Denver Law Review

Volume 88 | Issue 1 Article 9

December 2020

Vol. 88, no. 1: Full Issue

Denver University Law Review

Published by the
University of Denver
Sturm College of Law

2010 Volume 88 Issue 1

DENVER UNIVERSITY LAW REVIEW

ARTICLES

COMMENTS

REVITALIZING UNION DEMOCRACY: LABOR LAW, BUREAUCRACY, AND WORKPLACE ASSOCIATION

MATTHEW DIMICK†

Do core doctrines of labor-relations law obstruct the internal democratic governance of labor unions in the United States? Union democracy is likely an essential precondition for the broader strategic and organizational changes unions *must undertake in order to recruit new union members—the labor movement's cardinal priority. Yet according to widely accepted wisdom, the weakness of democracy within labor unions is the unavoidable outcome of an "iron law of oligarchy" that operates in all such membership-based organizations. This Article challenges this conventional thinking and argues that the triumph of oligarchy over democracy in U.S. labor unions is not inevitable, but conditioned on the nature of American labor law. The main message is that labor law will directly or indirectly undermine what I call "workplace association," a decisive strategic component in the florescence of union democracy, when, as in the U.S., it: (1) provides for exclusive representation; (2) establishes institutions and procedures for collective bargaining; and (3) inhibits the use of economic "self help" as alternatives to such procedures. To reach this conclusion, the Article develops a game-theoretic model of union democratization, formalized in the Appendix, that highlights the role of union bureaucracy and workplace association in the success or failure of union democracy. The Article then uses the model to analyze the impact of U.S. labor law on this game of union democracy, and makes comparison to Great Britain, where labor law has contrasted dramatically, with equally divergent results for union democracy.*

INTRODUCTION

In January of 2009, the Service Employees International Union (SEIU) placed its United Healthcare Workers-West (UHW) affiliate into trusteeship, removing all of its elected leaders.[1] Prominent friends of the

† Law Research Fellow, Georgetown University Law Center. E-mail: md498@law.georgetown.edu. I wish to thank Hannah Alejandro, Peter Byrne, Nancy Chi Cantalupo, Kenneth Dau-Schmidt, Alexa Schindel Dimick, Ivan Ermakoff, William Forbath, Robert Freeland, Scott Gehlbach, Gregory Klass, James Montgomery, Joel Rogers, Logan Sawyer, Katherine Stone, Mark Suchman, Christopher Tomlins, Robin West, Erik Olin Wright, Jonathan Zeitlin, seminar participants at Georgetown University Law Center, and conference participants at the 2009 annual meetings of the Midwest Law and Economics Association for helpful comments on earlier versions of this Article.

1. Michelle Amber & Joyce E. Cutler, *Former Leaders of UHW Local Quit SEIU, Plan to Form New Health Care Workers Union*, 17 Daily Lab. Rep. (BNA) A-17 (Jan. 29, 2009).

labor movement met the news with dismay.[2] By all accounts, the large, 150,000-member UHW was a "model" and "democratic" local union.[3] The trusteeship ended a long internal feud between SEIU and UHW over a plan to sever 65,000 members from UHW and merge them with workers from two other affiliates.[4] As the conflict reached its apogee, the UHW leadership said they would only accept the reorganization plan if members of the local were allowed to vote on it.[5] The trusteeship was imposed the next day.[6] At the heart of the conflict, according to one service clerk and former elected leader of the UHW, is union democracy: whether UHW members will "be part of a union that they control democratically" or "one that is led by a handful of outsiders from Washington, D.C."[7] Received opinion would regard this outcome as just another example of a general and inherent organizational dynamic toward oligarchy in labor unions—unfortunate but characteristic, and perhaps even inevitable. This Article contests that wisdom and argues that labor law critically conditions this organizational tendency toward oligarchy.

What makes the dissolution of UHW particularly tragic is that not only was it a model, democratic union, but that it was also enormously successful in organizing new members. Recruiting new workers into unions is the labor movement's cardinal priority.[8] However, success in organizing new members is not easy to achieve. To do so, labor unions must make far-reaching strategic and organizational transformations—a process scholars call union "revitalization."[9] The first requirement is high levels of membership participation and commitment in the comprehen-

2. *Id.* (describing the reactions of labor historian Nelson Lichtenstein and labor researcher Kate Bronfenbrenner).

3. *Id.*; *see also* Paul Pringle, *Breakaway Union Could Prompt War of Attrition with SEIU*, L.A. TIMES, Feb. 2, 2009, http://www.latimes.com/news/local/la-me-union2-2009feb02,0,6911054.story ("[T]he UHW had been widely viewed as one of the most vibrant and successful locals in the SEIU . . . .").

4. Amber & Cutler, *supra* note 1.

5. *Id.*

6. *Id.* The ousted leaders of the UHW responded the day after the trusteeship by forming a new union, the National Union of Healthcare Workers (NUHW). *Id.* The NUHW has had notable success in winning over former UHW members in contested NLRB elections with the SEIU. *See* Michelle Amber, *Employees at Two Facilities Vote for NUHW After NLRB Finds Elections Should Be Held*, 196 Daily Lab. Rep. (BNA) A-2 (Oct. 14, 2009).

7. Amber & Cutler, *supra* note 1 (quoting Angela Glasper, a former UHW elected leader).

8. This has been the official stance of the AFL-CIO since 1995; the slow change of pace was the pivotal reason several major unions left the AFL-CIO to form the Change To Win Federation in 2005. *See* Joesph A. McCartin, *Reframing US Labour's Crisis: Reconsidering Structure, Strategy, and Vision*, 59 LABOUR/LE TRAVAIL 133, 133 (2007), *available at* http://www.historycooperative.org/journals/llt/59/mccartin.html (describing how the Change to Win Federation left the AFL-CIO due its failure to devote more resources to recruitment). As it is well known, union density, the proportion of the workforce that are union members, has declined dramatically over the past several decades. *See* Press Release, Bureau of Labor Statistics, U.S. Dep't of Labor, Union Members Summary (Jan. 22, 2010), *available at* http://www.bls.gov/news.release/union2.nr0.htm (showing that in 2009 the number of workers belonging to unions was 15.3 million and that union members accounted for 12.3 of employed wage and salary workers, down from 20.1 percent in 1983; also reporting that union density is 7.2 percent in the private sector and 37.4 percent in the public sector).

9. Kim Voss & Rachel Sherman, *Breaking the Iron Law of Oligarchy: Union Revitalization in the American Labor Movement*, 106 AM. J. SOC. 303, 304 (2000).

sive use of "rank-and-file intensive" tactics.[10] In their widely-cited study of 14 organizing unions, Kim Voss and Rachel Sherman showed that this kind of strategic orientation demands dramatic organizational changes as well: shifting resources into organizing requires reducing the size of the union's traditional, bureaucratic-professional staff and increasing the amount of voluntary self-representation of members within the workplace.[11] Coincidentally (or not), the UHW was almost certainly one of a handful of "fully" revitalized labor unions highlighted as exemplars in Voss and Sherman's study.[12]

The UHW's combination of robust union democracy with an aggressive organizing posture seems to have been far from accidental. Eliciting high levels of membership participation requires that members have a voice in the decisions that make such heavy demands on their time and attention.[13] Downsizing the union's traditional professional staff can catalyze resistance from those on whom oligarchic leaders most depend.[14] Most important, a mobilized rank-and-file can not only better organize new members, it can also organize the opposition that removes the old guard from office.[15] Entrenched union officials with parochial prerogatives and interests, fortified in relatively undemocratic unions, will therefore assiduously avoid adopting an organizing posture, even as this strategic choice sacrifices the larger interests of the labor movement. Consequently, lack of union democracy may be thwarting union revitalization.[16]

What is more, the successful case of UHW is very rare. Actual commitment to organizing still does not match the rhetoric in the vast

10. Kate Bronfenbrenner, *The Role of Union Strategies in NLRB Certification Elections*, 50 INDUS. & LAB. REL. REV. 195, 198 (1997). *See also* Kate Bronfenbrenner & Robert Hickey, *Changing to Organize: A National Assessment of Union Strategies*, *in* REBUILDING LABOR: ORGANIZING AND ORGANIZERS IN THE NEW UNION MOVEMENT 17, 19 (Ruth Milkman & Kim Voss eds., 2004); Steven H. Lopez, *Overcoming Legacies of Business Unionism: Why Grassroots Organizing Tactics Succeed*, *in* REBUILDING LABOR: ORGANIZING AND ORGANIZERS IN THE NEW UNION MOVEMENT 114, 115 (Ruth Milkman & Kim Voss eds., 2004).

11. Voss & Sherman, *supra* note 9, at 313, 315.

12. *Id.* at 315. In accordance with their research protocol, Voss and Sherman did not reveal the identities of the local unions they studied. Nevertheless, they conducted their research on "almost all the major Northern California locals affiliated with SEIU." *Id.* Given the UHW's reputation and the fact that it was based in Oakland, California, the conclusion that the UHW was one of Voss and Sherman's model "revitalized" locals seems inescapable.

13. Teresa Sharpe, *Union Democracy and Successful Campaigns: The Dynamics of Staff Authority and Worker Participation in an Organizing Union*, *in* REBUILDING LABOR: ORGANIZING AND ORGANIZERS IN THE NEW UNION MOVEMENT 62, 63 (Ruth Milkman & Kim Voss eds., 2004) (explaining that workers "are more likely to stay involved if they feel a sense [of] ownership over the direction and outcome of the organizing drive").

14. Voss & Sherman, *supra* note 9, at 321–22.

15. *See id.* at 322.

16. The necessity of internal organizational change also implies that current reform efforts, such as the Employee Free Choice Act, may be insufficient to spark the kind of organizing revival needed to restore the prospects of the labor movement. *See* William B. Gould IV, *The Employee Free Choice Act of 2009, Labor Law Reform, and What Can be Done About the Broken System of Labor-Management Relations Law in the United States*, 43 U.S.F. L. REV. 291, 299–300 (2008) (arguing that the Employee Free Choice Act is not the best answer to labor law reform).

majority of unions.[17] And for the most part, labor unions in the US remain at best nominally democratic, governed like "one-party states."[18] And even more dismaying, prevailing opinion says that the absence of union democracy is the normal, even inexorable, organizational outcome for labor unions.[19] Many scholars believe the struggle to preserve and enhance union democracy is futile and ineffectual.[20] These scholars rest their claims on a long tradition of empirical and theoretical scholarship that credits these outcomes to an "iron law of oligarchy."[21] This research argues that large-scale, membership-based organizations, such as labor unions, require the installation of bureaucracies in order to efficiently function.[22] But bureaucracy enhances the capacities of the officialdom while it simultaneously promotes membership powerlessness and inactivity.[23] As the balance of power changes, so does the governance of the organization, as the leaders seek to erode the democratic constraints that inhibit their personal interests from prevailing over the interests of the members and larger movement. According to the iron law of oligarchy, the prospect for union revitalization is grim.[24]

This Article will contend that the permanence of oligarchy in labor unions is far less assured than widely supposed. More precisely, it will argue not only that union democracy is possible,[25] but also that labor law

17. *See* Bronfenbrenner & Hickey, *supra* note 10, at 55 (explaining that "[e]ven the country's most successful unions" must organize on an unprecedented scale "if they are going to make any significant gains in union density"); DAN CLAWSON, THE NEXT UPSURGE: LABOR AND THE NEW SOCIAL MOVEMENTS 45 (2003) ("Unions talk about committing 30 percent of their resources to organizing . . . but almost no unions in fact do so."); Voss & Sherman, *supra* note 9, at 324 (stating that "[the amount of resources devoted to organizing] is so low it's almost embarrassing. . . . We're lucky if we're doing three [percent]") (alterations in original).

18. Samuel Estreicher, *Deregulating Union Democracy*, 21 J. LAB. RES. 247, 247 (2000). *See also* Stewart J. Schwab, *Union Raids, Union Democracy, and the Market for Union Control*, 1992 U. ILL. L. REV. 367, 368–70 (1992); Clyde W. Summers, *Democracy in a One-Party State: Perspectives From Landrum-Griffin*, 43 MD. L. REV. 93, 93–95 (1984).

19. Schwab, *supra* note 18, at 371.

20. Estreicher, *supra* note 18, at 247–48 ("The pursuit of union democracy is counterproductive . . . .").

21. ROBERT MICHELS, POLITICAL PARTIES: A SOCIOLOGICAL STUDY OF THE OLIGARCHICAL TENDENCIES OF MODERN DEMOCRACY 377–92 (Eden Paul & Cedar Paul trans., Dover Publications 1959) (1915).

22. *See id.* at 25–27.

23. *See id.* at 50–51, 60–61, 69–72, 80–82, 130; *see also* Summers, *supra* note 18, at 93–95 (summarizing the iron law of oligarchy).

24. Voss & Sherman, *supra* note 9, at 304.

25. Sustained union democracy has been shown to be possible in several other studies. *See, e.g.*, Margaret Levi, *Inducing Preferences Within Organizations: The Case of Unions*, *in* PREFERENCES AND SITUATIONS: POINTS OF INTERSECTION BETWEEN HISTORICAL AND RATIONAL CHOICE INSTITUTIONALISM 219, 228–36 (Ira Katznelson & Barry R. Weingast eds., 2005) [hereinafter Levi, *Inducing Preferences*]; SEYMOUR MARTIN LIPSET ET AL., UNION DEMOCRACY: THE INTERNAL POLITICS OF THE INTERNATIONAL TYPOGRAPHICAL UNION 15 (1956); JUDITH STEPAN-NORRIS & MAURICE ZEITLIN, LEFT OUT: REDS AND AMERICA'S INDUSTRIAL UNIONS 12, 161 (2003); Margaret Levi et al., *Union Democracy Reexamined* 37 POL. & SOC'Y 203, 208 (2009). In most of these studies, the factors sustaining union democracy are acknowledged to be unique to the particular cases.

plays a pivotal role in *conditioning* the possibility of union democracy.[26] By comparing the labor laws of the United States and Great Britain, where labor unions are significantly more democratic than their American counterparts, this Article concludes that foundational doctrines of ostensibly pro-union U.S. labor law forcefully inhibit the realization of union democracy within American labor unions. In particular, democracy in labor unions is less likely to survive or thrive where labor law, as in the U.S.: (1) grants to labor unions the exclusive right to represent a given group of workers, (2) establishes or provides support for professional and institutionalized procedures to resolve disputes in collective bargaining, and (3) prohibits or discourages the use of strikes or other forms of economic "self help" as alternative ways to address those disputes.

To reach this conclusion, the Article develops a game-theoretic model that explains how union democracy and labor law are related through two key mediating variables: the size of the union's bureaucracy and the extent of membership self-organization in the workplace—what I will call "workplace association."[27] Union leaders prefer oligarchy while the membership would rather prefer that the union be run democratically. Nevertheless, leaders may concede to union democracy when members can threaten to disrupt the normal functioning of the union. Union democracy then solves a commitment problem for union leaders, and thereby avoids disruptions that are costly to both sides. Critically, members' capacity to threaten disruption depends on the strength of their own autonomous forms of self-organization: workplace association. Workplace associations are groups of union member-workers who are organized in their workplaces to improve conditions of work, as distinguished from the union's full-time officers and employees—the union's bureaucracy—who typically perform the same tasks.[28]

---

26. Previous research on law and union democracy has either focused on the direct regulatory impact of the Labor-Management Reporting and Disclosure Act, *see, e.g.*, Alan Hyde, *Democracy in Collective Bargaining*, 93 YALE L.J. 793, 833–34 (1984); Edgar N. James, *Union Democracy and LMRDA: Autocracy and Insurgency in National Union Elections*, 13 HARV. C.R.-C.L. L. REV. 247, 248–49 (1978), or on the law governing elections for certification by the National Labor Relations Board, *see, e.g.*, Craig Becker, *Democracy in the Workplace: Union Representation Elections and Federal Labor Law*, 77 MINN. L. REV. 495, 497–98 (1993); Matthew T. Bodie, *Information and the Market for Union Representation*, 94 VA. L. REV. 1, 3–5 (2008); Kye D. Pawlenko, *Reevaluating Inter-Union Competition: A Proposal to Resurrect Rival Unionism*, 8 U. PA. J. LAB. & EMP. L. 651, 656–57 (2006); Schwab, *supra* note 18, at 367. Research has left unexplored the impact of major principles and doctrines of the National Labor Relations Act on union democracy.

27. While the combination of game theory with a case-study comparison of the U.S. and U.K. may strike the reader as odd, these two methodologies share an underlying concern with the specification of causal mechanisms in social-science explanations, as distinct from the more typical general, covering-law approach. For an introduction to this "analytic narrative" methodology, see generally Robert H. Bates et al., *Introduction*, *in* ANALYTIC NARRATIVES 3, 10–12 (1998). For the contrast between causal mechanisms and general laws, see Peter Hedström & Richard Swedberg, *Social Mechanisms: An Introductory Essay*, *in* SOCIAL MECHANISMS: AN ANALYTICAL APPROACH TO SOCIAL THEORY 1, 1–26 (Peter Hedström & Richard Swedberg eds., 1998).

28. The definition is refined below. *See infra* Part I.A.3.

However, workplace association and union bureaucracy are "substitutes" in production of the collective goods that unions and workers strive for, such as increases in wages and benefits, or improvements in conditions of work. As such, the strength of workplace association depends inversely on the size of the union's bureaucracy. In addition, a large bureaucracy is an important source of bargaining power with which the union leadership can deflect democratizing threats from rank-and-file insurgents. In sum, union democracy is more likely when workplace association is stronger, but a strong workplace association is only possible with a smaller union bureaucracy.

Labor law affects the relative strengths of union bureaucracy and workplace association and through them influences the outcome for union democracy. Exclusive representation blocks the ability of workplace associations to reach bargains with employers independent of the formally-recognized union. Exclusive representation also underpins powerful incentives to invest in bureaucracy by removing a free-rider between unions that would arise should they attempt to jointly represent employees in a given workplace. In addition, law that facilitates legally-supported, institutionalized procedures for resolving disputes with employers lowers the costs of union bureaucracy and therefore encourages its growth. Finally, excluded as legitimate and recognized bargaining agents and unable to make use of the technical and arcane procedures of institutionalized collective bargaining, workplace associations must turn to economic self help to address their grievances; but legal restrictions on the ability to strike or take other economic action foreclose precisely those alternatives.

Part I.A of the Article introduces the two cases and describes how unions in the U.S. and Britain vary across the three organizational dimensions we have introduced: democracy, bureaucracy, and workplace association. Part I.B considers and rejects several reasons why U.S. and U.K. unions might differ significantly across these variables. Part II of the Article explores in more detail the intuition for the game-theoretic explanation of union democracy and oligarchy I have just introduced. Part III then uses this model as an analytical framework for exploring the effects of the three areas of labor law highlighted above: exclusive representation, legally-institutionalized collective-bargaining procedures, and restrictions on economic-action alternatives to those procedures. This Part will demonstrate how labor law in the U.S. and Britain has differed dramatically in these three areas and how these legal differences help sustain distinct organizational configurations in American and British unions. The Conclusion offers a brief review of issues necessary for a future normative and policy debate in light of the theory's arguments.

The Appendix provides the formal version of the game-theoretic model upon which the analyses in Parts II and III are derived.

The primary goal of this Article is to develop an explanatory, or "positive," theory of the possibility of union democracy and show how particular configurations of central labor-law doctrines may frustrate this possibility. Consequently, this Article treats labor law as an "independent variable" and asks how it affects the relationship between union leaders and members, rather than asking how union leaders, members, employers, or other interested actors may have played a role in creating or preserving the U.S. or U.K. systems of labor law. In addition, while union democracy may be critically important for union organizing success, that proposition is still hotly debated. By centering on the questions of whether and how labor law conditions the possibility for union democracy, this Article does not directly engage that debate. Nevertheless, the importance of the question is not diminished: if it is impossible to sustain union democracy, what use is there in debating its implications for union organizing? Furthermore, although the Conclusion will address some of the normative issues that arise from the explanation, the focus of the Article remains on the positive analysis because establishing the link between labor law and union democracy is a necessary first step to the important normative questions that follow.

I. U.S. AND U.K. LABOR-UNION DIFFERENCES

Although the claims made by the iron law of oligarchy are still considered to be significant,[29] enough instances of union democracy have been observed to question its unexceptional universality.[30] The persistence of union democracy in British labor unions is one such instance. This Part describes the democratic differences between U.S. and British unions, as well as their distinctions in two other organizational characteristics: the size of their bureaucracies and the strength of their workplace associations. Apart from these illuminating contrasts, additional reasons make the U.S. and U.K. a good case comparison. Although their labor laws and the organization of their labor unions have differed greatly, they are otherwise similar in other factors which one would like to control. They share more broadly a common law legal heritage, as well as a similar tradition in labor movement organization and philosophy, based on craft unionism and voluntarism.[31] Thinking ahead about the consequences of union democracy, it is also worth mentioning at the outset that British unions have performed better on what are considered key

29. *See* Paul Osterman, *Overcoming Oligarchy: Culture and Agency in Social Movement Organizations*, 51 ADMIN. SCI. Q. 622, 623 (2006) ("A great deal of literature suggests that the iron law is a common outcome . . . .").

30. *See* LIPSET ET AL., *supra* note 25, at 404–05; Levi et al., *supra* note 25, at 222.

31. *See* William E. Forbath, *Courts, Constitutions, and Labor Politics in England and America: A Study of the Constitutive Power of Law*, 16 LAW & SOC. INQUIRY 1, 23 n.79, 24 n.80 (1991).

measures of union success. Chiefly, over the course of the post-World War II period, Britain has had both higher union density and greater coverage of employees working under a collective bargaining agreement.[32]

A. Comparing Labor Unions in the U.S. and U.K.

1. Union Democracy

The most important difference between U.S. and British unions lies in how democratic they are.[33] In a path-breaking comparative study of U.S. and British unions, David Edelstein and Malcolm Warner found significant differences in the level of democracy between British and American labor unions.[34] Their measure of democracy was the closeness of elections to fill vacancies in the union's top and next-to-top posts (typically the offices of president and secretary respectively).[35] An election was judged closer when the runner-up received a higher percentage of the victor's votes.[36] For Britain, Edelstein and Warner found that during the years 1949–1966 the mean closeness of elections was 53.9 percent for top vacant-post elections and 69.5 percent for next-to-top vacant-post elections.[37] For the same years in the United States, the results were 10.3 and 14.8 percent respectively.[38] Elections in British unions were therefore more competitive than those in U.S. unions.

---

32. *See* Michael Wallerstein & Bruce Western, *Unions In Decline? What Has Changed and Why*, 3 ANN. REV. POLI. SCI. 355, 358 tbl.1 (2000) (categorizing the U.K. as a "middle-density" country, with densities between 45.1 (1950), 56.3 (1980), and 41.3 percent (1992), and the U.S. as a "low-density" country, with densities between 28.4 (1950), 24.9 (1980), and 15.3 percent (1992); in 1990, contract coverage was 47 in the U.K. and 18 percent in the U.S.).

33. For methodological reasons, the definition of democracy used in this Article is a relatively formal one. If, as I hope to demonstrate, workplace associations help explain the likelihood of union democracy, one needs a definition of democracy that keeps those two concepts distinct. Nevertheless, if union democracy requires extensive member participation in workplace associations our overall notion of democracy is obviously more robust than the formal definition. By the same token, this idea of union democracy is not overly demanding: although it requires significant levels of member participation, it falls short of requiring direct democracy and recognizes that effective membership direction of the union may be indirect, e.g., via competitive elections for union office.

34. *See* J. DAVID EDELSTEIN & MALCOLM WARNER, COMPARATIVE UNION DEMOCRACY: ORGANISATION AND OPPOSITION IN BRITISH AND AMERICAN UNIONS 4 (Transaction Books 1979) (1975).

35. *Id.* at 95. One may question whether electoral opposition is a necessary element of democracy. For example, consider a unanimous faculty vote in favor of a single candidate for new department chair. Genuine consensus decisions such as these are democratic, perhaps even more democratic than a majority decision. However, the possibility of consensus and unanimity in large, heterogeneous organizations, such as labor unions seems remote.

