. This is an interesting interpretation, and slightly strange, given that the entire holding implied that the two conditions were required, " . . . [W]e hold that the implied certification theory can be a basis for liability, at least where two conditions are satisfied . . ." *Escobar*, 136 S. Ct. at 2001. At the very least, this difference in interpretation, as compared to the courts in *Northern Adult* and *Sanford-Brown*, foreshadows potential circuit splits and problems post-*Escobar*.

166 *Rose*, 2016 U.S. Dist. LEXIS 128269 at *17.

167 *Id.*

168 *Id.* at *18.

169 *Id.* at *20.

170 *Id.*

attempts to limit the theory's wide reach through a higher materiality standard have led to more questions than answers.

This comment advocates that the circuit courts should employ a modified form of the outcome determinative test, based on the Eight Circuit's decision in *Costner v. Urs Consultants*, to determine materiality.[171] The modified test acknowledges that overall, the materiality standard under *Escobar* calls for a fact-intensive balancing test but, in order to encourage uniform application among the circuit courts, the factor that the courts should give the most weight to is the government's past payment practices.[172]

A. *Benefits of A Heightened Standard for Materiality*

While there are issues with the Supreme Court's materiality standard in *Escobar* that will be addressed later on in the section, it is important to note that there are many benefits to the materiality standard as it stands.

1. Placing the Burden on the Plaintiff to Plead Materiality is More Economically Efficient

One of the benefits of a heightened materiality standard is the reduction of economic inefficiency due to the incentive issues between the *qui tam* relator and the government. This economic inefficiency in turn affects contractors, especially smaller contractors or new contractors who cannot risk liability claims under the FCA due to the punitive damages.

Under section 3730(d)(1), a *qui tam* relator has the ability to receive anywhere from 15% to 25% of the proceeds of the action or settlement of the claim if the government decides to support them.[173] However, if the government does not support the relator, he or she can receive from 25% to 30% of the proceeds of the action or settlement of the claim along with compensation for their court costs.[174] Given that the damages under the FCA include treble damages, even 15% of a successful claim can give the *qui tam* relators an immense payout.[175]

---

171 Costner v. Urs Consultants, 153 F.3d 667, 677 (8th Cir. 1998).

172 The government payment practices include claims that the government has consistently refused to pay based on noncompliance, claims that the government pay despite actual knowledge of violation, or when the government regularly pays a particular type of claim despite actual knowledge of violations and signals it will not change its position on payment.

173 31 U.S.C. § 3730(d)(1) (2010).

174 31 U.S.C. § 3730(d)(2) (2010).

175 *See e.g.* United States v. Sci. Applications Int'l Corp., 653 F.Supp.2d 87, 94 (D.D.C. 2009) (where even though the government was only awarded $78 in damages for the contract claim, it received $577, 500 in civil penalties and $5,921,518.83 in treble damages).

There are many reasons that factor into whether or not the government will pursue a claim under the FCA. The government must not only consider how taking on a case affects its resources and time,[176] but also consider both social and private interests.[177] However, a *qui tam* relator does not necessarily share the same incentives. Section 3720(d)(2) allows for *qui tam* relators to bring a case against a contractor under the FCA without needing the government to intervene, and they will also receive a larger percentage of the settlement or compensation for a successful case.[178]

The theory of implied false certification in the past created more opportunities for the *qui tam* plaintiff to pursue these incentives since it gives the plaintiff an easier method of reaching liability under the FCA. A report by the GAO noted that four of the U.S. Attorneys' Offices participating in an initiative had taken actions inconsistent with the DOJ's guidance on bringing claims under the FCA.[179] These offices, without gathering the evidence required by the guidance, had sent letters alleging or implying FCA violations to many hospitals.[180] The letters stated the penalties each hospital would face, and they encouraged settlement amounts that were two times more than the overpayments identified.[181] While the report does not state whether the violations were brought under an implied or express false certification, this report shows the possibility of abuse under the FCA regardless of what type of certifications was alleged. Since the theory of implied false certification widened the scope of liability under the FCA, it would have allowed a *qui tam* plaintiff and the government to more easily find an opportunity for liability.