36. The closeness of *vacant*-post elections is arguably a better measure of democracy than either turnover in posts or the closeness of elections involving incumbents. For instance, one could have frequent turnover without any opposition, such as a quick succession of union presidents who appoint their successors. In elections with incumbents, an incumbent who is repeatedly reelected does not necessarily indicate a frustration of majority will. The leader may be a genuinely effective and favored officer. The focus on vacant-post contests avoids the problems of either of these measures of democracy.

37. EDELSTEIN & WARNER, *supra* note 34, at 95.

38. *Id.*

Edelstein and Warner also compared the constitutions of British and American unions and found that those of British unions were more favorable to democracy than those of U.S. unions.[39] For instance, consider the rules governing a union's convention—the union's "legislature" and highest policy-making body. In British unions, conventions were held more frequently[40] and had fewer delegates, which facilitated their ability to act more as functioning decision-making bodies and less like large pep rallies for rubber-stamping predetermined back-room agreements.[41] British union constitutions also shielded conventions from the dominance of executive officers by restricting or prohibiting them from participating or acting as delegates.[42] In the U.S., field staff appointed by national officers are frequent participants and often delegates in union conventions.[43]

Similarly, constitutional rules governing the union's highest executive body were more democratic in Britain than in the U.S. In Britain, members of the union's executive council were much more likely to be elected by a subdivision of the union.[44] Electing executive councilors in this fashion gives them a reliable base of support from which to launch electoral challenges for higher union office.[45] By contrast, executive councilors in U.S. unions were more likely to be elected by the convention as a whole.[46] Union rules also influenced the power of the union's president.[47] Union presidents in Britain rarely had the power to appoint subordinate officials. By contrast, many in the U.S. had this power.[48]

British and American union constitutions were further distinguished by how they protected the civil liberties of union members and officers. British unions' constitutions were much less likely than U.S. unions' constitutions to have a disciplinary "blanket clause" expressing vague prohibitions against conduct "unbecoming to a union member" or acting "contrary to the interests of the union."[49] In British unions, final-appeals bodies considering matters of internal union discipline tended to exclude the participation of full-time officers, while officer participation was

39. *See id.* at 99, 101.
40. *Id.* at 105.
41. *Id.* at 104–05. This fact outweighs our more romantic considerations that having a large number of convention delegates better approximates an ideal of direct democracy.
42. *See id.* at 101.
43. *Id.* at 104.
44. *Id.* at 107.
45. Sara Gamm, *The Election Base of National Union Executive Boards*, 32 INDUS. & LAB. REL. REV. 295, 295 (1979).
46. EDELSTEIN & WARNER, *supra* note 34, at 107.
47. *See id.* at 99, 101.
48. *Id.* at 101.
49. *Id.* at 109.

more common in US unions.[50] Constitutional rules differed in many other relevant aspects as well.[51]

Perhaps most interestingly, Edelstein and Warner's study sought to show that a union's constitutional rules were in fact related to the competitiveness of elections for top officers.[52] Indeed, even within the sample of U.S. unions, they found that elections were closer when, among other rules, the time between conventions was shorter, the percentage of the union's executive council elected (or appointed) by subdivision was greater, the highest appeal body was independent from national officers, and the union's president was elected by direct vote rather than by the convention.[53] In short, not only do constitutional rules give the appearance of more democracy in British unions, but those rules actually matter for the competitiveness of union elections.[54]

2. Union Bureaucracy

Another striking difference between British and American unions is found in the size of their administrative bureaucracies. In Britain, union bureaucracies are notably smaller than in the U.S. A 1961 study of British trade unions found an average of one full-time union employee for every 1400 union members.[55] In contrast, contemporary estimates put the employee-member ratio in the U.S. between 1 to 273[56] and 1 to 300.[57] Compared with labor unions in other countries, the size of union bu-

50. *Id.* at 110.

51. For Edelstein and Warner's full discussion of overall British-American union differences, see generally *id.* at 87–114.

52. For their specification of mechanisms linking rules and election outcomes, see generally *id.* at 63–82.

53. *Id.* at 142–47. The claim that rules matter for election competitiveness still holds in Britain, but in a more complicated way because of the part of the heterogeneous British sample on which the comparative U.S. hypothesis was tested. *Id.* at 184–86.

54. The importance of rules, or more broadly institutions, is a significant but underappreciated point. In most discussions, the issue of union democracy is reduced to a simple agency dilemma, with the problem inhering in the difficulty union members have in monitoring their leaders. *See, e.g.*, Estreicher, *supra* note 18, at 248–50; Schwab, *supra* note 18, at 379–81. At the very least, however, the problem of monitoring is insufficient to address the issue of union democracy. Consider the election of executive councilors on an at-large versus regional basis. Electing executive councilors on a regional basis does not improve democracy by giving union members more information, but by exploiting the self-interest of executive officers for the members benefit (since when offices are open to competition candidates will be compelled to cater to members' preferences).

55. H.A. CLEGG, A.J. KILLICK & REX ADAMS, TRADE UNION OFFICERS: A STUDY OF FULL-TIME OFFICERS, BRANCH SECRETARIES AND SHOP STEWARDS IN BRITISH UNIONS 104 tbl.29 (1961). A full-time employee of the union includes full-time officers of the union—typically those who hold elective position in the union—as well as full-time members of the union's professional or administrative staff. *Id.* at 19–20.

56. SAMUEL LUBELL, FUTURE OF AMERICAN POLITICS 193 (Anchor Books 1955) (1952) (dividing the number of union members by the number of union employees).

57. RICHARD A. LESTER, AS UNIONS MATURE: AN ANALYSIS OF THE EVOLUTION OF AMERICAN UNIONISM 116 (1958).

reaucracies in Britain ranks low, while in the U.S., the size of union bureaucracies ranks high.[58]

The paucity of administrative resources in British unions relative to U.S. unions correlates with differences in financial strength as well. In Britain, average dues (or subscriptions in British terminology) are approximately 0.4 percent of the average manual earnings of full-time working males.[59] This places the dues rates of British unions below those of unions in most other European countries.[60] Only unions in countries such as France and Italy have had similarly low levels of dues.[61] Dues data are hard to come by in the United States—probably owing to the decentralized nature of dues policy in U.S. labor unions. But an estimate can be made that places dues in U.S. unions at around one percent of the average manual wage.[62] Thus, U.S. unions fair rather well financially compared to European unions, fitting somewhere at the bottom end of the high dues-rate category.[63]

3. Workplace Associations

British and American unions are also distinguished by the sophistication and autonomy of their shop-floor organizations, or what I will term more generally as workplace associations. Shop-floor or workplace

58. Jelle Visser, *In Search of Inclusive Unionism*, 18 BULL. COMP. LABOUR REL. 168, 168–69 tbl.24 (1990) reports data for most European trade union confederations for several years between the 1950s and 1980s). Staff to member ratios for some of the better-staffed confederations, with years (in parentheses) chosen closest to the U.S. and U.K. figures, are as follows: Netherlands, 1:485 (1952) and 1:478 (1970); Germany, 1:855 (1950) and 1:800 (1970); Sweden, 1:690 (1980). *Id.* These figures include only officers and employees of the national and confederal unions, and therefore exclude regional and local officers and employees. *Id.* Including the latter would undoubtedly make the actual densities higher. To give some sense of the possible discrepancy, compare the more inclusive figure we reported for Britain (1:1400, in 1961) to the more exclusive ones reported by Visser: 1:5000 (1950) and 1:2857 (1970). *Id.* If that discrepancy carries over, some of European labor union bureaucracies would be comparable and probably larger than those of the U.S. On the other hand, the highly decentralized structure of labor union organization in the U.S. raises union staff densities relative to European unions.

59. Paul Willman, *The Logic of "Market-Share" Trade Unionism: Is Membership Decline Inevitable?*, 20 INDUS. REL. J. 260, 268 (1989). The estimate is based on unions' per-capita income from subscriptions. *Id.*

60. Visser, *supra* note 58, at 166–67 tbl.23 (dividing European dues rates into three categories: "low" if less than 0.5% of gross wages, "medium" if greater than 0.5% but less than 1.0%, and "high" if greater than 1.0%).

61. *Id.*

62. Using monthly dues data from Charles W. Hickman, *Labor Organizations' Fees and Dues*, MONTHLY LAB. REV., May 1977, at 19, and average weekly earnings for all production workers from the Bureau of Labor Statistics' Current Employment Statistics, I estimate that dues were approximately 0.9% of the average manual wage in 1974. This estimation is very approximate. On the one hand, since only *ranges* for *minimum* dues were reported, the calculation was based on the low end of the range and did not account for the fact that dues are usually set higher than the minimum at the local level. On the other hand, the estimation does not account for the fact that union workers normally enjoy a wage premium over nonunion workers.

63. *Cf.* Visser, *supra* note 58, at 166–67 tbl.23 (placing high dues rate at one percent or greater of the gross wages). On the paradox of U.S. unions' strong financial and staff resources but weak organizing capacity, see Margaret Levi, *Organizing Power: The Prospects for an American Labor Movement*, 1 PERSP. ON POL. 45, 47 (2003).

associations are exemplified in Britain by its industrial relations tradition of shop-steward committees. Committees of shop stewards began to appear in particular British industries at the beginning of the twentieth century, and by the 1960s were a prominent feature across virtually all industries.[64]

Through such committees, stewards and union members in Britain play a much larger and more independent role in collective bargaining than they do in the U.S. In a 1994 study, 48.9 percent of full-time officers responded that the most common method of decision making about annual pay claims was for stewards and members to decide *alone*.[65] Another 23.3 percent of officers said that the most common method was for stewards and members to decide after consulting with a union officer, and only 1.1 percent of respondents said that the most common method was for a union officer to decide pay claims alone.[66] In the U.S., on the other hand, agreements with employers are negotiated by local union officials or committees of at least one or more full-time union representatives.[67] Typically, shop stewards in the U.S. are at most involved in low-level grievance handling, are carefully monitored by full-time officials, and more often act to relay communications downward from the union hierarchy to the membership rather than the opposite.[68]

The initiative for taking strike action also attests to the power and autonomy of British shop stewards. Between 1960 and 1964, strikes undertaken by members and shop stewards without authorization from union officials "accounted for nearly 95 per cent of all strikes in Great Britain, and 60 per cent of days lost from work because of strikes."[69] In the United States, thirty percent of all strikes in the U.S. were unauthorized and five percent of working time lost to strikes was a result of unauthorized strikes.[70] The terms "unauthorized" or "wildcat" strike carry the sting of opprobrium in the United States. Given the prominent role of

---

64. Michael Terry, *Shop Steward Development and Managerial Strategies*, *in* INDUSTRIAL RELATIONS IN GREAT BRITAIN 67, 67–68 (G.S. Bain ed., 1983).

65. JOHN KELLY & EDMUND HEERY, WORKING FOR THE UNION: BRITISH TRADE UNION OFFICERS 129 tbl.7.1 (1994).

66. *Id.* The other 26.7 percent of officers responded that the most common method of decision-making about annual pay claims was for stewards and officers to make a joint decision. *Id.*

67. Jack Steiber, *Unauthorized Strikes Under the American and British Industrial Relations System*, 6 BRIT. J. INDUS. REL. 232, 235 (1968); *see also* H.M. Douty, *Post-War Wage Bargaining in the United States*, 23 ECONOMICA 315, 318 n.3 (1956) (noting that in the British case multi-unionism generates shop-steward bargaining "divorced from the local union." In contrast, "[the shop steward in the U.S.] secures his representation at the plant level directly through his local union. This means in turn that the local union [rather than the shop steward] in the United States tends to be immediately and directly concerned with all aspects of the collective agreement . . . .").

68. Voss & Sherman, *supra* note 9, at 324 (quoting one participant in a partially revitalized local union as saying that the job of a steward "is pretty much to disseminate information and maybe observe if there's [*sic*] contract violations"). Shop stewards take a much more active role in fully revitalized locals. *Id.* at 313.

69. Steiber, *supra* note 67, at 232.

70. *Id.* at 234–35.

shop stewards in Britain, however, the occurrence of an unauthorized strike probably does not connote the same degree of dysfunction.

B. Possible Explanations for U.S.-U.K. Union Differences

What explains these stark differences in British and American labor unions, particularly in their levels of democracy? Are they related in a way that can explain why British unions are more democratic? Before presenting the Article's answer to these questions, it will be instructive to first consider a few alternative hypotheses.

The alternative explanation that no doubt leaps quickest to the reader's mind is some form of an American "exceptionalism" argument.[71] According to this view, one should not be surprised that British unions are more democratic than U.S. unions. After all, the American labor movement—with its conservative, bread-and-butter, "business-unionism" ideology—has lacked the kind of far-reaching socialist philosophy found in Europe that one might think would seek to truly empower workers democratically.[72] While intuitive, there is not much to support this American exceptionalism argument.[73] Without question, European labor movements sought to achieve *political* democracy by, among other things, extending the franchise.[74] Whether they maintained democracy within their own political parties and trade unions is a much more doubtful question. Indeed, Robert Michels formulated his "iron law of oligarchy" theory based on his analysis of the European labor movement and of the German Social Democratic Party in particular.[75] Moreover, this disjuncture between democratic goals and internal oligarchy may not have been as contradictory as it seems. Even early socialists were quick to defend the need for authority against the more democratic elements within their movement, lest internal infighting, as the argument went, became an obstacle in the struggle against employers and the state.[76]

A comparison of British and other European trade unions bears this reasoning out. European trade unions share many of the same democratic

71. The *locus classicus* of the American exceptionalism argument is WERNER SOMBART, WHY IS THERE NO SOCIALISM IN THE UNITED STATES? (Patricia M. Hocking & C.T. Husbands trans., Macmillan 1976) (1906).

72. For an argument that explains the conservative orientation of the American labor movement as historically contingent on the structure of law and legal institutions, see generally Forbath, *supra* note 31, and WILLIAM E. FORBATH, LAW AND THE SHAPING OF THE AMERICAN LABOR MOVEMENT 2–3 (1991).

73. For reconsiderations of the idea of American exceptionalism generally, see Sean Wilentz, *Against Exceptionalism: Class Consciousness and the American Labor Movement, 1790-1920*, 26 INT'L LAB. & WORKING CLASS HIST. 1, 3 (1984), and Aristide R. Zolberg, *How Many Exceptionalisms?*, *in* WORKING-CLASS FORMATION: NINETEENTH-CENTURY PATTERNS IN WESTERN EUROPE AND THE UNITED STATES 397, 398–99 (Ira Katznelson & Aristide R. Zolberg eds., 1986).

74. *See* Zolberg, *supra* note 73, at 414–15.

75. *See generally* MICHELS, *supra* note 21.

76. *See, e.g.*, Friedrich Engels, *On Authority*, *in* THE MARX-ENGELS READER 730, 731–32 (Robert C. Tucker ed., 2nd ed. 1978).

limitations that are evident in U.S. trade unions: absence of open and contested elections, conventions that meet too infrequently and are dominated by officials, and executive committees that lack independence.[77] The same holds true for the other aspects of labor organization as well. As mentioned, British union bureaucracies are remarkably small by European standards;[78] only French unions are in contention for similarly paltry amounts of administrative muscle.[79] While some "exceptional" factor may account for the relatively high bureaucratic density of U.S. unions (for instance, U.S. labor unions bargain for and administer health and pension benefits that are provided to Europeans through their governments),[80] this does not appear to make bureaucracies any larger in U.S. unions than in continental European unions. Moreover, this factor cannot account for the conspicuous bureaucratic gap found between British and other European labor unions. Finally, no other European country has the form of workplace associations found in Britain.[81] While Germany is famous for its works councils, these labor-organization counterparts are severely curtailed in their bargaining agendas and lack the autonomy and power possessed by British shop-stewards committees.[82] In terms of union democracy, union bureaucracy, and workplace association it is British unions, not American unions, that appear "exceptional" when measured against the rest of Europe. One is almost tempted to endorse a British version of the "exceptionalism" thesis.

Other alternative explanations can likewise be dismissed. First, scholars have criticized weaknesses in the Labor-Management Reporting and Disclosure Act (LMRDA), which seeks to guarantee certain democratic minima in labor unions, for being responsible for the dearth of democracy in U.S. labor unions.[83] But this cannot explain the union-democracy difference between the two countries because, whatever the LMRDA's shortcomings, there is *no* equivalent regulatory approach in

77. ANTHONY CAREW, DEMOCRACY AND GOVERNMENT IN EUROPEAN TRADE UNIONS 195–96 (1976) (concluding that "it seems fair to characterise the continental unions studied as being more dominated by officialdom than is usually the case in Britain").

78. *See* Visser, *supra* note 58, at 168–69 tbl.24, 170; *see also* KELLY & HEERY, *supra* note 65, at 37 (finding the ratio of full-time officers—as distinct from all full-time employees—to members in British unions is below the Western European average).

79. *See* Visser, *supra* note 58, at 170–71.

80. NELSON LICHTENSTEIN, STATE OF THE UNION: A CENTURY OF AMERICAN LABOR 142–44 (2002). By failing to distinguish union officers from all full-time union staff, Lichtenstein may overstate the degree to which the bureaucratic densities of U.S. unions exceed those of European unions. *See* Willman, *supra* note 59, at 268.

81. *See* CAREW, *supra* note 77, at 196–97.

82. *Id.* at 197 (describing works councils as "essentially non-union, even anti-union, bodies based on an assumed community of interests between employers and workers, and providing little scope for the pursuit of workers' sectional interests").

83. *See generally* Hyde, *supra* note 26, at 807 ("[The LMRDA] . . . said nothing by its terms about collective bargaining, but this very silence was cause for concern to some leading figures of the period who wanted to make sure that full democracy for dues increases, election of officers, and internal by-laws did not infect the collective bargaining process."); James, *supra* note 26, at 248, 251.

Britain. Second, the structure of American unions is staggeringly fragmented relative to European unions, and this has been offered as a reason for the surprisingly large bureaucratic sizes of U.S. unions.[84] However, the degree of centralization in U.S. and British unions is actually a point of similarity when compared to other European unions.[85] Furthermore, given the strong association of localism with democracy in the U.S., one would think that decentralization would enhance democracy in American labor unions. Third, the democratic differences of U.S. and British unions reported by Edelstein and Warner are national in character and consequently abstract away from important differences in industrial or occupational characteristics.[86] We can therefore most likely rule out explanations that attribute democratic variation to such factors.

Finally, might the much larger population size of the U.S. account for any organizational differences? The staff-member densities relied on above account for differences in union membership sizes across countries because bureaucracy is measured as a ratio of the number of full-time union officers and employees to the number of union members. In any case, we should expect bureaucratic densities to be smaller when union memberships are larger because there are administrative economies of scale.[87]

If these explanations cannot account for the difference in union democracy between the U.S. and Britain, what can?

II. A THEORY OF UNION DEMOCRACY

This Part provides the intuition for the game-theoretic explanation of union democracy presented in the Appendix. This explanation understands union democracy as a resolution to a commitment problem between union leaders and members, both of whom prefer the institutionalized procedures of union democracy to the disruption that members can create when they are sufficiently self-organized and therefore able to express their opposition to leadership policy. However, because union leaders would rather be unencumbered by union democracy, they will

84. *See* LICHTENSTEIN, *supra* note 80, at 142.

85. When seventeen countries are ranked by the degree to which wage bargaining is centralized, one finds that Britain and the U.S. are among the least centralized, ranking at 12 and 16, respectively. Lars Calmfors & John Driffill, *Centralization of Wage Bargaining*, 3 ECON. POL'Y 16, 18 tbl.1 (1988).

86. The British sample included thirty-one national unions that consisted of 203,000 members from both blue and white-collar occupations, while the U.S. sample included fifty-one national unions that consisted of 268,000 members, also from both blue and white-collar occupations. EDELSTEIN & WARNER, *supra* note 34, at 87–91.

87. *See* Willman, *supra* note 59, at 268 (reporting evidence that per capita administrative costs are negatively correlated with the size of the union's membership in British labor unions). Nevertheless, the highly decentralized U.S. labor movement often means that membership sizes are smaller than European labor unions. *See* EDELSTEIN & WARNER, *supra* note 34, at 88–90 tbls.4.1 & 4.2 (comparing roughly similar total union membership sizes for the U.S. and Britain despite a significantly larger number of U.S. unions).

only abide by democratic restraints when the threat of disruption is strong enough. The strength of workplace association matters, therefore, and this factor depends inversely on the size of the union's bureaucracy because they are substitutes in the production of collective goods union members care about. I will explain at some length why workplace association is a particularly important form of member self-organization. The causal relationships between union democracy, bureaucracy, and workplace association are summarized in Figure 1. We first address the positive relationship between the strength of workplace association and the likelihood of union democracy: What are workplace associations and what is their role in the democratization of labor unions?

A. Union Democracy and Workplace Associations

1. Secondary and Workplace Associations: Two Examples

To begin our discussion about the relationship between workplace association and union democracy, consider the following two historical examples. Beginning in the early 1960s, serious conflict emerged within the United Mine Workers of America (UMWA) between union leaders and members.[88] Democracy within the UMWA was moribund as the union's administration and collective-bargaining were centralized under the authority of its president.[89] After the contract settlement of 1964 between the UMWA and the mine operators, eighteen thousand miners struck for eighteen days to protest the agreement.[90] According to Paul Clark's account, "This demonstration took place not in protest of the basic wage increase of two dollars a day negotiated by the UMW[A] leadership, but rather in opposition to the lack of fringe benefits, pension increases, and improvement in health and safety provisions in the final settlement."[91] Incipient protests such as these contributed to the emergence of a rank-and-file movement—the Miners for Democracy—that succeeded in electing a reform slate of union officers.[92] Significant, pro-democratic changes were made to the UMWA constitution in 1973 and 1976.[93] Unfortunately, this movement for democracy was unable to sustain itself. In 1979, a new administration amended the constitution in ways that undermined the earlier democratic reforms.[94] Dismayed by the decline of the democratic reform movement in the UMWA, reform-

88. PAUL F. CLARK, THE MINERS' FIGHT FOR DEMOCRACY: ARNOLD MILLER AND THE REFORM OF THE UNITED MINE WORKERS 22–23 (1981).
89. *Id.* at 15–22.
90. *Id.* at 22–23.
91. *Id.* at 23.
92. *Id.* at 26–31.
93. *Id.* at 42–43, 85–89.
94. *Id.* at 143–45.

minded former officers concluded that, despite their efforts, "it's all right back where we started."[95]

Contrast this chronology of the UMWA with that of the British Amalgamated Society of Engineers (ASE).[96] Significant turmoil emerged within the ASE in the early decades of the twentieth century as the process of collective bargaining began to change.[97] Previously, wages and working conditions were set—often unilaterally—at the local level by union members within their respective shops or by district committees of the labor union.[98] As markets expanded and technology changed, employers sought to bargain with the union over terms of employment at a national, rather than local, level.[99] While national union officers sought an accommodation with these changes, this transformation provoked significant opposition from local union districts.[100] In one of the most significant examples, workers in Glasgow, Scotland, launched a strike in defiance of the ASE Executive Council's agreement with the employers on a demand for a wage reduction in the midst of an economic recession.[101] These events accumulated into "an amount of dissatisfaction and unrest unprecedented in the Society's history."[102]

At the same time that local-central conflict was growing within the ASE, committees of shop stewards began to appear throughout the union, often as more formalized expressions of the ASE's ancient craft tradition of shop-floor representation.[103] The sophistication of these shop-steward organizations was impressive: in 1919 they achieved formal recognition from employers, independent from the recognition earlier granted to the ASE.[104] These shop-steward organizations were also closely linked to democratic reform committees within the ASE that culminated in a number of thorough-going constitutional changes, including the creation of an annual National Committee, whose membership was limited to rank-and-file members and which was responsible for setting policy for the union's executive officers.[105] Such reforms helped ensure rank-and-file control over union policy and administration and reduced conflict between local and central organizations of the union. However, whereas

95. *Id.* at 155.
96. JAMES B. JEFFERYS, THE STORY OF THE ENGINEERS: 1800-1945 (1946).
97. *Id.* at 167–69.
98. *Id.* at 12, 98–100, 167; Jonathan Zeitlin, *The Triumph of Adversarial Bargaining: Industrial Relations in British Engineering, 1880-1939*, 18 POL. & SOC'Y 405, 412–13 (1990)
99. JEFFERYS, *supra* note 96, at 167; Zeitlin, *supra* note 98, at 410, 413.
100. JEFFERYS, *supra* note 96, at 167–69.
101. *Id.* at 167.
102. *Id.*; *see also id.* at 169–71.
103. *Id.* at 165–66, 181–86.
104. *Id.* at 185–86 .
105. *Id.* at 169, 193.

reform in the UMWA faltered, democratic change in the ASE proved durable.[106]

What lead to attempts at democratic reform in each union, despite the different outcomes? In each case, organizations that were autonomous from the official union provided a critical counterbalance to the authority of incumbent union officials. In the UMWA this was the Miners for Democracy. In the ASE, it was the shop-steward committees. In democratic theory such organizations are called "secondary associations." Secondary associations are best defined in the pioneering study of union democracy by Seymour Lipset and his colleagues as "organized or structured subgroups which while maintaining a basic loyalty to the larger organization constitute relatively independent and autonomous centers of power within the organization."[107] The definition, of course, is analogous to that used in theories of civil society and civic associations; the acknowledgement to Tocqueville is explicit.[108] The existence of secondary associations within an organization enhances the potential for democracy in several ways: they serve as arenas in which new ideas are generated, as alternative networks of communications, as training grounds for new leaders, as a means of encouraging participation in the larger union arena, and as bases of opposition to the central union authority.[109]

Not only must secondary associations be independent of the central union organization, but they must also be able to exercise some considerable form of power against it as well.[110] In both the UMWA and the ASE, the ability to carry out independent strike action without the authorization of union officials was an important example and demonstration of such power. Indeed, such "wildcat" or "unauthorized" strikes—while often protests against conditions of employment—are frequently demonstrations against the union's leadership as well.[111] Such actions can be highly disruptive, even catastrophic, to the union, and are therefore a potent means of sanctioning the union leadership.[112] Not only may they pose the loss of dues revenues, but they also disrupt established bargaining relationships, cause employers to question union leaders'

106. *See, e.g.*, EDELSTEIN & WARNER, *supra* note 34, at 263–318 (describing persistence of democracy within the ASE, later the AUEW).

107. LIPSET ET AL., *supra* note 25, at 15.

108. *Id.* at 73–76. Tocqueville's study remains the *locus classicus* of the discussion of civil society and civic associations. *See generally* ALEXIS DE TOCQUEVILLE, DEMOCRACY IN AMERICA (Henry Reeve trans., The Legal Classics Library 1988) (1835).

109. LIPSET ET AL., *supra* note 25, at 80.

110. *Id.*

111. *See* Complete Auto Transit, Inc. v. Reis, 451 U.S. 401, 403 (1981) (acknowledging that union-member respondents "commenced a wildcat strike, because they believed that 'the union was not properly representing them'").

112. William B. Gould, *The Status of Unauthorized and "Wildcat" Strikes Under the National Labor Relations Act*, 52 CORNELL L. REV. 672, 672 (1967) (quoting Secretary of Labor, W. Willard Wirtz, as describing the threat of wildcat strikes as "very, very dangerous for collective bargaining").

ability to keep their promises, and threaten the leadership's legitimacy. Under certain conditions, secessions of smaller union organizations from the larger body are another example of a similar disruptive mechanism.[113]

Nevertheless, while the democratic reforms in the ASE proved enduring, those in the UMWA did not. What explains the differing outcomes? The divergence can be attributed to the faltering of secondary association in the UMWA and to its tenacity in the ASE. In the UMWA, secondary association took the form of a single-issue organization based explicitly around the goal of union democracy. Tragically, once its purpose was thought to have been served, the rationale for its continued existence came into question. Indeed, even successful reform candidates at all levels of the UMWA united in calling an end to the Miners for Democracy once they took office: "The need for MFD has ceased to exist and we now must devote our time to rejuvenating the union that we all want to serve and must improve."[114] However, without a secondary association to *sustain* union democracy, the reform struggle eventually failed. By contrast, secondary associations in the ASE were rooted in the workplace organizations of shop-steward committees. Based in the struggle with employers for gains in the workplace, the rationale for shop-steward committees survived the movement for democratic reforms as did the system of shop stewards itself.

The more general conclusion one draws from the different outcomes in the UMWA and the ASE is that secondary associations are more likely to form and persist when they address employees' primary workplace interests, such as improvements in wages and working conditions. As has been frequently recognized, union democracy and matters of internal union governance are often much less salient to them than these first-order goods that unionization delivers.[115] Indeed, even the partial success of the Miners for Democracy is a notable exception in an American labor movement that could be characterized by its high number of largely marginal and failed democratic-reform organizations.[116] However, workplace associations do not face this "salience" obstacle, since they form precisely to improve and defend conditions of employment in the workplace.[117] Their role as democracy-supportive secondary associations is

113. STEPAN-NORRIS & ZEITLIN, *supra* note 25, at 68–69, 71–72 (showing that among unions founded in the New Deal era under the Congress of Industrial Organizations, "many more of the . . . unions born in a workers' insurgency than [those unions] led into the CIO by their top officers had organized factions and were highly democratic").

114. CLARK, *supra* note 88, at 34–35.

115. Estreicher, *supra* note 18, at 251.

116. See EDELSTEIN & WARNER, *supra* note 34, at 197, for a long list of such reform groups. Other cases of partial success, such as the Teamsters for a Democratic Union, owe their victories to significant government support. *See* Estreicher, *supra* note 18, at 251.

117. In addition, workplace associations are more likely to form than associations that fulfill workers primary non-work goals. This is because workplace gains present a more likely basis of shared interests. For instance, while workers may have many alternative outlets for satisfying their

incidental and unintended. Thus while secondary associations can take a number of forms within labor unions, *workplace* associations are more likely to form and succeed as a particular case of secondary associations. We reach the paradoxical conclusion—inferred from the civic-associations literature as well—that successful secondary associations are more likely *not* to have democracy as their primary *raison d'etre.*[118]

I have tended to use the term "workplace associations" to refer to this particular kind of secondary association rather than the more straightforward "shop-steward committees" because, as we have seen in the U.S. case, shop stewards do not always have the responsibilities and autonomy required to act as genuine secondary associations.[119] Furthermore, we can think of autonomous forms of workplace association other than shop-steward committees. For instance, sociologists and organizational theorists have long been fascinated with the nature and functioning of "informal work groups."[120] And in analyzing the legal standing of "employee caucuses" in nonunion work settings, Alan Hyde provides examples of some other forms of workplace associations: "(1) unorganized networking and griping; (2) internal pressure groups that form in protest of *ad hoc* [management] decisions; and (3) 'identity' groups like women's, Black, Latino, Asian, or gay and lesbian caucuses."[121]

2. Union Democracy and Credible Commitments

We have highlighted the capacity of wildcat strikes and union successions to create disruptions. But disruption by itself does not induce union democracy.[122] So how does the power of workplace associations actually contribute to democratic change within a union? Moreover, how did protests over health and safety in the UMWA and over wage reductions in the ASE lead to struggles for democratic reform? More generally, how were disputes over *policy* transformed into changes in constitutional *principle*?

To answer these questions, I propose we understand union democratization as a resolution to a commitment problem between union leaders and union members.[123] When the policy preferences of union leaders and

---

social needs, they have only each other when searching for ways to improve conditions on the shop floor. *See* Estreicher, *supra* note 18, at 252.

118. On the civic-associations inference, see ROBERT D. PUTNAM ET AL., MAKING DEMOCRACY WORK: CIVIC TRADITIONS IN MODERN ITALY 91–92 (1993).

119. *See supra* text accompanying notes 64–70.

120. *See, e.g.*, Fred E. Katz, *Explaining Informal Work Groups in Complex Organizations: The Case for Autonomy in Structure*, 10 ADMIN. SCI. Q. 204, 220 (1965).

121. Alan Hyde, *Employee Caucus: A Key Institution in the Emerging System of Employment Law*, 69 CHI.-KENT L. REV. 149, 150 n.3 (1993).

122. For the capacity to create disruption as a general form of power distinct from institutions such as democracy, see FRANCES FOX PIVEN & RICHARD A. CLOWARD, POOR PEOPLE'S MOVEMENT: WHY THEY SUCCEED, HOW THEY FAIL 27 (Vinatage Books 1979) (1977).

123. For a similar analysis, see generally DARON ACEMOGLU & JAMES A. ROBINSON, ECONOMIC ORIGINS OF DICTATORSHIP AND DEMOCRACY (2006). *See also* Douglass C. North &

members diverge,[124] the threat of a wildcat strike or secession gives a self-organized membership a weapon with which to contest leadership decision-making in a relatively undemocratic labor union. Wanting to avoid such a costly and disruptive confrontation, leaders will want to make concessions in policy. But promises from a leadership who calls all of the shots are simply that: promises without any binding force. Union members will also prefer to have their policy preferences implemented without having to carry through on their strike or secession threat, which may entail lost wages and other costs for them as well. When an oligarchic leadership's promises are not credible, however, a wildcat strike or secession might be the only possible outcome.[125]

Democracy is a solution to this credibility problem. Democracy entails rules and procedures that make leadership promises more binding and policy-making more favorable to the preferences of union members.[126] To recapitulate our previous discussion, competitive elections present leaders with the possibility of being voted out of office; executive-council members that are elected by and accountable to specific subdivisions of the union are freer to oppose presidential authority; limitations on a parent union's power of trusteeship restricts its ability to remove leaders elected by the local's members; and so forth. Democratization is therefore a process of institutionalization that transforms the de facto power of workplace associations into the de jure power of union rules that more effectively transmits member preferences into policy.[127] Both leaders and members prefer this arrangement to the alternative: costly wildcat strikes, successions, or other forms of disruption.

However, the key point is that union leaders will submit to democratic procedures, which restrict their policy preferences, only when workplace association is sufficiently strong to pose a potent enough threat. If workplace associations are too weak to mount any real opposition, then oligarchy prevails and the union leadership implements its most preferred policy choice (in which case, the leadership's credibility is irrelevant). A third, intermediate outcome is also possible. If workplace association is at a middling level of strength and if the union leadership's promises are sufficiently credible, then the insurgent union members might be willing to forgo their strike threat and accept the leadership's concessions. This may indeed be a common outcome in U.S.

Barry R. Weingast, *Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England*, 49 J. ECON. HIST. 803, 804 (1989).

124. As in the "iron law of oligarchy" and agency theory, a misalignment of preferences between leaders and members is presumed. *See* Summers, *supra* note 18, at 93–95 (summarizing the iron law of oligarchy). There would hardly be a democratic dilemma without such a misalignment.

125. For a discussion of commitment problems in political democratization, see ACEMOGLU & ROBINSON, *supra* note 123, at 133–42.

126. Of course, union democracy would not fully eliminate agency costs. All that is necessary for the argument is that agency costs are lower under union democracy than under oligarchy.

127. ACEMOGLU & ROBINSON, *supra* note 123, at 173–81.

labor unions. For instance, labor law practitioners are familiar with the way that the threat of wildcat strikes can serve as an informal goad to encourage the union to process workers' grievances.[128] The lack of perfect credibility diminishes the value of the leaderships' concessions, but this is the best a weaker workplace association may be able to achieve. These three distinct union-governance outcomes—oligarchy, democracy, and oligarchy with concessions—and the conditions of workplace-associational strength and union-leadership credibility likely to satisfy them, are displayed in Table 1.[129]

B. The Role of Union Bureaucracy

The democratic outcome of the labor union therefore depends critically on the strength of workplace associations. But likelihood of union democracy and the strength of workplace associations both depend on the size of the union's bureaucracy (among other factors that shall be discussed). In our comparison of union bureaucracy and workplace association in the U.S. and Britain, we saw that while union bureaucracies were large in the U.S. and small in Britain, the opposite was the case for the strength of workplace association.[130] In Britain, shop-steward organizations are far more sophisticated, autonomous, and central to its industrial-relations system than they are in the U.S. These correlations raise the questions of whether bureaucracy and workplace association are "substitutes" and by extension whether the size of bureaucracy is inversely related to the likelihood of democracy. This section will argue that this is exactly the case.

In union democracy and revitalization studies, some negative relationship between union bureaucracy and union democracy or workplace association is often presumed, but the links articulating them are often unspecified.[131] The game-theoretic model helps identify two mechanisms linking a larger bureaucracy with a reduced likelihood of union democracy.[132] Returning attention to Figure 1, these relationships are visually summarized as the two arrows leading from union bureaucracy to workplace association and union democracy, respectively.[133] The first mechanism has a negative, indirect effect on union democracy: a larger union

128. *See* David E. Feller, *A General Theory of the Collective Bargaining Agreement*, 61 CAL. L. REV. 663, 751 (1973) ("[T]he grievance and arbitration machinery [must] operate with reasonable speed. . . . If it does not, and grievances accumulate, the system becomes unacceptable [to workers]. This, in turn, may lead to wildcat strikes or work-slowdowns.").

129. One can regard this Table as a simplified version of the equilibrium characterization of the union-democracy game found in the Appendix. *See infra* Appendix B.2.

130. *See* discussion *supra* Part I.A.2–3.

131. *See, e.g.*, STEPAN-NORRIS & ZEITLIN, *supra* note 25, at 161 (taking "union democracy" and "monolithic, bureaucratic unions" as mutually exclusive alternatives).

132. Another possible mechanism, recognized elsewhere, is patronage: a larger bureaucracy can buy off the opposition or incorporate those talented workers most able to lead workplace associations. *See* MICHELS, *supra* note 21, at 185–87.

133. *See infra* Figure 1.

bureaucracy "crowds out" the possible strength of workplace association.[134] The second mechanism has a negative direct effect on the likelihood that union democracy will emerge: a larger bureaucracy gives the union leadership greater "capacity" to make more attractive policy concessions and bargain its way out of a full, de jure democratization of the union.[135]

Before investigating these mechanisms in detail, we should provide a microfoundation for why union leaders seek to establish bureaucracies in the first place.[136] Bureaucracies represent significant and costly investments in professional and administrative staff. The union's staff performs work that benefits the union members: the bargaining and enforcement of contracts, the resolution of grievances, and the production of other collective goods. The question becomes, why would unaccountable union leaders ever make such costly investments to produce collective goods that benefit union members? One answer to this question is that union leaders care greatly about revenues.[137] The source of union revenues, however, is primarily dues—the monetary contributions that members make to the union. Dues in turn depend on the level of wages of union members. If union leaders care about revenue, they will therefore have substantial incentives to bargain wage increases and therefore to provide the administrative means to achieve them. Union leaders may have other motives for building bureaucracies, ranging from empire building to a genuine commitment to the union movement. Nevertheless, the assumption of revenue maximization is a useful and reasonable first approximation of union leader behavior.

Given these incentives, how does union bureaucracy affect workplace associations? The first, "crowding-out" mechanism works in the following way. Since workplace association is another way of producing the collective goods that unions deliver, bureaucracy and workplace association are *substitutes*. Indeed, the high number and relative sophistication of shop stewards in Britain are seen as necessary to replace the dearth of full-time professional staff.[138] Consequently, the more of the collective good the union produces, then the more likely that workplace association will run into decreasing returns. Wages, for instance, can only be raised so high before they threaten to put the employer out of

134. *Id.*

135. *Id.*

136. There is a large literature on labor union behavior, the engagement with which is beyond the scope of this Article. For an overview of that literature, see generally Henry S. Farber, *The Analysis of Union Behavior*, *in* HANDBOOK OF LABOR ECONOMICS, 1039–89 (Orley Ashenfelter & Richard Layard eds., 1986); Bruce E. Kaufman & Jorge Martinez-Vazquez, *Monopoly, Efficient Contract, and Median Voter Models of Union Wage Determination: A Critical Comparison*, 11 J. LAB. RES. 401 (1990).

137. Union revenues may be a source of personal perquisites for union leaders, or they may be used to pursue other goals.

138. KELLY & HEERY, *supra* note 65, at 119 (explaining that with "limited resources, British unions require lay activists to shoulder much of the burden of day-to-day representation").

business. In a large-bureaucracy environment that delivers high wages, workers will find that the costs of workplace association outweigh the potential gains for further wage increases.

A large bureaucracy that "delivers the goods" does not, unfortunately, mean that member preferences will be fully satisfied—even while it will still effectively obstruct the formation of workplace associations. Better union contracts may increase the scope for certain kinds of leadership rent seeking. For instance, higher wages can support higher dues payments, *higher* than necessary to pay for the larger bureaucratic effort. In this case, members could be made better off by lowering dues while holding the level of wages and bureaucracy constant. Large pension funds can generate personal income for union leaders, in different ways and in sometimes more or less legal forms.[139] The history of the Teamsters union is a prime example of the compatibility of favorable contracts with high costs in corruption.[140] In these examples, there would still be little point trying to form workplace associations in order extract greater gains in the workplace, and yet members are not receiving all of the gains the union wins from the employer.

In addition, the mix of collective goods the union negotiates may not be the one that union members prefer. Since union leaders have an interest in maximizing revenue, they may favor wage increases over other collective goods from which it may be more difficult to measure and extract monetary rents[141]—for example, recall the members' health-and-safety concerns in the case of the UMWA.[142] If wages are at least partial substitutes for other union goods, wage increases will tend to crowd out workplace associations even though the final contract is not the one union members would have chosen.

Finally, a strong collective-bargaining agreement that delivers gains in the workplace will crowd out workplace association, but leave unsatisfied issues that concern union members *outside* the direct employment relationship. The best example of this is the organization of new members, discussed in the introduction.[143] The union's bureaucratic effort may be allocated toward improving conditions of employment in a single plant or firm, edging out workplace associations, when members' inter-

139. Benefit funds generate income for trustees but can also serve as a source of money for all manner of nefarious schemes, from simple embezzlement to the use of funds in questionable ventures. ROBERT FITCH, SOLIDARITY FOR SALE: HOW CORRUPTION DESTROYED THE LABOR MOVEMENT AND UNDERMINED AMERICA'S PROMISE 25–28 (2006) (describing benefits and corruption associated with union benefit funds).

140. *See* LEVI, *Inducing Preferences*, *supra* note 25, at 231–32.

141. Edward E. Lawler, III & Edward Levin, *Union Officers' Perceptions of Members' Pay Preferences*, 21 INDUS. & LAB. REL. REV. 509, 515 (1968) (finding that "officers tend to greatly overestimate the members' desire for additional cash" relative to economic security benefits), *cited in* Schwab, *supra* note 18, at 383 n.73.

142. *See supra* text accompanying note 91.

143. *See supra* pp. 1–7.

ests might be better served by organizing the wider industry.[144] Members may also care about the union's legislative agenda or its role in politics more generally, about its relationship to other unions and the larger labor movement, or about strategy and tactics leading up to the negotiation of agreements. We could add to this a long list of noninstrumental or intrinsic benefits that union democracy could bring to union members.[145] A large bureaucracy that "delivers the goods" may therefore crowd out workplace associations even while still failing to fully satisfy union members' interests.

The second mechanism that links union bureaucracy with the reduced chance of union democracy is called the "capacity" effect.[146] As we saw previously, union leaders will want to avoid democratization by offering policy concessions to insurgent workers.[147] In sum, the capacity effect says that those policy concessions will be more attractive to union members when the union's bureaucracy is larger. If bureaucracy can produce a unit of the collective good for members at a lower cost than workplace association, then greater gains from efficiency will be associated with a larger bureaucracy.[148] An efficient bureaucracy may help union leaders more effectively pursue their own goals, distinct from those of the membership. But, by the same token, a more efficient bureaucracy's greater "capacity" can also more effectively achieve members' interests, should the threat of disruption provoke the leadership to make concessions. Thus, when bureaucracy is more efficient, the union leadership can more easily "compensate" union members for the privilege of retaining an oligarchy when a rank-and-file insurgency is threatened.[149] For example, a more efficient bureaucracy could bargain better health-and-safety rules in place of high wages. On the other hand, with an inefficient bureaucracy, the union's capacity for only modest wage gains can only be traded off for limited health-and-safety improvements. The success of these promised changes in policy of course depends on the credibility of the leaders; but the point is that a more efficient bu-

144. On union members' interests in organizing, see *supra* text pp. 5–6.

145. As "schools of democracy," democratic labor unions give union members the training and experience in a broader role of citizenship.

146. *See infra* Appendix B.2.

147. *See supra* Part II.A.2.

148. Will union bureaucracy be more efficient than workplace association? Certainly bureaucracy presents its own information problems and agency dilemmas. And when bureaucracy is used for patronage, nepotism, or empire building, a larger bureaucracy may indeed be associated with greater inefficiency. But abstracting from these latter effects, bureaucracy is probably at least relatively more efficient than workplace association. Whereas bureaucracy entails full-time specialization and hierarchical monitoring, workplace associations, whose members voluntarily contribute effort on a part-time basis, tend to lack those efficiency-promoting attributes. There is also a heuristic reason to assume efficient bureaucracy. One of the foundational assumptions of Michels' "iron law of oligarchy" was the bureaucracy was more efficient than other forms of organization. If it can therefore be demonstrated that union democracy is possible even when union bureaucracy is more efficient than workplace organization, so much the stronger for the theory. MICHELS, *supra* note 21, at 187.

149. *See id.* at 389.

reaucracy increases the value of concessions at any *given* level of credibility.

The capacity effect has an important implication that is relevant for a normative and policy discussion of union democracy. If union democracy is more likely when bureaucracy is less efficient, then there will be some welfare loss associated with the conditions necessary for union democracy. Indeed, British unions are sometimes portrayed this way: poor in staff and finances and with informal and imprecise collective agreements negotiated by lay shop stewards.[150] The question is then raised: would concessions produce a better outcome for workers than union democracy? After all, concessions are easier to sustain than democracy as the efficiency of bureaucracy increases. And concessions go some distance toward satisfying members' policy preferences, while at the same time allowing them to benefit from the efficiency gains of union bureaucracy. Offsetting the attractiveness of this alternative is the problem of ensuring the credibility of union leaders without union democracy, since this credibility is crucial to the effectiveness of the concessionary strategy. In addition, union leaders have no incentive to concede more than necessary to avoid a de jure democratization of the union. For these reasons, union members may indeed be willing to trade away some of the efficiency of bureaucracy for a reduction of agency costs under union democracy.

While bureaucracy threatens to undermine democracy and crowd out workplace association, there may indeed be limits to these effects. Dues and benefit funds may increase the opportunities for economic rents, but "shirking," or not working as hard as leaders could because of the difficulty in monitoring their efforts, is a real problem in some unions as well.[151] For instance, even if leaders were solely concerned with economic rents, the non-profit nature of unions restricts those incentives since leaders are unable to capture the full value of their efforts.[152] Furthermore, high wages may partially substitute for other collective goods that members care about, but by the same token, if they are only partial substitutes, members will still have some incentive to form associations to make further improvements in the workplace. However, the strength of workplace associations does not depend solely on the size of the un-

150. OTTO KAHN-FREUND, LABOUR AND THE LAW 199–200 (1972) (writing that the language of many collective agreements in Britain "is so vague that a court may have to hold them to be 'void for uncertainty'"); Steven Tolliday & Jonathan Zeitlin, *Shop-Floor Bargaining, Contract Unionism and Job Control: An Anglo-American Comparison*, *in* THE AUTOMOBILE INDUSTRY AND ITS WORKERS: BETWEEN FORDISM AND FLEXIBILITY 99, 106–07 (Steven Tolliday & Jonathan Zeitlin eds., 1987) (arguing that the frequent job actions and extreme decentralization of shop-steward organizations' in Britain dissipates collective power and exacerbates sectionalism, ultimately curtailing their abilities to challenge wage inequities and employment insecurities).

151. Schwab, *supra* note 18, at 395. This informational rent is the primary agency problem discussed by Schwab. It is perfectly possible for both economic and informational rent seeking to occur simultaneously.