The possibility of a lawsuit that carries with it treble damages and statutory penalties could discourage new contractors or small business contractors from entering into the market. This leads to economic inefficiency in that businesses that should be taking on contracting jobs would be afraid to do so in fear of being sued for immense damages. Additionally, well-established contractors could overcompensate for a possible lawsuit *ex ante* by either submitting more expensive bids, costing the government more in the long run. Thus, limiting the reach of the theory by requiring plaintiffs at the pleading stages to not only comply with the Federal Rules of Civil Procedure 8 and 9(b) but to also plead with a higher standard of materiality prevents the economic inefficiency skewed against defendants.

---

[176] Michael Lawrence Kolis, *Settling for Less: The Department of Justice's Command Performance Under the 1986 False Claim Amendments Act,* 7 ADMIN L. J. AM. U. 409, 438 (1993).

[177] *See* Ben Depoorter & Jef De Mot, *Whistle Blowing: An Economic Analysis of the False Claims Act*, 14 SUP. CT. ECON. REV. 135, 154 (2006).

[178] 31 U.S.C. § 3720(d)(2) (2010).

[179] U.S. GOV'T ACCOUNTABILITY OFFICE, GAO/HEHS-00-73, Medcare Fraud and Abuse: DOJ has Made Progress in Implementing False Claims Act Guidance (2000).

[180] *Id.*

[181] *Id.*

2. Limiting the Reach of the Theory of Implied False Certifications would Promote Fairness

The stakes of liability under the False Claims Act are extremely high given that the Act works not only to compensate the government but also to severely punish transgressors with treble damages and other penalties.[182] The Supreme Court itself has commented on the punitive nature of the FCA, stating that the increase from double damages to treble damages had turned a remedial scheme into "an essentially punitive one."[183] Due to its punitive nature, there must be fair notice that alerts a contractor that noncompliance will result in a punitive sanction.[184] A broad theory of implied false certifications does away with fair notice, leaving contractors vulnerable to a possible suit under the False Claims Act for failing to comply with one of the thousands of statutes and regulations under their contract. The Court in *Escobar* also conveyed that concern, stating multiple times that the FCA was not "an all-purpose antifraud statute" designed to catch even the most trivial of breaches.[185]

Some might argue that doing away with the theory of implied false certifications would favor the contractors too much, and it would allow cases of actual fraud to slip through the cracks. Doing so would go against the purpose of the FCA, which is to reach all fraudulent claims that cause the government to pay more money than it should.[186] However, cabining the theory of implied false certifications does not mean that the FCA will no longer be able to achieve its purpose. The 1986 amendments to the FCA broadened the abilities of the relator to bring actions against contractors acting fraudulently, as well as made it easier for relators to file claims.[187] Incentives for the relators were also bolstered by increased financial rewards, ensuring that there would still be a market for FCA claims.[188] Additionally, the government has other methods that work in tandem with the FCA to ensure that contractors comply with their contracts and that dishonest or fraudulent contractors are punished through suspension and debarment under the Federal Acquisition Regulations (FAR) and breach of

---

182 *See* 31 U.S.C. § 3729(a)(1) (2009).

183 *See* Cook County v. U.S. *ex rel.* Chandler, 538 U.S. 119, 120 (2003) ("[T]hat the change from double to treble damages turned what had been a "remedial" provision into an "essentially punitive" one."); Vermont Agency of Nat. Res. v. U.S. *ex rel.* Stevens, 529 U.S. 765, 800 (2000).

184 *See* Philip Morris USA v. Williams, 549 U.S. 346, 354 (2007) (where concerns about punitive damages are the risks of arbitrariness, uncertainty, and lack of notice); BMW of N. Am., Inc. v. Gore, 517 U.S. 559, 574 (1996) ("Elementary notions of fairness . . . dictate that a person receive fair notice not only of the conduct that will subject him to punishment, but also of the severity of the penalty that a State may impose.").