152. *Id.* at 396.

ion's bureaucracy, but on other costs as well, as depicted in Figure 1.[153] For example, as collective entities, workplace associations depend on the resolution of a collective-action problem. More central to this Article's focus are the costs of workplace association that are erected by legal statutes and rulings. These legally-imposed costs of workplace association will now be discussed in Part III.

III. LABOR LAW AND UNION DEMOCRACY

This Part examines how labor law influences union bureaucracy and workplace association, and hence the prospects for union democracy. As mentioned in the Introduction, we will examine three different areas: (1) the rights of labor unions to exclusively represent a given group of workers, (2) legal establishment or support for professional and institutionalized procedures to resolve disputes in collective bargaining, and (3) legal prohibitions on the use of strike action—economic "self help"—as alternative means of addressing those disputes. As each area of labor law is discussed, the U.S. version of the law will be compared to its British counterpart. Because the data used to compare the different elements of British and American union organization are available only from the 1960s and 1970s, the legal analysis will largely be confined to a similar period. The significant changes in British labor law—which have taken place since the late 1970s—therefore remain outside the scope of the analysis.

These three areas are critically important components of labor law. But they may not be the only important components, and one should not presume the list to be exhaustive. Further, whether the presence or absence of one area depends on the presence or absence of another is a question that is not addressed, although it will be possible to infer such an interdependency from the subsequent analysis. Whether or not this is the case, the three areas nevertheless have distinct causal implications. Figure 2 expands upon Figure 1 to display schematically the causal relationships between labor law and union bureaucracy and workplace association, and thence union democracy.[154] Each mechanism will be elaborated in more detail as each area of labor law is discussed.

A. Exclusive Representation

Clyde Summers has called the principle of exclusive representation the "fundamental ordering principle which shapes American labor law and collective bargaining."[155] I would add that the principle of exclusive representation is also the fundamental ordering legal principle that influ-

153. *See infra* Figure 1.

154. *See infra* Figures 1 & 2.

155. Clyde W. Summers, *Exclusive Representation: A Comparative Inquiry into a "Unique" American Principle*, 20 COMP. LAB. L. & POL'Y J. 47, 47 (1998).

ences the fate of union democracy in American labor unions. As shown in Figure 2, exclusive representation has two momentous consequences for union democracy. First, exclusive representation powerfully influences the union leadership's incentives to invest in building a bureaucracy. Second, exclusive representation determines which agency of the workers—the union leaders and its bureaucracy or workplace association—is granted privileged access to the institutionalized procedures both under the National Labor Relations Act (NLRA) and arising from collective agreements.

1. Exclusive Representation and Union Bureaucracy

The principle of exclusive representation is articulated in section 9(a) of the NLRA:

> Representatives designated or selected for purposes of collective bargaining by the majority of the employees in a unit appropriate for such purposes, shall be the exclusive representatives of all employees in such unit for the purposes of collective bargaining in respect to rates of pay, wages, hours of employment, or other conditions of employment . . . .[156]

When a labor union is certified by the National Labor Relations Board (NLRB) under Section 9(a) as the exclusive representative, "the presence of a majority union precludes the employer from bargaining collectively with a minority union."[157] The Supreme Court has acknowledged in numerous instances that an employer must "treat with no other"[158] than the exclusive bargaining representative and that a minority union may only bargain with an employer in the absence of an exclusive bargaining representative.[159] In the well-known case of *J.I. Case Co. v. NLRB*,[160] the Supreme Court recognized that employers commit unfair labor practices, specifically by refusing to bargain collectively with chosen representatives (Section 8(a)(5))[161] and by interfering with protected

156. National Labor Relations Act (NLRA) § 9(a), 29 U.S.C. § 159(a) (2006).

157. Summers, *supra* note 155, at 47.

158. Virginian Ry. Co. v. Sys. Fed'n No. 40, 300 U.S. 515, 547–48 (1937) (exploring the ramifications of the Railway Labor Act). For the NLRA specifically, see NLRB v. Jones & Laughlin Steel Corp., 301 U.S. 1, 44 (1937).

159. *See, e.g.*, Int'l Ladies' Garment Workers' Union v. NLRB, 366 U.S. 731, 741 n.1 (1961) (Douglas, J., dissenting in part) ("[A]bsent an exclusive agency for bargaining created by a majority of workers, a minority union has standing to bargain for its members."); Consol. Edison Co. v. NLRB, 305 U.S. 197, 237 (1938) ("[T]here is nothing to show that the [noncertified union] has been superseded by any other selection by a majority of employees of the companies so as to create an exclusive agency for bargaining under the statute, and in the absence of such an exclusive agency the employees represented by the [noncertified union], even if they were a minority, clearly had the right to make their own choice.").

160. 321 U.S. 332 (1944).

161. National Labor Relations Act (NLRA) § 8(a)(5) ("It shall be an unfair labor practice for an employer to refuse to bargain collectively with the representatives of his employees . . . .").

employee rights (Section 8(a)(1)),[162] when they attempt to bargain in circumvention of a certified and exclusive bargaining representative.[163]

The effect of exclusive representation on incentives to invest in union bureaucracy works in the following way. When unions lack exclusive representation, unions are free to recruit members on an individual basis. Two or more unions may seek to recruit members in the same workplace. This scenario presents several dilemmas to unions. First, because of the collective nature of the union good, workers will join the union which offers the lowest dues, *ceteris paribus*, since they get the good regardless of their union affiliation. This competition for members places downward pressure on the level of dues. Second, because the union-provided good is a collective good, each union has an incentive to *reduce* the amount of effort it provides (the size of its bureaucracy)—the opposite outcome in the market for a private good. In contrast, when a union has exclusive representation, workers have no incentive to join an outside, competing union. The result is both smaller bureaucracies and lower levels of dues in the absence of exclusive representation.

These mechanisms go a long way toward explaining the administrative and financial differences between U.S. and British unions. The lack of exclusive representation has made an enormous impact on Britain's system of industrial relations. First, what is called multi-unionism, the presence of multiple unions within a given workplace or firm, often competing to represent similar categories of workers, is a prominent feature of British labor relations.[164] In the U.S., by contrast, similarly situated workers belong to the same bargaining unit represented by a single, exclusive union.[165] Second, rival unionism has a depressing effect on union dues in Britain. Even within Britain, unions have lower per capita dues income where competition for members is fiercer.[166] One sees this phenomenon in other European labor movements when competitive un-

---

162. National Labor Relations Act (NLRA) § 8(a)(1) ("It shall be an unfair labor practice for an employer to interfere with, restrain, or coerce employees in the exercise of the rights guaranteed in section 157 of this title . . . ."); *see also* § 7 ("Employees shall have the right to self-organization, to form, join, or assist labor organizations, to bargain collectively through representatives of their own choosing, and to engage in other concerted activities for the purpose of collective bargaining or other mutual aid or protection . . . .").

163. *See J.I. Case Co.*, 321 U.S. at 334, 339; *see also* Medo Photo Supply Corp. v. NLRB, 321 U.S. 678, 679–80 (1944). Both *J.I. Case Co.* and *Medo Photo Supply Corp.* involved the question whether the employer could bargain directly with employees in circumvention of the exclusive representative. The principle is virtually unquestioned where the employer attempts to bargain with a minority union.

164. *See* Robert Kilroy-Silk, *Royal Commission on Trade Unions and Employer Associations*, 22 INDUS. & LAB. REL. REV. 544, 551 (1969) (discussing a report of the Royal Commission on Trade Unions and Employer Associations which recommended a reduction in multi-unionism, which is present when two or more unions exist); KAHN-FREUND, *supra* note 150, at 85 (interunion disputes are "of considerable importance in British industry").

165. EDELSTEIN & WARNER, *supra* note 34, at 17–18 (contrasting the US with Britain and stating that "[i]n the United States, virtually all the manual workers in a given workplace belong to the same union, which is the sole bargaining agent for such employees").

166. Willman, *supra* note 59, at 268.

ionism prevails. Both France and Italy have had multiple union confederations that compete to recruit members in the same industries and occupations. Like Britain, they have also had smaller bureaucracies and lower dues rates than European labor movements without such competition.[167]

The effects of exclusive representation can therefore be quite substantial. As will shortly be examined, the legally-supported institutionalization of collective bargaining can lower the costs of bureaucracy, and thereby influence its growth. But whereas the costs of union bureaucracy affect its size at the margin, the absence of exclusive representation and the accompanying collective-action problem between unions undermines incentives to invest in bureaucracy regardless of its costs.

2. Exclusive Representation and Workplace Association

Exclusive representation not only affects the size of the union's bureaucracy, but the strength and potential effectiveness of workplace association as well. This is because the principle of exclusive representation not only precludes an employer from bargaining with a minority union, but it also prohibits an employer from bargaining directly with its employees. In addition, the courts have concluded that the NLRA does not protect employees' use of economic action to pressure the employer into bargaining with them in circumvention of the recognized union representative because direct bargaining would violate the exclusivity principle. In sum, this raises the costs of workplace association, making them less likely emerge or become a meaningful and vibrant alternative to achieving workplace gains.

In an early case, *Medo Photo Supply Corp. v. NLRB*,[168] the Supreme Court concluded that an employer may not bargain directly with the employees, in circumvention of the exclusive agent, even when the employees initiate the bargaining.[169] In that case, a labor union was recognized as the bargaining representative for a designated unit of workers.[170] Prior to negotiations, however, a majority of employees in the unit sought to negotiate wage increases without the union's intervention.[171] The Court held that the employer's bargaining violated Sections 8(a)(1) and 8(a)(5) of the NLRA, stating: "Bargaining carried on by the employer directly with the employees, whether a minority or majority . . . would be subver-

167. *See* Visser, *supra* note 58, at 166–67 tbl.23, 170–71. Italy is particularly revealing of the causal relationship between competition for members and union dues, since dues went up after contending confederations established more a cooperative relationship.

168. 321 U.S. 678 (1944).

169. *Id.* at 685.

170. *Id.* at 681.

171. *Id.*

sive of the mode of collective bargaining which the statute has ordained . . . ."[172]

Not only will an employer who bargains directly with employees be subject to unfair labor practice sanctions, but he or she is also free to terminate employees who seek to press the employer into negotiations without the participation of the exclusive bargaining agent through the use of economic sanctions. This conclusion was most famously stated in *Emporium Capwell Co. v. Western Addition Community Organization.*[173] In that case, employees who filed grievances with the union against the employer that alleged racial discrimination became dissatisfied at the pace of progress.[174] As a result, they sought to engage the employer directly and commenced picketing their place of employment; after a warning, the employer discharged them.[175]

The Court in *Emporium Capwell* first addressed the question whether the employees were merely attempting to present a grievance to, rather than bargain with, their employer.[176] If it is clear that the NLRA bars employers from bargaining with unions rival to the exclusive representative, the direct relationship between employer and employees raises more of a question. The remainder of Section 9(a), following the part quoted above, contains a proviso stating:

> *Provided*, That any individual employee or a group of employees shall have the right at any time to present grievances to their employer and to have such grievances adjusted, without the intervention of the bargaining representative, as long as the adjustment is not inconsistent with the terms of a collective-bargaining contract or agreement then in effect: *Provided further*, That the bargaining representative has been given opportunity to be present at such adjustment.[177]

However, despite this proviso, the Court affirmed the Board's finding that the employees' conduct was "no mere presentation of a grievance but nothing short of a demand that the [Company] bargain with the picketing employees for the entire group of minority employees."[178]

---

172. *Id.* at 684.

173. 420 U.S. 50 (1975). The ramifications of *Emporium Capwell* reach far beyond exclusive representation. For a contextual history of the case exploring the relationship between racial discrimination in employment, labor law, and the fate of New Deal liberalism, see generally Reuel E. Schiller, *The Emporium Capwell Case: Race, Labor Law, and the Crisis of Post-War Liberalism*, 25 BERKELEY J. EMP. & LAB. L. 129 (2004).

174. *Emporium Capwell*, 420 U.S. at 53–54.

175. *Id.* at 55–56.

176. *Id.* at 60–61.

177. National Labor Relations Act (NLRA) § 9(a), 29 U.S.C. § 159(a) (2006).

178. *Emporium Capwell*, 420 U.S. at 57 (quoting the Board's adoption of the Trial Examiners' findings and conclusions); *see also id.* at 60–61 (affirming the finding of Board).

It also noted that the intention of the proviso was to allow employees to present grievances to the employer without exposing the employer to liability for bargaining in circumvention of the exclusive bargaining representative.[179]

The second question raised in *Emporium Capwell* was whether the termination of the picketing employees violated their rights under the NLRA.[180] Section 7 of the NLRA grants employees rights "to engage in other concerted activities for the purpose of collective bargaining or other mutual aid or protection" and section 8(a)(1) makes it an unfair labor practice for an employer to interfere with the exercise of those rights.[181] However, the Court concluded that the terminations of the workers did not violate the Act because their attempt to circumvent the union and bargain directly with the employer was inconsistent with section 9(a)'s exclusivity principle.[182] The Court conjectured at length on the "fragmentation" that would happen with "separate" bargaining, and the diminution of the union's bargaining power that would accompany it.[183]

These decisions place workplace associations in an insuperable position. If workers are dissatisfied with their rights established under the collective-bargaining agreement or the procedural means for securing them, they may seek to engage the employer directly—using "unofficial" means such as economic sanctions—to bring the employer to negotiations. However, the use of such self-help alternatives does not receive the same level of protection that other concerted activities are given under the NLRA. Moreover, even if the use of economic action is effective and induces the employer to make concessions, the law ties the employer's hands. If the employer does bargain a separate agreement with the workplace association, the union can charge the employer with violating its duty to bargain. Both of these outcomes significantly obstruct the ability of workplace associations to achieve gains independent of the official union. For workplace association to be successful it must overcome not only the absence of statutory protection, but also the employer's fears of liability.

3. The Closed Shop as Alternative in Britain?

The British counterpart to the American rule of exclusive representation was very simple: no principle of exclusive representation ever existed during the relevant time period. Nevertheless, union-security agreements—such as the closed shop, which make membership in the

179. *Id.* at 61 n.12.
180. *Id.* at 71.
181. National Labor Relations Act (NLRA) §§ 7, 8(a)(1). For the text of these provisions, see *supra* note 162.
182. *Emporium Capwell*, 420 U.S. at 65–70.
183. *See id.*

union a condition of employment—were legal.[184] And if a closed-shop agreement requires that employees join only a single union, then the union has exclusive bargaining rights.[185] Furthermore, there are cases illustrating the use of the closed shop to inhibit workplace associations from directly bargaining with an employer.[186] Was the closed shop in Britain therefore an effective counterpart to exclusive representation in the U.S.? The answer is no, for the simple reason that only about two-fifths of union members were covered by closed-shop agreements in the 1960s.[187] By itself, this is not a trivial proportion. But it approaches nowhere near the degree of pervasiveness that exclusive representation operates compared to the United States, where virtually the only way unions achieved recognition was through an NLRB certified election that conferred exclusive-representation status where the union prevailed.[188]

B. Legally-Institutionalized Collective Bargaining

Inscribed in the very purposes of the Wagner Act is the goal of "removing certain recognized sources of industrial strife and unrest, by encouraging practices fundamental to the friendly adjustment of industrial disputes."[189] In accordance with these objectives, the NLRA, greatly facilitated by subsequent judicial support and elaboration, establishes opportunities to substitute legal, administrative, or arbitral process for economic self help at virtually every stage of the collective-bargaining process. These stages include: the organization of workers, recognition of the union, the employer's duty to bargain with the union, and interpretation and enforcement of the collective-bargaining agreement.[190] The contrast with British labor law in the post-war period could not be greater. Whereas the Wagner Act system of collective bargaining endowed un-

184. KAHN-FREUND, *supra* note 150, at 199–200. Union-security agreements come in a variety of forms. Broadly speaking, in American usage, a closed shop requires the employer to hire and keep in employment only members of the particular union. A union shop allows the employer to hire anyone, but requires the new hire to become a member of the union within a specified time period, such as thirty days. An agency shop requires employees to pay a "fair share" of representation fees to the union as a condition of employment without requiring or conferring full, formal membership rights. A maintenance-of-membership agreement requires those employees who were union members at a given date to maintain union membership as a condition of employment. In British usage, the closed shop can be used to refer to either a closed- or union-shop agreement, with the distinction sometimes made between the "pre-entry" closed shop and the "post-entry" closed shop. *Id.* at 198–99.

185. Union-security agreements need not be exclusive. For example, a closed-shop agreement could require that all employees simply become a member of one of several possible unions. Kahn-Fruend mentions this kind of agreement in Britain, but does not say how common such agreements are. *Id.* at 199.

186. *See, e.g.*, Morgan v. Fry, (1968) 2 Q.B. 710, 721–23 (involving employer dismissing member of a breakaway union in order to avoid strike trouble from the union with which the employer had an informal exclusive bargaining agreement), *cited in* KAHN-FREUND, *supra* note 150, at 203.

187. W. E. J. MCCARTHY, THE CLOSED SHOP IN BRITAIN 28 (1964).

188. CHARLES J. MORRIS, THE BLUE EAGLE AT WORK: RECLAIMING DEMOCRATIC RIGHTS IN THE AMERICAN WORKPLACE 88 (2005).

189. National Labor Relations Act (NLRA) § 1, 29 U.S.C. § 151 (2006).

190. *See id.* §§ 7, 8(a), 8(d), 9(a).

ions and workers with a set of "positive" rights, British labor law left trade unions and workers merely with "negative" liberties. British labor law famously became characterized as a system of "collective *laissez faire*."

The main consequence of the substitution of legal and orderly procedure for economic strife is to lower the costs of collective bargaining. At any one of these points in the collective-bargaining process a dispute with the employer is possible. Without a binding procedure to address that dispute, unions have only the threat of strike or other economic action to enforce their claims. Yet strikes are enormously costly. An opportunity to resolve disputes through an institutionalized procedure therefore presents unions and workers with a lower-cost alternative to economic action.[191]

However, the union leadership is in a better position to benefit from the substitution of process for self help than is workplace association. This is for two reasons. First, as the exclusive representative, the union has privileged access to these institutionalized alternatives. In addition the costs of learning and accommodating collective-bargaining procedures are lower for the union bureaucracy—with its comparative advantage in skill, knowledge, and specialization—than for associated union members on the shop floor. Thus, institutionalized collective-bargaining procedures are expected to lower the costs of union bureaucracy rather than those of workplace association. And lowering the costs of bureaucracy gives union leaders further incentive to expand these administrative apparatuses, with all the consequences for union democracy that were discussed in the previous Part.

1. Organization, Recognition, and Bargaining

The NLRA provides legal support to labor unions attempting to organize workers. Before a union has been recognized, employees have rights under Section 7 to "form, join, or assist labor organizations."[192] Section 8(a)(1) makes it an unfair labor practice for an employer to interfere with an employee's Section 7 rights.[193] And Section 8(a)(3) makes it an unfair labor practice for an employer to discourage membership in a labor organization by discrimination in hiring, tenure, or conditions of employment.[194] A voluminous amount of law has developed under these provisions in the context of labor union organizing. In Section 9, the NLRA establishes procedures for determining an appropriate bargaining unit and certifying an exclusive representative of that unit through a

191. For a fuller discussion of these "two logics" of union behavior, see generally Claus Offe & Helmut Wisenthal, *Two Logics of Collective Action: Theoretical Notes on Social Class and Organizational Form*, 1 POL. POWER & SOC. THEORY 67 (1980).

192. National Labor Relations Act (NLRA) § 7. For text, see *supra* note 162.

193. *Id.* § 8(a)(1).

194. *Id.* § 8(a)(3).

Board-directed representation election.[195] Once recognized, an employer has a duty to bargain with the union under Section 8(a)(5).[196]

Thus, in each of these collective-bargaining stages—organization, recognition, and bargaining—the law provides a procedural substitute for the need to mobilize a large number of workers for strike action. Labor unions and their supporters justly complain about the inadequacy of the available remedies to redress employer violations of unions' and employees' rights under Section 8. Nevertheless, unions' consistent recourse to the protections of the NLRA indicates that they are the preferred method of addressing disputes, whatever their shortcomings.[197]

In contrast to the United States, during the relevant period in British labor law there were no legally meaningful rights to organize or rights to labor union recognition. Although under the British conception of collective laissez faire labor unions were free to engage in a broad range of economic actions, this nevertheless left "a glaring contrast between the wide scope of this freedom and the absence of any legislation seeking to guarantee its exercise."[198] Thus, there was no legal protection against anti-union discrimination or the interference by employers in the establishment or functioning of labor unions.[199] There was no obligation for an employer to bargain with a union.[200] Ultimately, the only sanction trade unions could bring to bear on employers in order to enforce a claim to recognition or against acts of discrimination or interference was the one of economic action.[201] British labor law therefore failed to provide its unions with the same kind of cost-reducing procedural alternatives that the Wagner Act presented to unions in the U.S.

2. Interpretation and Enforcement of Collective Bargaining Agreements

Once a union has successfully organized, been recognized, and has concluded an agreement with the employer, U.S. labor law also lends its support to disputes arising from the interpretation and enforcement of these agreements. First, Section 301(a) of the Labor Management Rela-

195. *Id.* § 9.

196. *Id.* § 8(a)(5). The duty to bargain is subject to a good-faith standard as established in the NLRA section 8(d). Section 8(d) also requires that, where a collective-bargaining agreement is in effect, a party desiring a termination of modification of the agreement give written notice to the other party, offer to meet and confer, and continue the existing agreement for sixty days after giving such notice (or until the expiration of the contract if that occurs later) before resorting to economic action. *Id.* § 8(d)(1)–(4).

197. *See* MORRIS, *supra* note 188, at 83–86 (discussing how rapidly labor unions came to embrace Board procedures in the case of representation elections). The law's failure to fully protect the right to organize, thus making the ability to organize more costly, could explain why unions tend to allocate bureaucratic resources to servicing existing members rather than organizing new ones.

198. KAHN-FREUND, *supra* note 150, at 172.

199. *Id.*

200. *Id.* at 78.

201. *Id.* at 249.

tions Act (LMRA) makes collective-bargaining agreements enforceable in federal courts.[202] In the important decision of *Textile Workers Union v. Lincoln Mills*,[203] the Supreme Court upheld the constitutionality of section 301.[204] However, the Court also went further and declared that section 301 gave the federal courts the mandate to create a federal substantive law of collective bargaining.[205] The policy of substituting institutionalized procedures for economic strife could not have been stated more clearly:

> Plainly the agreement to arbitrate grievance disputes is the quid pro quo for an agreement not to strike. Viewed in this light, the legislation does more than confer jurisdiction in the federal courts over labor organizations. It expresses a federal policy that federal courts should enforce these agreements on behalf of or against labor organizations and that industrial peace can be best obtained only in that way.[206]

Such legal support gives unions the assurance that disputes can be effectively addressed through procedural means, without the need to mobilize and resort to economic compulsion and self-help. Further, this policy was forcefully elaborated in a series of subsequent decisions, known collectively as the *Steelworkers Trilogy*. In those decisions, the Supreme Court held that courts should enforce agreements to arbitrate regardless of the underlying merit of the dispute,[207] that the agreement should be interpreted to cover the dispute even where the scope of the agreement was ambiguous,[208] and that arbitration awards would be enforced as long as they drew their essence from the collective agreement.[209]

The contrast with British labor law is again instructive. For the most part, collective agreements were legally unenforceable because they lacked contractual intent.[210] This was as the parties desired, and is the

---

202. 29 U.S.C. § 185(a) (2006) ("Suits for violation of contracts between an employer and a labor organization representing employees in an industry affecting commerce as defined in this chapter, or between any such labor organizations, may be brought in any district court of the United States having jurisdiction of the parties, without respect to the amount in controversy or without regard to the citizenship of the parties."). Prior to the passage of the LMRA, some state courts had enforced collective agreements under various legal theories. *See, e.g.*, Goldman v. Cohen, 227 N.Y.S. 311, 312–14 (N.Y. App. Div. 1928), *cited in* Katherine Van Wezel Stone, *The Post-War Paradigm in American Labor Law*, 90 YALE L.J. 1509, 1520 (1981); *see also* Schlesinger v. Quinto, 194 N.Y.S. 401, 409–10 (N.Y. App. Div. 1922).

203. 353 U.S. 448 (1957).

204. *Id.* at 457–58.

205. *Id.* at 455–57.

206. *Id.* at 455.

207. United Steelworkers v. Am. Mfg. Co., 363 U.S. 564, 567–68 (1960).

208. United Steelworkers v. Warrior & Gulf Navigation Co., 363 U.S. 574, 580–81 (1960).

209. United Steelworkers v. Enter. Wheel & Car Corp., 363 U.S. 593, 597 (1960).

210. *See* KAHN-FREUND, *supra* note 150, at 132 (citing Ford Motor Co. v. Amalgamated Union of Eng'g & Foundry Workers, (1969) 2 Q.B. 303, 330–31 (holding that an agreement between Ford and a number of other unions could not be enforced because of the factual finding that the agreement lacked contractual intent)).

consequence of long tradition in British industrial relations.[211] There was an exception to this rule, which was that certain provisions of collective agreements could be incorporated into the terms of individual contracts of employment through a theory of prevailing custom and usage.[212] However, this theory only applied to the terms of the collective agreement that could and were intended to be terms of contracts of employment and that gave rise to rights and duties that could be enforced through law-of-contract remedies.[213] Hence, parts of collective agreements covering the making of employment contracts themselves were not applied, for example, agreements over job allocation or the promise to return to work after a strike. The theory also did not cover jointly created institutions, such as committees for whatever purpose or pension funds. Nor did it apply either to established collective bargaining rules or "peace obligations," known in the United States as no-strike agreements.[214]

More generally, British law establishes no rules promoting the use of legal or quasi-legal procedure to resolve disputes under the collective agreement. In the 1970s, British labor-law scholar Otto Kahn-Freund even felt justified in concluding that although British trade unions had achieved de facto recognition over a broad swath of industries, "[t]he law . . . had no share in the advancement of collective bargaining."[215] Collective-bargaining parties did establish their own dispute-settlement procedures, but as just shown, such procedures lacked any legal status: the *ultima ratio* for any dispute regarding the dispute procedures themselves always remained the resort to economic self-help.[216] Further, British law did provide for state-initiated conciliation of disputes, but in practice this service tended to reinforce rather than undermine the dominant British model of collective laissez faire. That is, state intervention could not dictate, but only facilitate, a settlement and the procedure was only triggered after the parties' own negotiation machinery had failed.[217] Finally, just as in the U.S., the British state did provide employers and unions with voluntary, nonbinding arbitration services. Significantly, however, the law establishing the arbitration service never created a permanent board that could develop an expertise or a set of principles to apply in the

211. KAHN-FREUND, *supra* note 150, at 132.

212. *Id.* at 146 ("[T]he elementary rule of contractual construction . . . that the parties are, in the absence of an express term to the contrary, deemed to have implicitly incorporated the substance of the prevailing usages or customs, remains the principal link between collective agreements and contracts of employment.").

213. *Id.* at 149.

214. *Id.*

215. *Id.* at 77.

216. *See supra* text accompanying note 198–201.

217. *Id.* at 98–100.

settlement of workplace disputes.[218] In sum, the exceptions in British labor law tended to prove the rule.

C. Restrictions on Economic-Action Alternatives

Key to the Wagner Act's voluntarist system of labor law is the use of "economic weapons," such as strikes and other economic actions. Since U.S. labor law does not dictate any of the terms to which unions and management agree, the resulting bargaining must be one determined by the relative economic power of the parties. The Wagner Act's commitment to "industrial peace and stability" also dictates that economic actions will be an important object of regulation.[219] The governance of these economic weapons is therefore a main concern of U.S. labor law.

The regulation of economic action also critically affects the fortunes of workplace associations.[220] We have previously examined two reasons why workplace associations will tend to resort to strategies of economic self help. First, they lack access to formal proceedings either before the Board or made available under the collective-bargaining agreement because in the presence of a union they do not have exclusive-representation status.[221] Second, union-member workers do not have the same opportunities to acquire knowledge or specialize in the arcane and formal procedures under the NLRA or the collective agreement.[222] If workplace associations wish to achieve their aims, they will therefore need to resort to the use of strikes or other economic actions. This was exactly the conundrum faced by the workers in the *Emporium Capwell* case.[223] Yet U.S. labor law places restrictions on the ability to strike and take other economic action and therefore increases the costs of precisely those alternatives to which workplace associations must resort to advance their demands. This further reduces the strength and possibility of

218. *Id.* at 100. Britain also experimented with a system of compulsory arbitration arising out of wartime industrial-relations experience and which nevertheless proved unenduring. Regarding Britain's final experiment with compulsory arbitration, which lasted from 1951 to 1959, Kahn-Freund remarked, "If this system had been more important in practice than it was it would have been inconceivable for the employers to put up with it for more than seven years." *Id.* at 117. The British government also experimented during the two world wars with government-sponsored efforts to establish industry-wide procedure agreements for collective bargaining by instituting Joint Industrial Councils (JICs) where unions had not been established. Though notable as an exception to the British system of collective laissez faire, the JICs remained peripheral to the established trade union sections and did not contribute to the formal institutionalization of collective bargaining. *See* PAUL DAVIES & MARK FREEDLAND, LABOUR LEGISLATION AND PUBLIC POLICY: A CONTEMPORARY HISTORY 39–43 (1993). *But see* K.D. Ewing, *The State and Industrial Relations: "Collective Laissez-Faire" Revisited*, 5 HIST. STUD. INDUS. REL. 1, 17–20 (1998); CHRIS HOWELL, TRADE UNIONS AND THE STATE: THE CONSTRUCTION OF INDUSTRIAL RELATIONS INSTITUTIONS IN BRITAIN, 1890-2000, at 72–73 (2005) (portraying a more positive impact of the JICs).

219. *See* Textile Workers Union v. Lincoln Mills, 353 U.S. 448, 454–55 (1957); NLRB v. Local Union No.1229, Int'l Bhd. of Elec. Workers, 346 U.S. 464, 476 (1953) (stating the declared purpose of the Wagner Act is to promote "industrial peace and stability").

220. *See infra* Figure 2.

221. *See supra* text accompanying notes 168–83.

222. *See supra* Part III.B.

223. *See* Emporium Capwell Co. v. W. Addition Cmty. Org., 420 U.S. 50, 53–56 (1975).

workplace associations. In Britain by contrast, legal restrictions on economic action are unknown.

1. The Law of Economic Action in the United States

Although Sections 7 and 13 of the NLRA purported to give employees a broad right to strike,[224] this right has been vastly restricted in a number of ways. Important for our story are three main changes: (1) the statutory restrictions on organizational and recognition strikes and secondary boycotts added to the NLRA by the Taft-Hartley Act of 1947; (2) judicial deprivation of protection for what I will call "shop-floor tactics"; and (3) the effect of unions' no-strike agreements on workplace associations.

a. Statutory Restrictions on Economic Action

The possibilities for workplace association can be quite favorable at the organization and recognition stages of collective bargaining, when the position of the labor union is less established.[225] In this respect, the most salient restrictions on economic action are the Taft-Hartley Act's limitations on organizational and recognition strikes. Among these provisions is Section 8(b)(4)(C), which bans the use of economic action to compel an employer to recognize or bargain with a particular labor organization in the case where another labor organization had already been certified as the employees' representative.[226] A violation of any of 8(b)(4)'s prohibitions may be redressed under Section 303 of the LMRA with an action for damages in federal court, a remedy not available for any employer unfair labor practice.[227] Thus when employees seek to bargain directly with an employer in circumvention of the exclusive union representative, not only does an employer face an unfair labor practice sanction and and not only is employees' economic action unprotected,[228] as we have already seen, but economic action with such an object may be patently illegal as well. This further raises the costs of workplace associations.

A possible distinction is the restriction of the 8(b)(4)(C) prohibition to "labor organizations," which conceivably could be interpreted not to

224. *See* National Labor Relations Act (NLRA) § 7, 29 U.S.C. § 157 (2006); *id.* § 13 ("Nothing in this [Act] . . . shall be construed so as either to interfere with or impede or diminish in any way the right to strike, or to affect the limitations or qualifications on that right.").

225. *See* Gould, *supra* note 112, at 680.

226. National Labor Relations Act (NLRA) § 8(b)(4)(C) (making it an unfair labor practice for a labor organization to engage in economic action when the object is "forcing or requiring any employer to recognize or bargain with a particular labor organization as the representative of his employees if another labor organization has been certified as the representative of such employees under the provisions of section 159").

227. Labor Management Relations (Taft-Hartley) Act § 303, 29 U.S.C. § 187 (2006) (making unlawful "any activity or conduct defined as an unfair labor practice in Section 158(b)(4)" and creating a right of action for damages in US district courts).

228. *See supra* text accompanying notes 168–83.

cover employees or their associations. However, the Act defines labor organization quite broadly, as including "any organization of any kind, or any agency or employee representation committee or plan" that at least in part deals with employers over terms and conditions of employment.[229]

A case from the First Circuit demonstrates both the negative impact of Section 8(b)(4)(C) on workplace associations as well as the breadth of the Act's definition of "labor organization." In *Simmons, Inc. v. NLRB*[230] the circuit court vacated the Board's order dismissing an employer's complaint that a committee of workers had violated Section 8(b)(4)(C) by initiating a strike demanding that the employer bargain with the committee, although a labor union had already been elected and certified.[231] At issue was whether the committee of workers was a "labor organization."[232] The court concluded that it was, reasoning that if the committee "sought to have itself recognized or bargained with, then it acted as a labor organization."[233] Thus, the same conduct of the committee that made its action prohibited under the NLRA also defined it as a labor organization.

In addition, the Taft-Hartley Act also prohibits the use of secondary tactics by labor unions.[234] Secondary actions are the use of economic or other pressure on an employer who is not the primary party to a labor dispute.[235] Such restrictions are potent, for as is recognized, secondary actions are "one of the most effective weapons in labor's economic arse-

229. National Labor Relations Act (NLRA) § 2(5); *see also* NLRB v. Cabot Carbon Co., 360 U.S. 203, 210–14 (1959) (holding that although employee committees did not bargain with employers in "the usual concept of collective bargaining," they were nevertheless labor organizations because they existed in part for the purpose of dealing with employers concerning grievances).

230. 287 F.2d 628 (1st Cir. 1961).

231. *Id.* at 631.

232. *Id.* at 628.

233. *Id.* at 629, 631.

234. National Labor Relations Act (NLRA) § 8(b)(4)(B) (making it an unfair labor practice for a labor organization to engage in economic action when the object is "forcing or requiring any person to cease using, selling, handling, transporting, or otherwise dealing in the products of any other producer, processor, or manufacturer, or to cease doing business with any other person . . . Provided, That nothing contained in this clause (B) shall be construed to make unlawful, where not otherwise unlawful, any primary strike or primary picketing").

235. Section 8(b)(4)(B)'s proviso to preserve the primary right to strike or picket has led the Board to develop a labyrinthine set of rules governing the secondary-activity ban that, if anything, increases uncertainty and confusion for workers. For instance, under certain conditions, unions can target worksites that the primary employer shares with another employer (a "common situs"). Economic action can target "allies" (interpreted narrowly by the Board as those employers so integrated with the struck employer that they form a "common enterprise" and entities that perform struck work of the primary employer). And under the *Moore Dry Dock* rules, striking workers can follow and picket supervisors, nonstrikers, and replacements that work at the location of a secondary employer. *See* Sailor's Union of the Pac. v. Moore Dry Dock Co., 92 N.L.R.B. 547, 549 (1950). There is also some permissiveness for what are broadly called "sympathy strikers"; Section 8(b)(4) does not prohibit strikers at a primary employer's location from asking delivery drivers, vendors, or outside contractors from honoring their picket line. *See* Chauffeurs, Teamsters & Helpers Local Union No. 175 v. NLRB, 294 F.2d 261, 262 (D.C. Cir 1961) (per curiam).

nal."[236] Although one may question whether workplace-based employee associations would be able to coordinate multi-firm economic actions, it is clear that where developed such associations can achieve a remarkable degree of interplant organization, which would allow them to exploit these potent tactics.[237]

b. No Protection for "Shop Floor" Economic Actions

In addition to these legislative prohibitions, the courts have also deprived from the protection of the NLRA certain forms of economic action in which workplace associations may have a comparative advantage over union bureaucracy. As we shall shortly see, *unprotected* economic actions leave workers in the U.S. no worse off from a strictly legal perspective than their counterparts in Britain, where labor law has provided no "positive" right to strike. Nevertheless, the very distinction in U.S. labor law between protected and unprotected activities may create a normative (as contrasted with legally coercive) disincentive to strike. The lack of illegitimacy is absent under British labor law, where the absence of a right to strike is the "normal" status of a strike.

Among the set of unprotected activities most likely to affect workplace associations are what might be called shop-floor tactics: work stoppages that entail not just quitting work, but that also obstruct the ability of employers to restart production with replacement workers because striking workers remain in the workplace. Workplace associations have an advantage in such tactics because they typically require a high degree of workplace communication and coordination. And because they make it more difficult for the employer to restart production, they are highly effective. The most important decision in this category of cases is *NLRB v. Fansteel Metallurgical Corp.*[238] In that case, the Court held that sit-down strikes were not protected under the act and striking sit-down workers may be terminated without redress.[239] The decision in *Fansteel* could have been read to deprive strike activity of legal protection in the case when workers trespass on the employer's property.[240] However, later courts drew on *Fansteel* to examine a variety of different work-stoppage

236. 2 COMM. ON THE DEV. OF THE LAW UNDER THE NAT'L LABOR RELATIONS ACT, AM. BAR ASS'N, THE DEVELOPING LABOR LAW: THE BOARD, THE COURTS, AND THE NATIONAL LABOR RELATIONS ACT 1621 (Patrick Hardin et al. eds., 4th ed. 2001) [hereinafter DEVELOPING LABOR LAW].

237. *See* EDELSTEIN & WARNER, *supra* note 34, at 18 (commenting on "combine" committees of shop stewards linking workplace organizations across multiple plants in the British automobile industry); Terry, *supra* note 64, at 69 (noting the "considerable sophistication" of inter-plant coordination of workplace organizations in the British engineering industry).

238. 306 U.S. 240 (1939).

239. *Id.* at 256–58.

240. Craig Becker, *"Better Than a Strike": Protecting New Forms of Collective Work Stoppages Under the National Labor Relations Act*, 61 U. CHI. L. REV. 351, 367–68 (1994); *see also* James Gray Pope, *How American Workers Lost the Right to Strike, and Other Tales*, 103 MICH. L. REV. 518, 520–26 (2004).

tactics.[241] Decisions following *Fansteel* therefore found that intermittent work stoppages, slowdowns, and partial strikes were also unprotected concerted activities.[242]

c. No-Strike Agreements and Workplace Associations

Another important source of restrictions on economic action has been promises unions make not to strike during the term of a collective-bargaining agreement. Both scholars and activists have frequently highlighted no-strike agreements as a prime example of the way insular labor leaders undermine the prospects for more vigorous workplace associations in the U.S.[243] Yet, the relationship between no-strike agreements and workplace associations is more complicated than it would otherwise appear. First, individual union members are not liable for damages when they engage in unauthorized work stoppages in violation of a no-strike agreement.[244] Second, absent express contractual assumptions of responsibility, unions face reduced standards of liability for damages in cases of wildcat strikes that violate a no-strike promise.[245] To the extent that liability of the union encourages it to be more vigilant in policing the union's ranks for dissenters, these reduced standards weaken those disciplinary incentives.[246] Thus, no-strike agreements may actually have a fairly weak effect on workplace associations, in terms of the prospects of damages liability.

---

241. Becker, *supra* note 240, at 368–71.

242. *See* DEVELOPING LABOR LAW, *supra* note 236, at 1488–90.

243. *See* DAVID BRODY, WORKERS IN INDUSTRIAL AMERICA: ESSAYS ON THE TWENTIETH CENTURY STRUGGLE 185–195 (2d ed. 1993) (arguing that the "contractual logic" and no-strike promises of collective bargaining agreements "evolved into a pervasive method for containing shop-floor activism").

244. Labor Management Relations (Taft-Hartley) Act § 301(b), 29 U.S.C. § 185(b) (2006) ("Any money judgment against a labor organization in a district court of the United States shall be enforceable only against the organization as an entity and against its assets, and shall not be enforceable against any individual member or his assets."); *see also* Complete Auto Transit, Inc. v. Reis, 451 U.S. 401, 407 (1981) (holding individual union members not liable for damages in the express case of an unauthorized strike in violation of a no-strike obligation); Atkinson v. Sinclair Ref. Co., 370 U.S. 238, 247–48 (1962) (holding that section 301(a) of the LMRA does not authorize a damages action against individual union officers and members when their union is liable for violating a no-strike clause in a collective-bargaining agreement).

245. *See* Labor Management Relations (Taft-Hartley) Act § 301(b) (stating that labor unions "shall be bound by the acts of its agents"). *But see* § 301(e) (stating that in determining union responsibility for acts of its agents, the question of actual authorization or subsequent ratification shall not be controlling). The federal courts nevertheless differed on what was the union's standard of liability. *Cf.* Eazor Express, Inc. v. Int'l Bhd. of Teamsters, 520 F.2d 951, 962 (3d Cir. 1975) (holding the union liable for violation of a no-strike obligation for failure to use best efforts to end unauthorized strikes); United Constr. Workers v. Haislip Baking Co., 223 F.2d 872, 877–78 (4th Cir. 1955) (holding the union not liable for damages from an unauthorized strike when there was no evidence that the union "adopted the strike, that they encouraged it or that they prolonged it"). This tension was not resolved until the Supreme Court's decision in Carbon Fuel Co. v. United Mine Workers, 444 U.S. 212, 216–18 (1979) (holding that LMRA § 301(e) established that the common-law test of agency should be used to determine union liability for the acts of its agents). *Carbon* also held that absent an express agreement in the contract, there was no implied duty of the union to "use all reasonable means to prevent and end" unauthorized strikes. *Id.* at 216.

246. The subject of union discipline has received much attention.

Nevertheless, while the Supreme Court held in *Complete Auto Transit, Inc. v. Reis*[247] that individual union members were not liable for damages for an unauthorized strike that violated a no-strike agreement, the Court expressly declined to address the issue of whether an employer could obtain injunctive relief in such a situation.[248] Because of its immediacy, a temporary restraining order could more effectively hamper incipient workplace associations than would the more remote prospect of damage liability. Commentary suggests that an injunction is indeed the typical and accepted response to a wildcat strike.[249] Moreover, courts have long held that wildcat strikes are unprotected,[250] and again this distinction from protected activities may give such actions the cast of illegitimacy and induce a normative, if not coercive, constraint on economic action. Indeed, the fact that labor law practitioners refer to the terminations of wildcat strikers as the industrial-relations equivalent of "capital punishment" speaks in favor of this conclusion.[251] Therefore, despite the distant impacts of damages liability to wildcat strikers, injunctions and the normative-legitimacy constraints on authorized job actions undoubtedly inhibit workplace associations.

2. The Law of Economic Action in Britain

Unlike the Taft-Hartley prohibitions on organizational, recognition, and secondary strikes in U.S. labor law, no legal restrictions on strikes existed under British law. Also unlike U.S. law, the idea of legally-protected concerted activities was unknown. Rather, British workers and union members have had for much of the twentieth century only privileges and immunities: for a broad range of activities, economic action was neither legally prohibited nor legally protected. This system was the

247. 451 U.S. 401 (1981).

248. *Id.* at 417 & n.18. It is clear in the case of union-authorized strikes that employers can obtain an injunction. *See* Boys Mkts., Inc. v. Retail Clerks Union, Local 770, 398 U.S. 235, 252–54 (1970) (upholding an injunction for a strike in violation of a no-strike clause that was precipitated over a dispute subject to arbitration in the collective-bargaining agreement), *overruling* Sinclair Ref. Co. v. Atkinson, 370 U.S. 195, 199–201 (1962) (holding that the Taft-Hartley Act provision authorizing suits against unions did not impliedly repeal the Norris-LaGuardia Act prohibition against labor injunctions in federal courts); *see also* Buffalo Forge Co. v. United Steelworkers, 428 U.S. 397, 406–09 (1976) (refusing to expand *Boys Markets* to strikes in violation of a no-strike obligation over disputes not subject to grievance arbitration in the collective-bargaining agreement).

249. M. Jay Whitman, *Wildcat Strikes: The Unions' Narrowing Path to Rectitude?*, 50 IND. L.J. 472, 473 (1975) (explaining that in a wildcat strike, the "employer is typically content with a *Boys Market* injunction and a speedy resumption of production").

250. NLRB v. Draper Corp., 145 F.2d 199, 205 (4th Cir. 1944). Wildcat strikes need not violate a no-strike agreement to be unprotected. In *Draper*, the court concluded that the employees's actions were unprotected because wildcat strikes undermine the principle of exclusive representation, not because the strike violated a no-strike agreement. *Id.* at 202–04. However, for a strike to be "unauthorized," a duly recognized labor union must be present, so an unannounced walkout in an unorganized workplace is generally protected under section 7 of the NLRA. *See* NLRB v. Wash. Aluminum Co., 370 U.S. 9, 14–16 (1962).

251. *See* Feller, *supra* note 128, at 780.

product of two statutes that the British labor movement achieved in the culmination of a long legislative battle.[252]

First, economic action was immunized from criminal liability in 1875. The Conspiracy and Protection of Property Act (CPPA) immunized any act done "in contemplation or furtherance of a trade dispute" from the common-law doctrine of criminal conspiracy;[253] abolished certain individual crimes associated with strikes, such as "molestation" and "obstruction";[254] and sanctioned certain forms of picketing.[255] In 1890, British courts began entertaining actions alleging that trade unions were civilly liable for activity that had been decriminalized in the CPPA.[256] In response, British Parliament passed the Trade Disputes Act (TDA) in 1906 which immunized peaceful picketing from civil liability,[257] eliminated the tort of inducement to breach of contract in the context of a trade dispute,[258] and protected trade union funds from civil remedies.[259] The famous Section 1 of the TDA, enshrined in British labor law as the "golden formula,"[260] immunized strikes against civil conspiracy: "An act done in pursuance of an agreement or combination by two or more persons shall, if done in contemplation or furtherance of a trade dispute, not be actionable unless the act, if done without any such agreement or combination, would be actionable."[261]

This "extraordinary freedom"[262] conferred by the TDA proved to be broad. The Act's language granted immunity to actions done "in contemplation or furtherance of a trade dispute" and defined a trade dispute as a dispute between employers and workmen or between workmen "connected with the employment or nonemployment, or the terms of employment, or the conditions of labour of any person."[263] Importantly, unlike the U.S., this immunity was determined to cover recognition strikes and secondary disputes.[264] The immunities granted by the TDA, along with the CPPA, were also extremely durable: with only minor modification it "remained on the statute book for seventy-five years" and became the "bedrock of the British system of labour law."[265]

---

252. *See* Forbath, *supra* note 31, at 22–31 (providing a concise account of this story); *see also* Michael J. Klarman, *The Judges Versus the Unions: The Development of British Labor Law, 1867-1913*, 75 VA. L. REV. 1487, 1487–90 (1989) (providing a more detailed account of this story).

253. Klarman, *supra* note 252, at 1496.

254. *Id.*

255. *Id.*

256. *Id.* at 1505–21.

257. *Id.* at 1521.

258. *Id.*

259. *Id.* at 1535–36.

260. Forbath, *supra* note 31, at 24, 31.

261. Trade Disputes Act, 1906, 6 Edw. 7, c. 47, § 1 (Eng.), *available at* http://www.constitution.org/sech/sech_136.htm.

262. Forbath, *supra* note 31, at 31.

263. KAHN-FREUND, *supra* note 150, at 248.

264. *Id.* at 249.