185 Universal Health Servs. v. U.S. *ex rel.* Escobar,136 S. Ct. 1989, 2003 (2016).

186 United States v. Neifert-White Co., 390 U.S. 228, 232 (1986).

187 False Claims Act, Pub. L. No. 99-562, 100 Stat. 3153.

188 *Id.*

contract remedies by the agency. For example, the Environmental Protection Agency has their regulations codified in Title 40 of the CFR,[189] and the Food and Drug Administration has their regulations in Title 21 of the CFR.[190] If there are needs that the agency would have to address, they can issue rules through the "notice and comment rulemaking" process.[191] Additionally, the FAR, the main set of rules that governs the process by which executive agencies contract with appropriated funds from the government,[192] outlines all the requirements that a contractor would have to comply with, as well as providing penalties for noncompliance. Under the FAR, a contractor that repeatedly does not follow regulations can be suspended or debarred.[193]

Thus, cabining the theory of implied false certifications would ensure fair notice by limiting the liability of contractors to material claims without hindering the purpose of the FCA. This acknowledges that contractors cannot perfectly comply with the complex regulatory scheme they have to work with and that a breach of these terms does not immediately result in punitive damages.

B. *Circuit Courts Should Follow a Modified Form of the Outcome Determinative Test to Determine Materiality*

Unfortunately, the heightened materiality standard of *Escobar* operates as a double-edged sword. While strengthening the already existing materiality requirement is an effective way to limit the reach of the theory, especially given the strong combination of having both a statute and common law definition, the lack of a clear cut example of materiality has led to confusion as to what would be considered material. The factors that the Court provided are, at best, scant guidelines. By leaving it to the lower courts to shape what facts would determine materiality, the Supreme Court has provided an opportunity for a circuit split to emerge. To prevent a split from happening, the circuit courts should adopt a modified version of the outcome determinative test under the Eighth Circuit.

---

189 40 C.F.R. (2015).

190 21 C.F.R. (2015).

191 OFFICE OF THE FEDERAL REGISTER, A GUIDE TO THE RULEMAKING PROCESS (2011), *available at* https://www.federalregister.gov/uploads/2011/01/the_rulemaking_process.pdf.

192 48 C.F.R. § 1.101 (2015).

193 FAR 9.4 (2015).

1. The Outcome Determinative Test Under the Eighth Circuit and the Need for Modification

The Eighth Circuit in *Costner v. URS Consultants* established the outcome determinative test as one that "requires a showing that the alleged fraudulent actions had 'the purpose and effect of causing the United States to pay out money it is not obligated to pay, or those actions which intentionally deprive the United States of money it is lawfully due.'"[194] The court concluded that the test reached actions that had the purpose or effect of causing the government to pay money that it need not pay or actions that were intentionally made to deprive the government of money.[195] The reasoning for the test was derived from the text of the FCA, specifically section 3729(a)-(c).[196]

2. The Outcome Determinative Test is the Better Method of Determining Materiality Post-Escobar Among the Other Test of Materiality

Overall, there are two major tests determining materiality: the outcome determinative test and the natural tendency test. Under the natural tendency test, a court would consider "whether the false statement has a natural tendency to influence agency action or is capable of influencing agency action."[197] Thus, the natural tendency test is more concerned with the potential ramifications of the false statement or claim, rather than its actual effect. However, the outcome determinative test, as explained in the prior section, looks to whether a false statement or claim induced the government to pay more than it should or were made with the intention of depriving the government of money. Unlike the natural tendency test, the outcome determinative test is more concerned with the actual effect of the false claim.