265. DAVIES & FREEDLAND, *supra* note 218, at 15.

CONCLUSION

The goal of this Article has been to argue that the law makes a difference for the possibility of union democracy and that, in the United States, labor law restricts this possibility. The central claim is that labor law can obstruct or promote the amount of workplace association either directly, or indirectly through the law's effect on union bureaucracy; and that the strength of workplace association is critical for the maintenance and florescence of union democracy. Both exclusive representation and legally-supported procedures of collective bargaining promote the bureaucratization of unions, which negatively impacts the amount of workplace association. And both exclusive representation and restrictions on the use of economic action, which workplace associations often resort to in order to voice their grievances, further hinder the growth of worker self organization. Whereas such a constellation of rules accurately characterizes labor law in the U.S., the labor law contrasts in each respect in Britain. And these different legal configurations can explain contrasting patterns in labor organization between the two countries. British unions have small bureaucracies, strong workplace associations, and enjoy a greater measure of democracy. American unions have large bureaucracies, weak workplace associations, and suffer a democracy deficit.

The main thrust of this article has been to *explain* why U.S. labor law presents affirmative obstacles to the emergence of union democracy. However, before we embrace the *normative* conclusion that features of the British collective laissez faire system ought to be adopted if union democracy is the goal, a much fuller normative and policy discussion ought to be considered. While such a discussion is beyond the scope of this Article, some of the relevant points will at least be raised here.

The overriding consideration is the one that began the Article in the Introduction: the paramount need for the labor movement to organize new members. There is a contentious debate about whether union democracy hinders or helps unions generally and more specifically in the case of organizing.[266] Certainly, the claim that union democracy is ineffective is unpersuasive, as the example of the United Healthcare Workers, as well as many others, demonstrates so powerfully. In fact, if anything, the story of the UHW suggests that union democracy and successful organizing are essential for one another. Some would also point out that union members themselves can be an obstacle to union revitalization.[267] But in this case, there are strong indications that such behavior is a highly cultural and learned response to conditions of bureaucratic and

266. Estreicher, *supra* note 18, at 247 (claiming that the "pursuit of union democracy is . . . counterproductive because it . . . weakens (or complicates) [unions'] ability to wage economic struggle with employers").

267. Voss & Sherman, *supra* note 9, at 320–21.

undemocratic unions.[268] How members act in an undemocratic union is a poor guide to how they would respond in a democratic one.

The more troublesome objection is not that union democracy would induce poor union performance, but that aspects of the British collective laissez faire model would hinder unions in other ways. For example, the absence of exclusive representation negatively affects not only the size of union servicing bureaucracies in Britain but also their resources. And as the research on union organizing we introduced earlier has also shown, union organizing is a highly staff and resource intensive endeavor. While it requires a reduction in the union's traditional bureaucracy, and therefore an increase in workplace association, it also requires new organizing and research staff.[269] Smaller bureaucracy has a salutary effect on workplace associations, but abolishing exclusive representation may do too much, diminishing the union's resources and weakening bureaucracy across the board, including the organizing and research departments that might work more like *complements* rather than substitutes in the specific case of union organizing.

How and whether this conundrum can be resolved is a matter that will have to occupy further reflection and research. Can exclusive representation be amended to unburden workplace associations while preventing inter-union competition, in a way similar to that sought by the plaintiffs in *Emporium Capwell*? Should the logic of contractualism in the workplace be weakened? Will this undermine the traditional *servicing* bureaucracy sufficiently while avoiding the drastic effects of abolishing exclusive representation? Whatever the conclusion, the burden of this Article has been to demonstrate that the present configuration of labor law is likely one that is extremely unfavorable to revitalization and transformations unions must undertake to organize on a serious scale, and that certain changes are likely necessary if the labor movement wishes to once again become a vibrant and important part of our economic and political landscape.

APPENDIX

The formal analysis of union democratization is presented in two games: the union-bureaucracy game and the union-democracy game. The first game, the union-bureaucracy game, generates predictions of the sizes of union bureaucracy and workplace association given an initial, nondemocratic union. The other main goals of this game are to show the inverse relationship between union bureaucracy and workplace association and the effect of exclusive representation on union bureaucracy. Given the levels of union bureaucracy and workplace association determined in the union-bureaucracy game, the union-democracy game then

268. *Id.* at 321–23.
269. *Id.* at 313; *see also* Bronfenbrenner & Hickey, *supra* note 10, at 54.

makes predictions about the likelihood of democracy when the union leadership's credibility is at issue. Presenting the analysis in two games, rather than one, effectively means workers do not take into account the impact of workplace association in the second game, when they determine its level in the first game. This is necessary given our (more realistic) assumption that workplace association has only an unintended effect on the likelihood of union democracy.

A. The Union-Bureaucracy Game

1. Players, Actions, Order of Play, and Payoffs

Consider a game with two players, the leadership of a union, L, and a group of workers, W, of size n, who we treat as a unitary actor. In the first step of the game, the union leadership selects the level of monetary contributions, or dues, that workers will contribute to the union, $d \in [0,\infty)$, and the size of its bureaucracy, $b \in [0,\infty)$. In the second step, workers choose their level of association, $a \in [0,\infty)$, and whether or not to accept the union's offer. If workers reject the union's offer, leaders and workers each receive a reservation payoff normalized to 0.

Each worker receives a single, collective good, which is a function of the levels of bureaucracy and workplace association, $w(a,b)$. The key assumption is that a and b are perfect substitutes. Production of the collective good is assumed to take on the natural logarithmic form. Thus, $w(a,b) \equiv \ln(a+b)$. Although $\ln(0)$ is undefined, at least a or b or both will be strictly positive. The assumption of diminishing returns for the production of the collective union good seems natural, since there are definite limits to efficiency in the size of organizations, bureaucratic or otherwise. The union receives a dues payment from each worker. To capture limitations on the residual claim (e.g., unions' non-profit status), we assume that the leadership captures only a fraction of the dues revenue, $\sigma \in [0,1]$, so the union receives σnd. Bureaucracy and workplace association are costly; the cost of association is linear and is given by αa, while the cost of bureaucracy is convex and is given by $\beta b^2/2$. Also assuming convex costs for workers would not change the underlying intuition of the results, but the linear cost assumption is maintained to make those results clearer here.

2. Equilibrium

Using subgame-perfect Nash equilibrium to find the solution of this game, we begin with the workers' problem. The workers' utility function is given by:

$$V_W = n\ln(a+b) - nd - \alpha a$$

Workers choose a to maximize V_W. More precisely, assuming that their participation constraint is satisfied, workers set:

$$a^* = \begin{cases} n/\alpha - b, & \text{if } b < n/\alpha \\ 0, & \text{if } b \geq n/\alpha \end{cases}$$

where $n/\alpha - b$ is obtained by differentiating the workers' objective function with respect to α and setting it equal to zero. As is clearly seen, workers' associational effort is decreasing in its cost as well as in the size of the union's bureaucracy.

Since the reservation payoff is zero if workers do not accept the union's offer, workers' optimal choice of association can be substituted into their objective function, and their participation constraint can be written as $d \leq \ln(a^* + b) - \alpha a^*/n$. Since union leaders will want to set the level of dues as high as possible, we can define a maximum level of dues:

$$\bar{d} \equiv d = \ln(a^* + b) - \alpha a^*/n \quad (1)$$

Some further analysis would show that $\bar{d}$ will still be positive even if the amount of bureaucracy is zero. This captures the idea that the leadership in a nondemocratic union is able to transfer to itself "unearned" rents. We will see how the union members can remedy this situation in the union-democracy game.

Having solved the workers' problem, we can now back up and address the union leadership's problem. The union maximizes:

$$V_L = \sigma n d - \beta b^2/2$$

with respect to d and b. Given the workers' optimal solution, and assuming that dues are set as high as possible, the union's objective function can be rewritten as:

$$V_L = \sigma[n \ln(a^* + b) - \alpha a^*] - \beta b^2/2$$

Plugging in the workers' optimal solutions yields:

$$V_L = \begin{cases} \sigma[n \ln(n/\alpha) - (n - \alpha b)] - \beta b^2/2, & \text{if } b < n/\alpha \\ \sigma n \ln(b) - \beta b^2/2, & \text{if } b \geq n/\alpha \end{cases}$$

Differentiating each equation with respect to b and setting each equal to zero gives the union's optimal choice of bureaucracy:

$$b^* = \begin{cases} \sigma\alpha/\beta, & \text{if } b < n/\alpha \\ (\sigma n/\beta)^{1/2}, & \text{if } b \geq n/\alpha \end{cases}$$

As is clearly seen, the optimal level of bureaucracy is increasing in σ, the share of the surplus which the union can appropriate, increasing in α (for $b < n/\alpha$), the cost of workers' association, and decreasing in β, the cost of bureaucracy. Finally, note also, that if workers could influ-

ence the dues level, they would prefer that the union choose α/β (when $b < n/\alpha$).

3. Extension: Rival Unions

Now consider a version of the above game, but where there are two union leaderships, $L = \{1,2\}$. We can think of this situation as equivalent to the absence of exclusive representation and the previous problem with one union as equivalent to the presence of exclusive representation. Each union chooses its level of dues and bureaucratic effort and the workplace association can allocate any proportion of workers between the two unions or decide to reject both unions' offers. Since the workplace good is a collective one, workers receive the good regardless of which union they join. The collective good function is given by $w(a, b_1, b_2) \equiv \ln(a + b_1 + b_2)$. The union leadership receives a dues payment from each worker who joins its union, with $n_1 + n_2 = n$.

The workers' utility function is now given by:

$$V_W = n\ln(a + b_1 + b_2) - n_1 d_1 - n_2 d_2 - \alpha a$$

And the optimal choices of association are:

$$a^* = \begin{cases} n/\alpha - b_1 - b_2, & \text{if } b_1 + b_2 < n/\alpha \\ 0, & \text{if } b_1 + b_2 \geq n/\alpha \end{cases}$$

While the workers' optimal solution is similar to the previous game, the problem facing the unions is much different. Since workers obtain the collective good regardless of which union they are allocated to, they will join a union based only on its choice of dues. At any dues level that union 1 sets lower than union 2, $d_1 < d_2$, (or vice versa), the number of workers joining union 2 is zero, $n_2 = 0$, and union 2's payoff becomes $-\beta b_i^2/2$, which is inconsistent with its participation constraint. Each union will therefore set dues as low as possible.

How low will each union set dues? Answering this question will also help us understand each union's choice of b. When $d_1 = d_2$, the lowest dues that will just satisfy each union's participation constraint is such that $\sigma n_i d_i - \beta b_i^2/2 = 0$. Solving this for d_i gives a minimum level of dues:

$$\underline{d} \equiv d_i = \beta b_i^2/2\sigma n_i \quad (2)$$

This minimum dues level is clearly increasing in the level of bureaucracy. Therefore, even when each union's dues are equal and positive, each will have an incentive to lower bureaucratic effort by a fraction, which from the above equation will allow the union to lower its minimum dues by a fraction. But then all workers will join that union and the union's revenues will increase dramatically. My intuition is that

this process will continue until both the level of effort and dues are equal to zero. In reality we observe positive dues and effort between competing unions, as in Britain. But this reasoning yields the correct qualitative insight that competition between unions will both lower dues *and* bureaucratic effort.

B. The Union-Democracy Game

1. Players, Actions, Order of Play, and Payoffs

The union-democracy game is once again a game between two players, the union leadership (L) and the union members (W). In the first stage of the game, the union leadership decides whether to concede to the installation of democracy (D) or to maintain an oligarchy (O). In the second step, the level of union dues is set: d_O denotes the dues rate set by the leadership in a nondemocratic union and d_D denotes the dues rate set in a democracy by the union members. The set of possible dues levels is the interval $[\underline{d},\overline{d}]$, where $\underline{d}$ and $\overline{d}$ are as defined in equations (1) and (2) at the optimal level of bureaucracy, b^*, as determined in the previous game. (Thus, in the union-democracy game b^* becomes a parameter of the model.) If the leadership chooses D at the first stage, union members determine the level of dues, but if it chooses O, then the leadership determines dues. Following the dues decision, the union members choose whether to stage a revolt. In a revolt, union members cease paying dues. A revolt always succeeds if attempted, but a tradeoff is faced since members depend solely on their own associational capacity, a, which is determined by members' optimal choice of effort, a^*, in the union-bureaucracy game. Without loss of generality, I assume that a^* always takes on its nonnegative value, $n/\alpha - b^*$, where b^* is again determined by the union's optimal choice of bureaucracy in the union-bureaucracy game. Since b^* can take any value in $[0,\infty)$, we can capture the full range of outcomes in the union-democracy game.

Therefore, if workers undertake a revolt, their payoffs are:

$$V_W(R) = n\ln(n/\alpha - b^*) - (n - ab^*)$$

Since the union is deprived of revenue, the payoff to the union is $V_L(R) = -\beta(b^*)^2/2$. Without loss of generality, we say that the revolution constraint is binding if the workers obtain more in a revolt than when the leadership chooses its ideal dues level, $\overline{d}$. Therefore the revolution constraint is binding if $V_W(R) > V_W(O,\overline{d})$. Note that since the union sets dues as high as possible consistent with workers' participation constraint, then $V_W(O,\overline{d}) = 0$. The revolution constraint then simply reduces to the condition that the collective good that workers can produce with their own effort is greater than the costs of doing so, or:

$$n\ln(n/\alpha - b^*) > n - \alpha b^*$$

Because of the logarithmic collective-good function we have chosen,[270] this constraint may or may not bind. For instance, as $b^* \to n/\alpha$, and in particular as $(n/\alpha - b^*) \to 1$, then the condition tends toward $0 > \alpha$, which will fail to hold for any positive α. On the other hand, when $b^* \to 0$, then the condition goes to $\ln(n/\alpha) > 1$, which will hold for n large enough and α small enough. In other words, the revolution constraint is more likely to bind when both the size of union bureaucracy and the cost of workplace association are smaller. This is one way of seeing the crowding-out effect in operation. If workers undertake a revolt, the game ends with payoffs $(V_W(R), V_L(R))$.

If democracy has been created and there is no revolt, then the game ends with dues set at the level preferred by the membership. Members want the lowest dues possible, $d_D = \underline{d}$. Therefore, in the case of democracy, payoffs are:

$$V_W(D) = n\ln(n/\alpha) - n\underline{d} - (n - \alpha b^*)$$

$$V_L(D) = 0$$

Rather than democracy, the leadership can choose a nondemocratic form of union governance and set the level of dues themselves. In this case, in an attempt to stave off a revolt, the leadership will choose $d_N = \hat{d}$, $\underline{d} \leq \hat{d} \leq \overline{d}$. Following this decision, we are again at the stage where workers choose to revolt, with payoffs the same as before if workers in fact revolt. In a nondemocracy, however, whether the dues level that the union sets becomes the effective dues level depends on whether the leadership can credibly commit to maintain its promised level. Therefore, if workers choose not to revolt, nature then moves and determines $p \in [0,1]$, where, with probability p, the promise that the union gave with respect to dues stands, but with probability $1 - p$, the union reneges and resets the dues level. If the leadership's promise is credible, then payoffs are:

270. It would possible to find collective good functions such that the revolution constraint would bind for any positive level of association. This would not change the main results.

$$V_W(O, d_O = \hat{d}) = n\ln(n/\alpha) - n\hat{d} - (n - \alpha b^*)$$

$$V_L(O, d_O = \hat{d}) = \sigma n\hat{d} - \beta(b^*)^2/2$$

However, if the leadership's promises aren't credible, then after the workers have decided not to revolt, the leadership can do no better than to set dues at their most preferred level. When nature lets the leadership reset the dues level, then payoffs are:

$$V_W(O, d_O = \bar{d}) = n\ln(n/\alpha) - n\bar{d} - (n - \alpha b^*)$$

$$V_L(O, d_O = \bar{d}) = \sigma n\bar{d} - \beta(b^*)^2/2$$

A game-tree depiction of the order of play and payoffs of the union-democracy game is found in Figure 3.

2. Equilibrium

Since workers are unsure whether the leadership is credible or not, the expected payoff to workers not revolting in an oligarchy is $pV_W(O, d_O = \hat{d}) + (1-p)V_W(O, d_O = \bar{d})$, which we can also write as $V_W(O, d_O) = n\ln(n/\alpha) - n[p\hat{d} + (1-p)\bar{d}] - (n - \alpha b^*)$. If the leadership can choose a dues level such that $V_W(O, d_O = \hat{d}) \geq V_W(R)$, then such a concession will be sufficient to stop a revolt. This condition may or may not hold. To see this, let $\hat{d} = \underline{d}$, which is the best concession the union can make to the workers (i.e., if the condition fails at $\underline{d}$, it fails at any $\hat{d}$). Substituting payoffs into the condition and simplifying a bit, we get:

$$\ln(n/\alpha) - [p\underline{d} + (1-p)\bar{d}] \geq \ln(n/\alpha - b^*)$$

Recall that the left-hand side is the payoff from concessions and the right-hand side the payoff from revolt; the only difference is that workers' costs of association, $(n - \alpha b^*)$, on both sides cancel each other out.

Consider first the case where the leadership's promises are perfectly credible, that is $p = 1$. Then the condition reduces to $\ln(n/\alpha) - \underline{d} \geq \ln(n/\alpha - b^*)$. As $b^* \to 0$, then $\underline{d} \to 0$, and the condition becomes $\ln(n/\alpha) \geq \ln(n/\alpha)$. Therefore, workers are indifferent between concessions and revolt when the size of union bureaucracy is zero and leadership promises are perfectly credible. However, as $(n/\alpha - b^*) \to 1$, the condition becomes $\ln(n/\alpha) - \underline{d} \geq 0$, and we know that the left hand side is positive, since the concession payoff must still satisfy workers' participation constraint, the costs of workplace association costs are absent, and $\underline{d} < \bar{d}$. In this case, it is again easy to observe the crowding out effect. As bureaucracy increases, the payoff to

revolt (the right-hand side) decreases: bureaucracy crowds out workplace association and reduces the threat of revolt.

On the other hand, when the leadership's promises are imperfectly credible, the condition for concessions may fail to hold. Consider the case where the leadership's promises are perfectly incredible, that is, $p = 0$. Then the condition becomes identical to the revolution constraint previously examined, which we know will bind with a sufficiently small bureaucracy. Similarly, as we let the leadership's credibility vary, concessions will not be sufficient to stop a revolution for a sufficiently smaller bureaucracy. Note in particular that as $b^* \to 0$, the condition tends toward $\ln(n/\alpha) - \overline{d} \geq \ln(n/\alpha)$, where the union's maximum dues is $\overline{d} > 0$. In this case, the condition must fail to hold.

We can also use the condition for concessions to illustrate the capacity effect, which is more difficult to observe. It is easiest to see if we rewrite the condition for concessions and hold the crowding-out effect constant, at workers' best revolt payoff, i.e., with bureaucracy at zero. When leaders' promises are perfectly credible and leaders make the best possible concession, the condition is: $n\ln(n/\alpha) - n\underline{d} - (n - \alpha b^*) \geq n\ln(n/\alpha) - n$. After some simplification and rearranging, the condition becomes $\alpha b^* \geq n\underline{d}$. The term on the left-hand side is the effect of bureaucracy on workers' costs of association while the term on the right is the minimum dues level, which is equivalent to the costs of bureaucracy. Substituting in the values for $\underline{d}$ (evaluated at b^*) and b^*, the condition then becomes:

$$\sigma\alpha^2/\beta \geq \sigma\alpha^2/2\beta$$

which must always hold with inequality or at equality when the level of bureaucracy is at zero (e.g., when the share of dues revenue the leadership takes is zero). In other words, as bureaucracy increases the reduction in the costs of workplace association is greater than the increase in costs of bureaucracy. Intuitively, because union bureaucracy is more efficient than workplace association, the leadership can produce the collective good with a decreasing cost as bureaucracy increases. The union's offer to pass on this efficiency to workers in the form of concessions makes those concessions more attractive and the threat of revolt less appealing. Because we have assumed that leaders are perfectly credible, this capacity effect is by itself sufficient to thwart a revolt; however, with imperfect credibility this will not be the case in general.

Since the condition for concessions may or may not hold, we can define a critical value of bureaucracy, b^{**}, such that workers are indifferent between revolt and concessions: $V_W(R, b^{**}) = V_W(O, d_O = \underline{d})$. This can be written more fully as:

$$\ln(n/\alpha - b^*) = \ln(n/\alpha) - \left[p\underline{d} + (1-p)\overline{d}\right] \quad (3)$$

If $b^* < b^{**}$, then even at the best possible dues level, the promises of the leadership are not sufficient to forestall a revolt; that is, $V_W(R) > V_W(O, d_O = \underline{d})$. In order to stop a revolt, the leadership will therefore have to democratize. Democratization is a feasible strategy if democracy leaves workers at least as well off as revolt. This is the case when $V_W(D) \geq V_W(R)$, which is equivalent to:

$$\ln(n/\alpha) - \underline{d} \geq \ln(n/\alpha - b^*)$$

This condition is identical to the condition where the union leadership offers workers concessions with perfectly credibility. We therefore know that this condition will always hold.

On the other hand, when $b^* \geq b^{**}$, then a revolt is sufficiently unrewarding that the leadership can prevent democratization by making dues concessions. In this case, the leadership will set the dues level at the amount which makes workers indifferent between revolting or not; in this case, $V_W(R) = V_W(O, d_O = \hat{d})$.

We can now state succinctly the union-democracy game's equilibrium structure in terms of b:

1. If $n\ln(n/\alpha - b^*) \leq (n - \alpha b^*)$, then the revolution constraint does not bind and the leadership can stay in power without democratizing or changing the dues level.
2. If $n\ln(n/\alpha - b^*) > (n - \alpha b^*)$, the revolution constraint binds. In addition, letting b^* be defined as in (3), then:
   a. If $b^* \geq b^{**}$, the leadership does not democratize and sets the dues level to concede enough to avoid a revolt.
   b. If $b^* < b^{**}$, dues concessions are insufficient to avoid a revolution and the union is democratized.