While the language of the natural tendency test can be found in the text of the FCA, the test runs into the issue of being extremely vague. Anything can be influential, but the test given as is does not elaborate on the parameters.[198] This could allow the government or relators to bring cases for garden-variety breaches of contract or regulatory violations, which

---

194 United States *ex rel.* A+ Homecare, Inc. v. Medshares Mgmt. Group, Inc., 400 F.3d 428, 445 (6th Cir. 2005) (quoting Costner v. URS Consultants, Inc., 153 F.3d 667, 677 (8th Cir. 1998)).

195 *Id.*

196 *Costner*, 153 F.3d at 677.

197 United States v. Norris, 749 F.2d 1116,1122 (4th Cir. 1984).

198 Megan Hoffman noted in her comment that "any small error in the claim or in a statement could "potentially" affect the government's decision to pay, but what is unclear is the level of potentiality the statement or claim must have before it is considered material." Hoffman, *supra* note 14, at 201.

is not the purpose of the FCA.[199] Given the punitive nature of the FCA, due to the treble damages that defendants could face, a vague standard that almost encourages variances among courts creates uncertainty and does not solve the problems of the materiality standard in *Escobar*.

Additionally, there has also been another test proposed, the substantial weight test. The test proposes that:

> while the government need not actually have paid out on the false claim, the alleged false statement must be the type often considered by the government in the decision-making process, and its effect must be given substantial weight in the government's decision with respect of the payment of the claim.[200]

Though this test is better and more specific than the natural tendency test, it creates an unnecessary restriction by requiring that the false statement be a type often considered by the government in its decision-making. Hoffman seems to connect claims given substantial weight as claims that have been typically considered in prior cases.[201] However, this reasoning runs the risk of being under-inclusive. Contracts will vary among the contractors that the government works with. Relying on how often a provision is brought up in the government's decision-making process does not necessarily indicate that the provision itself is important. A newer, more important provision might be negotiated in; however, the substantial weight test would bar its entry since it would not be the term that has been often considered by the government.

Thus, compared to the natural tendency test and the substantial weight test, the outcome determinative test is the best vehicle from which to define materiality under *Escobar*. The outcome determinative test is clear, focusing on the economical detriment to the government, which follows the purpose of the FCA. In doing so, the test places the burden on the plaintiff to show that the government would have acted differently if it had known of the omission, which is a clear showing of materiality because there is a definite link between the claim and how it influenced the government to act one way or another.[202] Hoffman criticizes the outcome determinative test as being too restrictive, "put[ting] substantial weight on the government's actions to determine the materiality of the claim and not on the nature of the defendant's statements or claims."[203] However, that is not a fair depiction

---

199 Based on the Supreme Court in *Escobar*: "The False Claims Act is not 'an all-purpose antifraud statute' . . . or a vehicle for punishing garden-variety breaches of contract or regulatory violations." *Escobar*, 136 S. Ct. at 2003.

200 Hoffman, *supra* note 14, at 200.

201 *Id.* at 205.

202 Additionally, this also ensures that the plaintiff is able to sustain a pleading in line with Federal Rules of Procedure 8(b) and 9, which the Supreme Court in *Escobar* has emphasized needed to be fulfilled in order to find for materiality.

203 Hoffman, *supra* note 14, at 203.

of the test. The nature of the defendant's statements or claims is important in the sense that there must be a showing of fraud or falsity. However, showing fraud does not necessarily mean that the claim itself is material. The government's action is a basis for materiality because it clearly shows what the government considered to be material.

However, despite being the best option of the three tests, the outcome determinative test still does not conform to the standard of materiality under *Escobar*. Modifications are needed to bring the test, which was developed in 1998, in line with the more modern interpretation of the FCA.

3. The Need for Modification to the Outcome Determinative Test under the Eighth Circuit

The Eight Circuit's test in *Costner* tracks the general tone of the Supreme Court's standard for materiality, placing emphasis on an influential act or an act that is meant to be influential. However, *Costner*, viewed under the lens of *Escobar*, would be too restrictive. *Escobar* clearly intended for materiality to be balanced among various factors. The Supreme Court does not pronounce a clear bright-line rule, but instead, opts for a case-by-case evaluation. Ascribing to *Costner*'s original holding, that materiality depends on whether or not the government had suffered a financial detriment and ultimately whether the plaintiff can prove that the government would have acted differently had it known the claim was false,[204] ignores other factors the Supreme Court found important to mention.[205]

The modification to the outcome determinative test is slight. The test will still require that there be a showing of financial detriment that could have been avoided if the government knew that the claim was false. However, to avoid vagueness, the factor to focus on is the government's payment practices.[206] This strikes a happy medium between the too restrictive test of the *Costner* court but reins in the vagueness of the *Escobar* materiality standard by giving courts a factor to divert their attention towards. This does not mean that other factors would not be considered. Rather, the other factors may outweigh the government's payment practices if they were overwhelmingly for or against a claim's

---

204 *Id.*

205 These factors include: the defendant's knowledge of what the government considers important, what conditions in the contract itself that are express or labeled as important, or if the government would have had the option to decline payment if it knew of the defendant's noncompliance.

206 There are multi-factor tests that also place emphasis on one or two specific facts. For example, the test for a limited-purpose public figure is a five-factor test. However, the second and third factors are often conflated and are regarded as the more important factors. *See* Carr v. Forbes, Inc., 259 F.3d 273, 280 (4th Cir. 2001). Thus, having a test that favors one or two factors over the others is not unprecedented and is beneficial in helping the courts focus their analysis.

materiality. Additionally, if the government payment practice does not speak too strongly for or against materiality, another factor or factors may be used to help tip the scale.

However, by focusing on the government's payment practices, it will place the burden on the government to show what terms to them are important. The government is in the best position to show what is material to them, and payment practices provide a visible and definite measure of materiality. For example, if the government has repeatedly paid for a claim despite failure on the contractor to fulfill a term of the contract, it is most likely that the condition is not material.

Thus, this modified form of the outcome determinative test provides a uniform point on which all courts can focus their analysis. The courts then can accordingly balance the rest of the factors they consider up next evidence of the government's payment practices.

CONCLUSION

For years, the circuits have been split on whether or not to accept the theory of implied false certification and what exactly the scope of the theory was. However, the Supreme Court laid the split to rest with its decision in *Universal Health Services v. United States ex rel. Escobar*. Although the Supreme Court accepted that the theory of implied false certifications, it limited the reach of the theory by heightening the materiality standard and stating two preconditions that need to be fulfilled: (1) the contractor submits claims for payment, "mak[ing] specific representations about the goods or services provided" and (2) the contractor "knowingly fails to disclose [its] noncompliance with a statutory, regulatory, or contractual requirements makes those representations misleading half-truths."[207]

The Supreme Court's limitations on the theory of implied false certifications are a step in the right direction. Restricting the theory promotes fairness as well as being more economically efficient. However, the vagueness surrounding what materiality exactly means has led to varying decisions in cases post-*Escobar*. To remedy what might potentially be a new circuit split on the theory of implied false certifications, this Comment advocated that the circuit courts should employ a modified form of the outcome determinative test, based on the Eight Circuit's decision in *Costner v. Urs Consultants*, to determine materiality.[208] The modified test acknowledges that overall, the materiality standard under *Escobar* calls for a fact-intensive balancing test, but, in order to encourage uniform

---

[207] Universal Health Servs. v. U.S. *ex. rel.* Escobar, 136 S. Ct. 1989, 2001 (2016).

[208] Costner v. Urs Consultants, 153 F.3d 667, 677 (8th Cir. 1998).

application among the circuit courts, the factor that the courts should give the most weight to is the government's past payment practices.

In applying the heightened materiality standard in this way, it allows the courts to focus their attention on one factor: the government's past payment practices. This does not mean that this is the only factor that matters. However, past payment practices are a definitive and clear example of materiality. This also shifts the burden on the government to provide what terms are material to them, which is something the government is in the best position to provide.

The Circuit Courts and lower courts should thus consider employing this test to reach more uniform decisions.