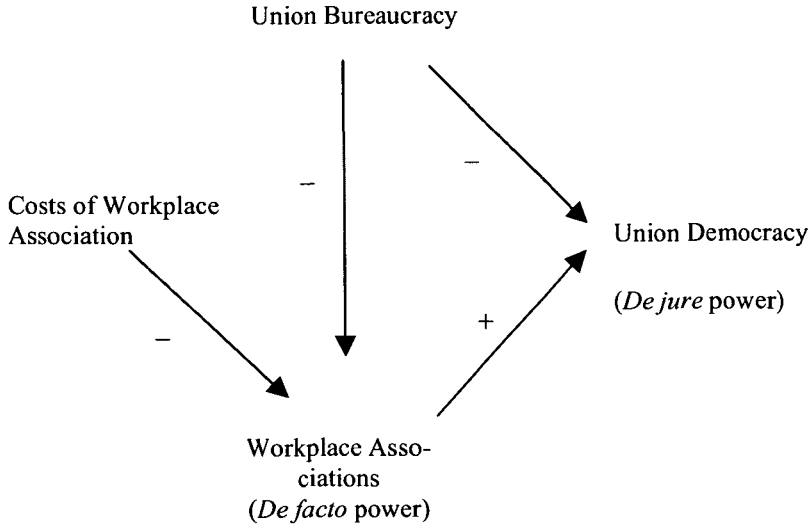


Figure 1. Relationships Among Union Bureaucracy, Workplace Associations, and Union Democracy

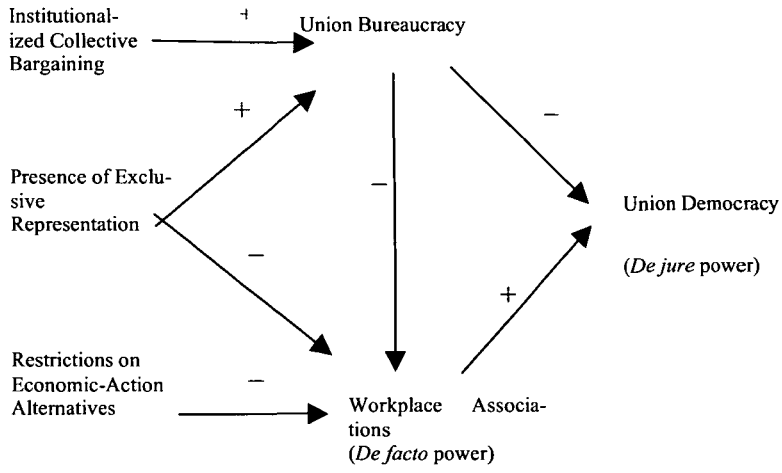


Figure 2. Effects of Legal Rules on Union Democracy

STRENGTH OF WORKPLACE ASSOCIATION	CREDIBILITY OF LEADERSHIP	UNION GOVERNANCE OUTCOME
Low	N/A	Oligarchy
Medium	High	Oligarchy with concessions
High	Low	Democracy

Table 1. Union Governance Outcomes for Given Credibility and Associational Conditions

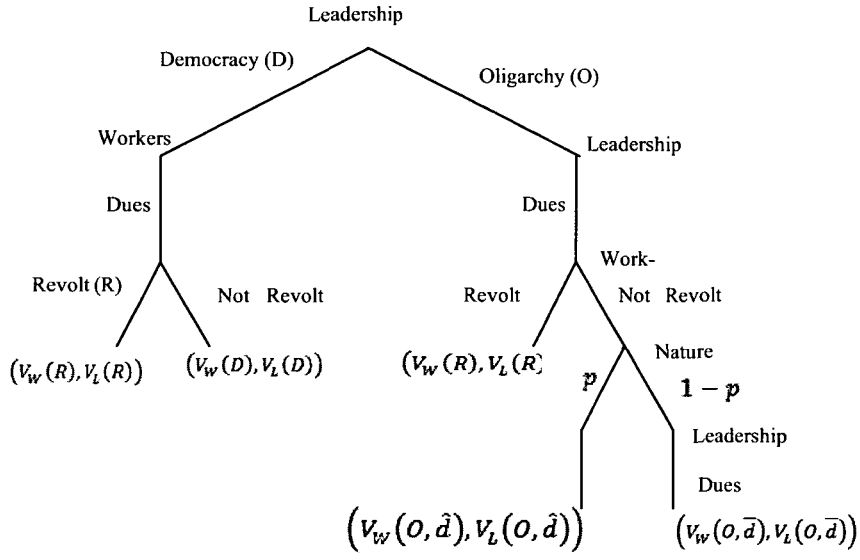


Figure 3. The Union-Democracy Game

ADDRESSING THE SPECIAL PROBLEMS OF MENTALLY ILL PRISONERS: A SMALL PIECE OF THE SOLUTION TO OUR NATION'S PRISON CRISIS

MICHAEL VITIELLO†

INTRODUCTION

After years of neglect, policymakers must confront a crisis in our prisons created by the increasing number of mentally ill prisoners.[1] Mentally ill prisoners are both vulnerable and troublesome. Often out of control, they may need physical restraint, creating a risk to themselves and to prison guards.[2] Other prisoners fear and target the mentally ill, as well.[3]

Apart from their special needs, they are an increasing segment of the prison population.[4] While many mentally ill individuals end up in a nursing home or become homeless, their numbers have risen roughly in proportion with the release of the mentally ill from mental hospitals and the closing of those institutions.[5] Many people who received some form of mental health treatment in those settings are now in prison,[6] where they are unlikely to receive adequate mental health care.[7]

Around the nation, states are looking for ways to reduce prison costs.[8] Various mainstream organizations have been recommending a

---

† Distinguished Professor and Scholar, Pacific McGeorge School of Law; University of Pennsylvania, J.D., 1974; Swarthmore College, B.A., 1969. This paper is based on a presentation that I delivered at the University of Wyoming College of Law. I want to extend special thanks to Wyoming Professor Diane Courselle for inviting me to make that presentation and to my colleague Gerald Caplan for suggesting the topic and then for many helpful conversations about the thesis and comments on an earlier draft. I also want to thank Drs. Claude Arnett and Seymour Moscovitz for their suggestions and insights. I also want to thank my research assistants Lauren Knapp and Brittany Griffith for their help in research and footnoting this article.

1. *See* Coleman v. Schwarzenegger, No. CIV S-90-0520, 2009 WL 2430820, at *12 (E.D. Cal. Aug. 4, 2009).

2. William Kanapaux, *Guilty of Mental Illness*, PSYCHIATRIC TIMES, Jan. 1, 2004, *available at* http://www.psychiatrictimes.com/display/article/10168/47631.

3. Steven K. Hoge, *Providing Transition and Outpatient Services to the Mentally Ill Released from Correctional Institutions*, *in* PUBLIC HEALTH BEHIND BARS: FROM PRISONS TO COMMUNITIES 461, 470 (Robert Greifinger ed., 2007).

4. LANCE T. IZUMI ET AL., PACIFIC RESEARCH INST., CORRECTIONS, CRIMINAL JUSTICE, AND THE MENTALLY ILL: SOME OBSERVATIONS ABOUT COSTS IN CALIFORNIA 3 (1996), *available at* http://www.mhac.org/pdf/PacificResearchStudy.pdf.

5. *See* James Ridgeway & Jean Casella, *Locking Down the Mentally Ill*, THE CRIME REPORT (Feb. 17, 2010, 10:06 PM), http://thecrimereport.org/2010/02/17/locking-down-the-mentally-ill/.

6. Te-Ping Chen, *For Many With Mental Illnesses, Jail's the Only Treatment Option*, CHANGE.ORG (May 12, 2010, 9:23 AM), http://criminaljustice.change.org/blog/view/for_many_with_mental_illnesses_jails_the_only_treatment_option.

7. SASHA ABRAMSKY & JAMIE FELLNER, HUMAN RIGHTS WATCH, ILL-EQUIPPED: U.S. PRISONS AND OFFENDERS WITH MENTAL ILLNESS 110 (2003).

8. *See Cost-Cutting States Reduce Prison Populations: Number of State Inmates Drops For First Time Since 1972*, MSNBC.COM (March 17, 2010, 12:02 AM), http://www.msnbc.msn.com/id/

variety of reforms.[9] In California, the prison system has been subject to federal court litigation for over 20 years.[10] In 2009, a panel of three federal judges found that overcrowding has created health risks—prompting the court to order release of over 40,000 prisoners.[11] California may represent the worst-case scenario, but it is hardly unique. As a result of this national crisis, for the first time in decades, meaningful reform may be in the air.

But if reform takes place, it should be done right. Part of the problem with sentencing generally—as well as the dramatic increase in mentally ill prisoners—is that public policy has been driven by anecdotes and headline cases. As a result, legislation is driven by exaggeration rather than by careful analysis. This is obvious in laws like Three Strikes in California that resulted from the tragic kidnapping, rape and murder of Polly Klaas.[12] Less obvious is how misinformation led to the increase in mentally ill prisoners. And so this Article discusses how the movement to release the civilly committed mentally ill has resulted in the increased number of mentally ill prisoners.[13] The point of that inquiry is to learn some lessons about how we made mistakes.[14] Thereafter I apply those lessons to today's discussions about reforming the prison system as it relates to mentally ill prisoners.[15]

I. GOOD INTENTIONS GO AWRY

So how did we get to where we are today? *One Flew Over the Cuckoo's Nest* should be assigned viewing for anyone attempting to get a quick historical view about the current state of the law governing the mentally ill.[16] In Milos Forman's film, based on Ken Kesey's novel, Jack Nicholson plays a conman who ends up in a mental institution as a way to avoid doing hard labor.[17] Central to the film is his battle against Nurse

---

35903114/ns/us_news-crime_and_courts/39172744.

9. *See generally* MICHAEL E. ALPERT, THE LITTLE HOOVER COMMISSION, SOLVING CALIFORNIA'S CORRECTIONS CRISIS: TIME IS RUNNING OUT (2007), *available at* http://www.lhc.ca.gov/studies/185/Report185.pdf; Michael Vitiello & Clark Kelso, *A Proposal For A Wholesale Reform Of California's Sentencing Practice And Policy*, 38 LOY. L.A. L. REV. 903 (2004); Lauren E. Geissler, Creating and Passing a Successful Sentencing Commission in California (Jan. 27, 2006) (unpublished manuscript), *available at* http://www.law.stanford.edu/program/centers/scjc/workingpapers/LGeissler_06.pdf.

10. *See* Coleman v. Schwarzenegger, No. CIV S-90-0520, 2009 WL 2430820, at *12 (E.D. Cal. Aug. 4, 2009).

11. *See id.* at *115–16.

12. Michael Vitiello, *"Three Strikes" And The* Romero *Case: The Supreme Court Restores Democracy*, 30 LOY. L.A. L. REV. 1643, 1655 (1997).

13. *See infra* Part II.

14. *See infra* Part III.

15. *See infra* Part IV.

16. ONE FLEW OVER THE CUCKOO'S NEST (Fantasy Films 1975); *see also* David Pescovitz, *Cuckoo's Nest Hospital to be Demolished*, BOINGBOING (July 16, 2008, 9:32 AM), http://boingboing.net/2008/07/16/cuckoos-nest-hospita.html (explaining that the author of the original story, Ken Kesey, got many of his ideas from working in a mental institution earlier in his life).

17. ONE FLEW OVER THE CUCKOO'S NEST, *supra* note 16.

Ratched, the person effectively in charge of the mental institution.[18] The film captures several themes: it raises questions about whether those in mental institutions in fact are insane; it suggests that the diagnosis of insanity is in part used to suppress rebels, like Nicholson's character, Randall McMurphy; and it shows the debilitating effects of mental health treatments, including McMurphy's lobotomy.[19]

The film's view of mental illness was hardly unique to Kesey or Forman. It reflected powerful themes that had serious backing in the psychiatric community during that era. Emerging as a serious intellectual force in the 1960s, the "anti-psychiatry" movement challenged the most fundamental assumptions and practices of psychiatry.[20] Many prominent figures led an attack on psychiatry as it was then practiced.[21] Central to their claims were a number of premises. For example, they believed that definitions of many psychiatric disorders are vague and arbitrary, leaving too much room for interpretation by the observer and to too many misdiagnosed patients.[22] And the anti-psychiatrists could point to notorious failures and misuses of psychiatry.[23] The modern anti-psychiatrists argued that illnesses like schizophrenia reflected healthy attempts to cope with a sick society.[24] In effect, the diagnosis of mental illness was society's way to control and limit dissent.[25]

Another premise of the anti-psychiatry movement was that available treatments were far more damaging than helpful.[26] Treatment could be brutal. Existing techniques included electric shock therapy, involuntary

---

18. *Id.*

19. *Id.*

20. *See* EDWARD SHORTER, A HISTORY OF PSYCHIATRY: FROM THE ERA OF THE ASYLUM TO THE AGE OF PROZAC 277 (1997).

21. *Id.* at 274–276 (explaining that among the leaders in the movement were Michael Foucault, Ronald D. Laing, and Erving Goffman).

22. Heap v. Roulet (*In re* Estate of Roulet), 590 P.2d 1, 10–11 (Cal. 1979).

23. *See, e.g.*, THOMAS SZASZ, SCHIZOPHRENIA: THE SACRED SYMBOL OF PSYCHIATRY 152–53 (1976) (citing the ability of husbands to have their wives committed for disobedience despite their wives' sanity); SHORTER, *supra* note 20, at 303–04 (explaining that anti-psychiatrists could also point to the American Psychiatric Association's inclusion of homosexuality as a form of mental illness until the 1970's); Richard J. Bonnie & Svetlana V. Polubinskaya, *Unraveling Soviet Psychiatry*, 10 J. CONTEMP. LEGAL ISSUES 279, 279 (1999) (explaining the Soviet's use of mental institutions to deal with political opponents of the state); Ariela Gross, *Pandora's Box: Slave Character on Trial in the Antebellum Deep South*, 7 YALE J.L. & HUMAN. 267, 293 (1995) (explaining the 18th century diagnosis of a mental disease afflicting some slaves whose symptoms included their tendency to escape their masters).

24. SHORTER, *supra* note 20, at 276.

25. Bonnie & Polubinskaya, *supra* note 23, at 279 (explaining that the anti-psychiatry movement coincided with opposition to the Vietnam War and to civil rights and women's rights movements); *see* E. FULLER TORREY, OUT OF THE SHADOWS: CONFRONTING AMERICA'S MENTAL ILLNESS CRISIS 142 (1997) [hereinafter OUT OF THE SHADOWS] (explaining that a new generation of lawyers emerged with an interest in civil liberties and borrowed strategies from other civil rights litigation as well); Michael E. Staub, *Madness is Civilization: Psycho Politics and Postwar America* 4 (School Soc. Sci., Occasional Paper No. 34, 2008), *available at* http://www.sss.ias.edu/files/papers/paper34.pdf (explaining that as a result, claims that the mentally ill were victims of a sick society gained credibility).

26. *See* SHORTER, *supra* note 20, at 208.

commitment for long periods of time with few constraints, and lobotomies—often leaving the patient catatonic.[27] Combine those invasive practices with famous cases of misdiagnosis of different kinds. In some instances, a patient suffering from one mental illness was diagnosed with a different illness.[28] Even more frightening were cases where a perfectly sane individual was involuntarily committed and kept committed for a prolonged period of time.[29]

The system was certainly broken. Peaking in 1956, the population housed in state and local public mental health hospitals was about 560,000.[30] Many were warehoused in state institutions described as "snake pits," where they were at the mercy of poorly trained staff, which lacked adequate resources.[31] Back when Geraldo Rivera was a serious investigative reporter, he, among others, got the public's attention with exposés of the terrible conditions in mental institutions.[32]

This period was the setting for a dramatic expansion of the rights of the mentally ill and for the movement that led to deinstitutionalizing mental health care. Change came through various legislation and many lawsuits, several of which ended in the Supreme Court.[33] Several important principles emerged that expanded the rights of the mentally ill.[34] The net result was that involuntary civil commitment and compelled medica-

27. Sheldon Gelman, *Looking Backward: The Twentieth Century Revolutions in Psychiatry, Law, and Public Mental Health*, 29 OHIO N.U. L. REV. 531, 531–32 (2003).

28. *See* Heap v. Roulet (*In re* Estate of Roulet), 590 P.2d 1, 10–11 (Cal. 1979).

29. *See* SZASZ, *supra* note 23, at 149–51.

30. MICHAEL PUISIS, CLINICAL PRACTICE IN CORRECTIONAL MEDICINE 33 (2d ed. 2006) (stating that by comparison, today, there are about 80,000 people committed to such institutions).

31. *Psychiatry: Out of the Snake Pits*, TIME, Apr. 05, 1963, *available at* http://www.time.com/time/magazine/article/0,9171,830082-1,00.html.

32. *See* WILLOWBROOK: THE LAST DISGRACE (ABC 1972).

33. *See, e.g.*, O'Connor v. Donaldson, 422 U.S. 563 (1975).

34. *See* Wyatt v. Stickney, 325 F. Supp. 781, 785–86 (1971). For example, mentally ill patients who are involuntarily committed have due process interests in conditions of reasonable care and safety and reasonably nonrestrictive confinement conditions. They have the right to a range of services, including the right to treatment in a community setting. *O'Connor*, 422 U.S. at 574–76. Further, the Court has found that it is unconstitutional to detain someone involuntarily if that person is not a danger to himself or to others. Thus, a finding of mental illness, without more, does not justify continued confinement even if appropriate treatment is available. *Id.* at 575. Both lower federal courts and the Supreme Court have limited the state's ability to administer psychotropic medication in any setting. Involuntarily committed mental patients have a right to make their own treatment decisions and may not be forcibly medicated (subject to limited circumstances, notably emergencies and periods of incompetence). *See* Washington v. Harper, 494 U.S. 210, 221 (1990). An institution's decision to medicate is not justified solely on a finding that the patient is incompetent. The decision to medicate requires additional litigation and a specific finding that the patient is incompetent to make that decision for herself. *Id.* at 228. In the more recent past, some states have cut back on the rights of the mentally ill, often in reaction to a violent crime committed by a mentally ill individual. For example, New York enacted "Kendra's Law," N.Y. MENTAL HYG. LAW § 9.60 (McKinney 2010), after a schizophrenic man pushed a young woman onto subway tracks, leading to her death. PATRICIA E. ERICKSON & STEVEN K. ERICKSON, CRIME, PUNISHMENT, AND MENTAL ILLNESS: LAW AND THE BEHAVIORAL SCIENCES IN CONFLICT 23–25, 45–46 (2008).

tion became far more difficult.[35] Many of the same protections apply to mentally ill prisoners as well.[36]

Not only have the mentally ill gained legal protection, but at the same time, we experienced a movement away from publicly funded state mental institutions.[37] That change was not inevitable, but flowed from the horrible exposure of conditions in those institutions. Even those revelations may not have resulted in the closing of many of those institutions. After all, revelations about horrible prison conditions did not lead to closing those facilities.[38] But as indicated earlier, inspired in part by the anti-psychiatry movement, numerous reformers believed, in effect, that many mentally ill individuals were rebels against an oppressive society and that the state used mental institutions to suppress dissent.[39]

And not all of those interested in closing mental institutions were disability rights activists. In California, in the late 1960s, then-Governor Ronald Reagan signed legislation that paralleled developments elsewhere, and made involuntary commitment extremely difficult.[40] Mentally disabled rights activists called the California legislation "the Magna Carta of the mentally ill" and saw it as a step towards an eventual goal of eliminating involuntary commitment altogether.[41] As a result of the deinstitutionalization movement, mentally ill patients who were released from mental health facilities were sent back into their communities.[42]

---

35. *See* CAL. WELF. & INST. CODE § 5150 (West 2010). For example, under California's law, commitment was no longer justified simply based on a showing of the need for treatment but instead required a showing that the person was a danger to himself or to others. *Id.*

36. In 1990, the Supreme Court held that correction officials can administer such medication in compelling circumstances but cannot do so arbitrarily. *Washington*, 494 U.S. at 221. Thus, the state must show that the prisoner is gravely disabled or is a danger to himself or others. Under the Court's case law, an inmate has a right to refuse psychotropic medication under most circumstances. The net result of these various cases is a set of important procedural rights that make involuntary commitment and treatment difficult to compel.

37. Alfred Auerback, *The Short-Doyle Act: California Community Mental Health Services Program: Background and Status After One Year*, CAL. MED., May 1959, at 335, *available at* http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1577700/pdf/califmed00113-0095.pdf.

38. *See* Margaret Winter & Stephen F. Hanlon, *Parchman Farm Blues: Pushing for Prison Reforms at Mississippi State Penitentiary*, 35 LITIG. 1, 1–8 (2008), *available at* http://www.aclu.org/images/asset_upload_file829_41138.pdf (explaining that instead, for example, in prison litigation in the south, court supervision led to markedly improved conditions in notorious prisons like Parchman and Angola prisons in Mississippi and Louisiana).

39. Bonnie & Polubinskaya, *supra* note 23, at 279.

40. CAL. WELF. & INST. CODE § 5150 (West 2010).

41. E. Fuller Torrey & Kenneth Kress, *The New Neurobiology of Severe Psychiatric Disorders and Its Implications for Laws Governing Involuntary Commitment and Treatment* 51 (Bepress Legal Series Working Paper No. 423, 2004), *available at* http://law.bepress.com/expresso/eps/423; *see also* OUT OF THE SHADOWS, *supra* note 25, at 143–144. As with many political coalitions, not all of those who supported making civil commitment more difficult did so out of concern for the mentally ill. Some proponents of the legislation saw it as a way to reduce costs to the state.

42. *See* Antonia Moras, *Human Rights Watch: The Mentally Ill in U.S. Prisons*, ALASKA JUST. F., Spring 2004, at 2, 2, *available at* http://justice.uaa.alaska.edu/forum/21/1spring2004/b1_mentallyill.html. As observed by one author:

> State incentives for cost-shifting to the federal government reside almost exclusively in the discharge of patients from state hospitals, who then become eligible for SSI, Medicaid, food stamps, and other federal benefits. States gain nothing by ensuring that patients

The promise at the time was that community-based care would allow the mentally ill greater freedom without abandoning them to their own devices.[43]

So what went wrong? Closing institutions seemed humane and community-based care seemed like a sound way to treat the mentally ill. Adequately funded community based programs have worked: many patients see a dramatic improvement in their quality of life; many are able to hold steady employment and find housing.[44] However, in most places the development of the community-based programs lagged far behind the demand created by the release of the mentally ill.[45] The lack of adequate resources for community-based care has only grown worse over time—especially since states have confronted serious budget crises brought on by the recession.[46] As described below, these reforms, even with the best intentions, have come at a high cost to many mentally ill persons.

II. THE REVOLVING DOOR

Today, most state mental hospitals have closed or dramatically reduced available beds.[47] But what happens to the mentally ill? Since the elimination of most beds in state-run facilities, and the cutting of community health care resources offers a dramatic contrast to the world envisioned by the anti-psychiatrists and mental health care advocates, the result of many of the reformists' efforts have come at a cost to the mentally ill.

---

receive follow-up care following their hospitalization because readmission of the patients can be deflected to the psychiatric wards of general hospitals, where federal Medicaid will cover much of the costs.

OUT OF THE SHADOWS, *supra* note 25, at 102. Thus, the way in which federal funds are made available to the states provides states an incentive to discharge patients whether or not they are able to function on their own and to do so without regard to available aftercare.

43. *See* PHIL BROWN, THE TRANSFER OF CARE: PSYCHIATRIC DEINSTITUTIONALIZATION AND ITS AFTERMATH 67 (1985).

44. *See, e.g.*, *Direct Access to Housing*, CORP. FOR SUPPORTIVE HOUSING, http://www.csh.org/index.cfm?fuseaction=Page.viewPage&pageId=501 (last updated Aug. 2005).

45. H. Richard Lamb & Leona L. Bachrach, *Some Perspectives on Deinstitutionalization*, 52 PSYCHIATRIC SERVICES 1039, 1044 (2001), *available at* http://psychservices.psychiatryonline.org/cgi/reprint/52/8/1039. Some of the additional freedoms that the mentally ill gained have exacerbated the problem. Many mentally ill persons refuse medication that might otherwise enable them to live more stable lives and to stay out of trouble with the law. *Id.* at 1041.

46. *See* Rusty Selix, *State Budget Memorandum*, CAL. COUNCIL COMMUNITY MENTAL HEALTH AGENCIES (Jan. 10, 2008), http://www.cccmha.org/public_policy/state_budget.html; *see also* CAL. COUNCIL CMTY. MENTAL HEALTH AGENCIES, PRESERVE AB 2034 FUNDING: A MODEL PROGRAM THAT WORKS AND HAS CHANGED LIVES (2008), *available at* http://www.cccmha.org/documents/AB2034FACTSHEET--ProgramthatWorks.pdf. For a period of time, legislation made available federal matching grants for community health programs, including mental health care. California initially followed suit, but in the 1990's, it shifted the burden of responsibility for funding to local governments. For a time, it had in place pilot programs that were highly successful in reducing incarceration and homelessness among the mentally ill. But those programs were eliminated when budget cuts were made in 2007.

47. Hitesh C. Sheth, *Deinstitutionalization or Disowning Responsibility*, 13 INT'L J. PSYCHOSOCIAL REHABILITATION, no. 2, 2009 at 11, *available at* http://www.psychosocial.com/IJPR_13/Deinstitutionalization_Sheth.html.

The effect has been a change of venue for the mentally ill from mental hospitals to prisons, not just to nursing homes or the streets. While there are few data on incarcerations of mentally ill people prior to the deinstitutionalization movement,[48] evidence suggests that, since deinstitutionalization, the rate of incarceration of mentally ill people has increased significantly.[49] While estimates vary, studies are consistent that large numbers of those admitted to prison are mentally ill.[50] When states closed or reduced the population of mental health facilities, the prison system took in those mentally ill patients who required twenty-four hour supervision.[51] Due to the lack of community programs and adequate and affordable housing for the mentally ill patients who were released from the institutions, many of those released wound up homeless.[52] Because of a general public fear of those with mental illness, law enforcement was pressured into arresting and incarcerating the homeless mentally ill for petty crimes, such as public intoxication.[53] Further, illegal drug use among mentally ill people is common.[54] Mentally ill individuals often self medicate.[55] As a result, many of the mentally ill people living in a community—who would have once been institutionalized—are arrested for behavior that they engage in as a result of their illness.[56]

Further, unable to get adequate resources for mental health care treatment in state run institutions or community health care facilities, mentally ill individuals in prison have their symptoms exacerbated by being put in jail or prison, causing them to act out.[57] Prisons are seldom good places to receive mental health care treatment.[58]

Mentally ill inmates who are released have a difficult time getting into community mental health programs and public housing because of their criminal records.[59] Thus, for those who are released from prison, it becomes a vicious cycle of homelessness, to imprisonment, and back to homelessness. Without adequate treatment to allow the mentally ill to adapt to living in the community, many end up back in prison.[60]

---

48. Lamb & Bachrach, *supra* note 45, at 1042.

49. *Id.*

50. *Human Rights at Home: Mental Illness in U.S. Prisons and Jails, Hearing Before the Subcomm. on Human Rights and the Law of the S. Comm. on the Judiciary*, 111 Cong. (2009) (statement of Gary Maynard, Secretary of the Maryland Department of Public Safety and Correctional Services).

51. Lamb & Bachrach, *supra* note 45, at 1042.

52. *Id.* at 1040.

53. *See id.* at 1042.

54. *Id.* at 1041.

55. *Id.*

56. *See id.* at 1042.

57. Allan Schwartz, *Imprisoning the Mentally Ill*, MENTALHELP.NET, http://www.mentalhelp.net/poc/view_doc.php?type=doc&id=14284 (last updated Jan. 14, 2008).

58. *See* Kanapaux, *supra* note 2.

59. *Id.*

60. *See id.*; *see also* OUT OF THE SHADOWS, *supra* note 25, at 108.

III. LESSONS LEARNED?

California may be forced to reduce its overcrowded prison population. Reform may be possible for the first time in years because a three-judge panel has ordered California to reduce its prison population by about 40,000 inmates.[61] That may force California to come to terms with its bloated prison system.[62]

The Supreme Court has granted certiorari to review the order of the three-judge panel.[63] As is typical of this closely divided Court, predicting how it will resolve the dispute is a crapshoot. But we may be in familiar territory. As Adam Liptak wrote, the Constitution means what Justice Kennedy says it means.[64] Despite strong conservative leanings, Justice Kennedy may vote to uphold the order. For example, even after voting to uphold two sentences under California's Three Strikes law,[65] Justice Kennedy has been a vocal critic of mandatory minimum sentencing and the overuse of prisons.[66] He also authored a number of majority opinions striking down the death penalty[67] and, more recently, an opinion striking down true life sentences for offenders who were juveniles when they committed offenses other than homicide.[68] As a result, the conservative wing of the Court cannot count on his vote on criminal justice issues.

If the Supreme Court upholds the federal district court order, reform will have to take place, and California will need to find less expensive ways to handle prisoners generally and the mentally ill specifically.

So what lessons should policy-makers take from history? The reforms of the past several decades were suitable if the then-popular assumptions were true. As discussed above, those assumptions included the

---

61. Coleman v. Schwarzenegger, No. CIV S-90-0520, 2009 WL 2430820, at *115–16 (E.D. Cal. Aug. 4, 2009).

62. *See id.* The state has taken an aggressive litigation posture. It attempted to have the prisoner receiver removed, but was rebuffed by the Ninth Circuit. Julie Small, *Court Upholds Federal Oversight of California's Prison Medical Care*, S. CAL. PUB. RADIO (Apr. 30, 2010), http://www.scpr.org/news/2010/04/30/receiver-stands/. The state has also petitioned, now twice, to have the three judge panel's order overturned. Schwarzenegger v. Plata 130 S. Ct. 1140, 1140 (2010). If the Court finds that the three judge panel exceeded its authority, reform may be dead. The litigation may be the state's last-best hope for meaningful reform of its prison system. The legislature's response to prison overcrowding and massive spending on its prison system has been discouraging. For example, the senate passed a bill that included a sentencing commission, but the Democratic-controlled assembly refused to go along. Jack Chang, *Sentencing Panel Sets Off Alarms*, SACRAMENTO BEE, Aug. 20, 2009, at 1A, *available at* http://www.sacbee.com/2009/08/20/2124062/sentencing-panel-sets-off-alarms.html.

63. Schwarzenegger v. Plata, 130 S. Ct. 3413, 3413 (2010).

64. Adam Liptak, *Anthony M. Kennedy*, N.Y. TIMES, http://topics.nytimes.com/top/reference/timestopics/people/k/anthony_m_kennedy/index.html (last updated July 1, 2009).

65. Ewing v. California, 538 U.S. 11, 14, 30–31 (2003); Lockyer v. Andrade, 538 U.S. 63, 66, 77 (2003).

66. Pete Williams, *Justice Anthony M. Kennedy: End Minimum Sentences*, THE NOVEMBER COALITION (Aug. 9, 2003), http://www.november.org/stayinfo/breaking/Kennedy.html.

67. *See, e.g.*, Kennedy v. Louisiana, 554 U.S. 407 (2008); Roper v. Simmons, 543 U.S. 551 (2005).

68. Graham v. Florida, 130 S. Ct. 2011, 2034 (2010).

belief that diagnoses were routinely wrong,[69] that the mentally ill were capable of easy integration into the community,[70] and that psychotropic drugs and other treatments were dehumanizing,[71] and that institutions were so bad that they had to be abandoned.[72]

And all of those assumptions were true, but only to a point. Those who work with the mentally ill and the families of the mentally ill will tell you that the diseases are real and that adequate care can improve the quality of their lives.[73] And ask any family member of a mentally ill person whether today's system works well—many would describe their frustration in getting access to basic mental health care services.[74] Further, policymakers were unable to work through the unintended consequences of their decisions. That is, they did not recognize that they were basing policy on an incomplete view of the mentally ill and made overly optimistic assumptions about the ability for the mentally ill to live on their own without state supervision. They did not recognize the revolving door from homelessness to jail and prison to homelessness and back.[75]

Reformers should focus on these lessons of experience. As developed below, we have learned a great deal about mental illness and the needs of the mentally ill.[76] Applying current data should allow a more realistic approach to caring for the mentally ill.

IV. THE SHAPE OF REFORM

As indicated above, California may be forced to affect a reform of its prison system.[77] Part of that reform should focus on the special problems of mentally ill prisoners. Because of California's budget crisis,[78] anyone who comes forward with a proposal for reform must demonstrate that it will save the system money. Even given that constraint, this section argues that meaningful reform is possible.

As currently delivered, mental health care for prisoners is expensive and ineffective.[79] Treating the mentally ill in a variety of settings, like

---

69. Heap v. Roulet (*In re* Estate of Roulet), 590 P.2d 1, 10–11 (Cal. 1979).

70. *See* BROWN, *supra* note 43, at 67.

71. *See supra* text accompanying notes 26–27.

72. *See supra* text accompanying notes 31–32.

73. *Mental Illnesses*, NAT'L ALLIANCE ON MENTAL HEALTH, http://www.nami.org/Content/NavigationMenu/Inform_Yourself/About_Mental_Illness/About_Mental_Illness.htm (last visited Dec. 29, 2010).

74. *See* MARY BETH PFEIFFER, CRAZY IN AMERICA: THE HIDDEN TRAGEDY OF OUR CRIMINALIZED MENTALLY ILL 159–160 (2007).

75. I assume that they did not recognize those consequences because who would have chosen today's response to the mentally ill had they been able to foresee where we have ended up?

76. *See infra* Part VI.

77. Aaron Rappaport & Kara Dansky, *State of Emergency: California's Correctional Crisis*, 22 FED. SENT'G REP., no. 3, 2010, at 133.

78. Dan Walters, *Overview of California's Budget Crisis*, SACRAMENTO BEE (July 21, 2009, 12:50 PM), http://www.sacbee.com/2009/07/21/2044072/overview-of-californias-budget.html.

79. RISDON N. SLATE & W. WESLEY JOHNSON, THE CRIMINALIZATION OF THE MENTALLY ILL: CRISIS AND OPPORTUNITY FOR THE JUSTICE SYSTEM 289–296 (2008).

community-based facilities, is far less expensive than is warehousing them in prison and even less expensive than maintaining them in prison with adequate mental health care services.[80] Thus, using alternative settings for the mentally ill may be an effective alternative to incarceration.

If state officials adopt reforms that would enable a shift of mentally ill prisoners from prisons to community care facilities, they must do so in ways that protect the public. Here, they must fully appreciate the lessons from the past. As discussed above, policy makers and the public in the 1960s and beyond had a naïve view of mental illness.[81] They bought into stereotypes about the ability of the mentally ill to live independent lives. When many mentally ill failed to conform to reformers' hopes, we experienced a backlash that has resulted in the current situation where a person is more likely to receive mental health care in prison than in the community.[82] In effect, society replaced one stereotype of the mentally ill with other stereotypes. Thus, today many members of the public view the mentally ill as incapable of cure[83] or as malingerers,[84] individuals in need of punishment.

Any change in policy towards the mentally ill must be grounded in reality, rather than stereotypes. While providing care for the mentally ill in community-based treatment facilities can save the state money, not all mentally ill prisoners are capable of being reintegrated into society.[85]

To this point, I have spoken of mentally ill prisoners without making an essential distinction between two distinct kinds of mentally ill prisoners. Many criminals suffer from an assortment of mental illnesses, but would continue to violate the law even if they received adequate treatment.[86] Indeed, one suspects that treatment might make them more capable of carrying out criminal acts. By comparison, our prisons now house many prisoners whose mental illness has led to their criminal conduct.[87]

---

80. Mental Health Servs, Oversight & Accountability Comm'n, *Commission Meeting Minutes*, CA.GOV, 9 (June 26, 2008), http://www.mhsoac.ca.gov/Meetings/docs/Meetings/2008/Jul/MHSOAC_June08MeetingMinutes_2.pdf.

81. ERICKSON & ERICKSON, *supra* note 34, at 25.

82. John Gunn, *Future Directions for Treatment in Forensic Psychiatry*, 176 BRIT. J. PSYCHIATRY 332, 333 (2000).

83. Rohan Ganguli, *Mental Illness and Misconceptions*, POST-GAZETTE.COM (Mar. 18, 2000), http://www.post-gazette.com/forum/20000318gang1.asp.

84. SLATE & JOHNSON, *supra* note 79, at 290.

85. *Human Rights at Home: Mental Illness in U.S. Prisons and Jails, Hearing Before the Subcomm. on Human Rights and the Law of the S. Comm. on the Judiciary*, 111 Cong. (2009) (statement of Harley G. Lappin, Dir. of Fed. Bureau of Prisons).

86. Historically, mental health experts considered sociopaths and psychopaths as difficult, if not impossible to treat. CHARLES H. KNICKERBOCKER, HIDE-AND-SEEK: THE EFFECT OF MIND, BODY, AND EMOTION ON PERSONALITY AND BEHAVIOR IN OURSELVES AND OTHERS 90 (1967). Today, some researchers contend that even those mental illnesses are treatable. Randall Parker, *Psychopathic Brain Driven to Seek Rewards*, FUTUREPUNDIT (March 14, 2010, 11:14 AM) http://www.futurepundit.com/archives/007018.html.

87. OUT OF THE SHADOWS, *supra* note 25, at 39–40.

Many mentally ill individuals enter the criminal justice system because of drug abuse, often their way of self-medicating.[88] They may commit petty property crimes to feed themselves or to get money to buy drugs.[89] When delusional or disoriented, they may act in ways that frighten members of the public.[90] The literature is full of accounts of mentally ill individuals who end up in conflict with law enforcement agents.[91] Those confrontations may result from the person urinating in public or engaging in other antisocial conduct.[92] Otherwise non-violent, the mentally ill individual may resist arrest or otherwise challenge the police officer's authority.[93] Assaulting an officer may result in serious felony charges.[94]

In addition, these offenders are less able to deal with prison. Prisons require rigid rules and adherence to those rules.[95] They are more likely than other offenders to be written up for violations of prison rules.[96] But disoriented mentally ill inmates cannot understand the rules leading to what guards see as defiance and sometimes leading to guards using physical force against them.[97] They often end up in solitary confinement, making their illness worse.[98] As a result of their disruptive behavior, they tend to serve longer prison sentences than other offenders.[99] They may also be victimized by fellow inmates.[100] Suicide rates for mentally ill prisoners are high.[101] As quoted by one author, "the bad and the mad just don't mix."[102]

Reform efforts should focus on this group of mentally ill prisoners. As a matter of decency, the state should not subject them to the brutal conditions of prison, so ill-suited to their needs. Placing them in community-based care facilities would serve their needs far better than they are served in prison and the state would save money by doing so.

Such a proposal, however, begs other questions. First, one might appropriately ask about high rates of recidivism among mentally ill[103]

---

88. *Id.* at 35

89. MARCUS NIETO, CAL. RESEARCH BUREAU, MENTALLY ILL OFFENDERS IN CALIFORNIA'S CRIMINAL JUSTICE SYSTEM 4 (1999).

90. *See* OUT OF THE SHADOWS, *supra* note 25, at 38.

91. SLATE & JOHNSON, *supra* note 79, at 83, 109–177.

92. OUT OF THE SHADOWS, *supra* note 25, at 37–38.

93. *See* PFEIFFER, *supra* note 74, at 120–121.

94. *See* CAL. PENAL CODE § 243(c)(2) (West 2010).

95. SLATE & JOHNSON, *supra* note 79, at 60.

96. *Id.* at 60–61.

97. OUT OF THE SHADOWS, *supra* note 25, at 31.

98. SLATE & JOHNSON, *supra* note 79, at 295.

99. *Id.* at 60-61.

100. JOHN PARRY, CRIMINAL MENTAL HEALTH AND DISABILITY LAW, EVIDENCE AND TESTIMONY 27 (2009).

101. OUT OF THE SHADOWS, *supra* note 25, at 33.

102. *Id.* at 32.

103. SLATE & JOHNSON, *supra* note 79, at 197.

and why we should risk continued criminality among this group of offenders.

Here, a close look at how this group of individuals ends up in a cycle of release from prison back to the streets and back to prison helps to explain how adequate follow-up care can reduce recidivism. Unlike the overly optimistic view of the mentally ill that led to deinstitutionalization,[104] many mentally ill persons cannot function adequately merely left to their own devices. Currently, many mentally ill prisoners are stabilized on medication before their release from prison.[105] At discharge, they are given a small supply of medication and told to follow up with public health officials to receive more.[106] That may be the extent of follow-up that they receive upon release.

Even if they find some kind of housing, many recently released prisoners run out of medication and are too disorganized to continue treatment[107] or choose to go off medication.[108] As a result, they may be evicted from their housing or otherwise choose to go back on the street.[109] Once homeless, they often find themselves in conflict with law enforcement again and back into the criminal justice system.[110]

At least for individuals who are going to be placed on parole, one obvious solution is to make continued compliance with a regimen of therapy and medication a condition of release.[111] Further, the state needs to stop releasing the mentally ill back into the community without resources. Instead, it needs to expand various housing options for the mentally ill where their compliance with terms of release can be enforced.[112] For individuals not yet in prison, similar rules should be put in place that would allow alternative disposition of charges against the mentally ill.[113] That is, the state should expand the options open to sentencing judges to place the mentally ill in appropriate facilities where they can be monitored, but where they are not subject to the dehumanizing conditions that they would otherwise face in prison.[114]

Some advocates for the mentally ill might object to restrictive terms of release.[115] But given the current state of the law, the options are lim-

104. ERICKSON & ERICKSON, *supra* note 34, at 25.
105. THE RELEASED (PBS Home Video 2009).
106. *Id.*
107. *Id.*
108. *See* PFEIFFER, *supra* note 74, at 25.
109. THE RELEASED, *supra* note 105.
110. *Id.*
111. OUT OF THE SHADOWS, *supra* note 25, at 160–61. Studies demonstrate that conditional release increases individuals' compliance with treatment plans, including continued use of medication, and reduces their violent behavior. *See id.*
112. SLATE & JOHNSON, *supra* note 79, at 183–97.
113. *Id.* at 131–34, 156. Some jurisdictions already have in place mental health courts. Studies suggest that these courts have better outcomes than would occur otherwise.
114. PARRY, *supra* note 100, at 191–92.
115. OUT OF THE SHADOWS, *supra* note 25, at 162.

ited: untreated, the individual is likely to end up in prison again. That option is far less desirable than imposing lesser limitations on the individual's autonomy.

My proposal begs two additional closely related questions. Does such a proposal adequately protect the public? And can we really distinguish between the bad and the mad or those who are mentally ill who would continue to commit dangerous criminal act and those whose untreated mental illness is responsible for their criminal conduct?

A great deal is at stake. As I developed above, misperceptions about the mentally ill led to the current state of affairs, with large numbers of mentally ill persons in prison.[116] If policymakers fail to learn the lessons from our earlier experience with deinstitutionalization, we will simply end up with the inhumane and costly alternative of dealing with the mentally ill in our prisons. Releasing dangerous mentally ill persons into the community who commit violent crimes will quickly undo any reform efforts.[117]

In partial answer to the first question, the mentally ill are not typically violent, despite sensationalized reports in the media.[118] And that is especially true if the individual receives adequate follow-up care.[119]

The related question is whether we are able to distinguish between those who get involved in the criminal justice system as a result of inadequately treated mental illness and those who are likely to continue to pose a risk of harm even if treated. Or, as argued by the anti-psychiatrists, is the state of the art inadequate to make accurate diagnoses of mental illness?

A great deal has changed over recent decades. At a minimum, data collection is more sophisticated than in the past. In the area of criminal sentencing, for example, advocates of evidence-based sentencing have demonstrated that predictions about future criminal conduct are increasingly reliable.[120] Researchers have developed testing instruments that measure traits like the inability to feel remorse and the individual's level of impulsivity.[121] Researchers have also been able to determine factors

---

116. Gunn, *supra* note 82, at 333.

117. *Sacramento Early Release Inmate Kevin Peterson Arrested for Attempted Rape: Said Release Wasn't A "Bad Deal"*, NEWS 10 (Feb. 3, 2010), http://www.news10.net/news/local/story.aspx?storyid=74615.

118. PARRY, *supra* note 100, at 23–24.

119. Liesel J. Danjczek, *The Mentally Ill Offender Treatment and Crime Reduction Act and Its Inappropriate Non-violent Offender Limitation*, 24 J. CONTEMP. HEALTH L. & POL'Y 69, 103 (2007).

120. ROGER K. WARREN, NAT'L CTR. FOR STATE COURTS, EVIDENCE BASED PRACTICE TO REDUCE RECIDIVISM: IMPLICATIONS FOR STATE JUDICIARIES 2 (2007); Richard E. Redding, *Evidence Based Sentencing: The Science of Sentencing Policy and Practice*, 1 CHAPMAN J. CRIM. JUST. 1, 5–6 (2009).

121. *See generally* Kent A. Kiehl et al., *An Event Related Potential Investigation of Response Inhibition in Schizophrenia and Psychopathy*, 48 BIOLOGICAL PSYCHIATRY 210 (2000).

that predict violent behavior among the mentally ill.[122] Further, studies of the brain through various kinds of measurements have generated knowledge that we have lacked in the past. For example, using an MRI allows measurement of changes in the structure and function of the brains of the mentally ill, allowing a health care professional to determine objectively that the person is suffering from mental illness.[123]

Not only has our ability to diagnosis mental illness improved, but treatment has improved as well. Lobotomies and electric shock treatments as administered up until the 1970s are no longer routine.[124] The availability of Thorazine in the 1950s aided the movement to deinstitutionalize the mentally ill,[125] but proved less effective than hoped for the mentally ill because of its debilitating effects.[126] While some individuals experience side effects from psychotropic drugs,[127] they may be reduced by adjusting the dosage[128] or by finding an alternative medication.[129] Further, newer medications may be more acceptable because of different side effects.[130]

It would also be a mistake to think that medication alone is the answer to the problem posed by mentally ill prisoners. Some studies raise questions about the effectiveness of many medications that have been touted by psychiatrists and the pharmaceutical companies as miracle cures.[131] Many mental health care professionals recognize that the best outcomes require treatment in combination with medication.[132] Availability to adequate therapy, as envisioned when our society began closing state hospitals, remains an essential component to any meaningful reform.

---

122. OUT OF THE SHADOWS, *supra* note 25, at 53 (stating that "overwhelming evidence" demonstrates that "a small subgroup of the mentally ill have a propensity toward violence," and also that "a persons' past history of violence, concurrent abuse of drugs and alcohol, and failure to take medications are risk factors for violent behavior").

123. *Id.* at 4.

124. M. PADOLINA & C. SANCHEZ, COUNSELING AND PSYCHOTHERAPY: THEORIES, TECHNIQUES AND APPLICATIONS 197 (1997); LINDA GASK, A SHORT INTRODUCTION TO PSYCHIATRY 18 (2004). Since almost disappearing as a method of treatment, electric shock therapy reemerged in the 1990's. In 1999, the Surgeon General endorsed it. About one-hundred thousand patients a year receive electric shock therapy in the United States. As one author states, "the treatment has been refined and made gentler by lowering the amount of electricity delivered and changing where the scalp the leads are placed." DANEIL J. CARLAT, UNHINGED 167 (2010).

125. OUT OF THE SHADOWS, *supra* note 25, at 8.

126. ROBERT WHITAKER, MAD IN AMERICA 147–159 (2010).

127. SLATE & JOHNSON, *supra* note 79, at 58 (noting that side effects include, "dry mouth, weight gain, tiredness, and depression . . . [a]ntipsychotic medications may also cause Akathisia, Dystonia, Parkinsonianism, Tardive Dyskinesia, and Arganulocytosis").

128. NAT'L INST. OF MENTAL HEALTH, MENTAL HEALTH MEDICATIONS 12 (2010).

129. *Id.*

130. *See* OUT OF THE SHADOWS, *supra* note 25, at 5–6.

131. For a particularly disturbing view of America's belief in the silver bullet theory of such drugs, see generally ROBERT WHITAKER, ANATOMY OF AN EPIDEMIC (2010).

132. *See, e.g.*, CARLAT, *supra* note 124.

Thus, as part of a larger reform of California's prison system, addressing the special problems of the mentally ill may be a way to save the state money and improve the quality of the lives of many individuals who would otherwise do hard time in prison.

CONCLUSION

At the outset, I argued that the deinstitutionalization movement began with some truths, like the dehumanizing conditions in state institutions and inaccurate diagnoses, but that reforms were based on exaggerations of those truths.[133] As a result, the cure created a new set of problems that now confront policymakers.[134] Today's policymakers should avoid the same kind of naiveté that led to the current dilemma.

As a result, I must underscore that releasing or diverting some mentally ill individuals from prison is only one measure to address prison over-crowding and to reduce expenditures. All mentally ill prisoners are not suitable candidates for conditional release.[135] Not all mentally ill individuals respond to treatment; and some may pose a risk of violence that justifies their continued incarceration.[136] Releasing mentally ill prisoners who make headlines by committing violent acts will undo any reform that may be in place.[137]

Despite that, meaningful, if incremental, reform is possible. It requires careful risk assessment of whether a prisoner can be successfully integrated into the community,[138] and devotion of resources for follow-up care, including finding or creating housing, and for assuring that they comply with a regimen of treatment.[139] Critics of compelled treatment should recognize that the alternative currently is incarceration, a cruel option for a person who may have difficulty making an informed choice for herself. Critics of prison reform must recognize that years of get-tough-on-crime has bloated our prisons beyond our ability to afford them and that when applied to the mentally ill, those sentences are particularly cruel and often unnecessary.

---

133. *See supra* Part II.
134. *See* Coleman v. Schwarzenegger, No. CIV S-90-0520, 2009 WL 2430820, at *12 (E.D. Cal. Aug. 4, 2009).
135. Kobley, *supra* note 117.
136. *See supra* notes 85–86.
137. OUT OF THE SHADOWS, *supra* note 25, at 54–56.
138. *See supra* notes 120–23.
139. SLATE & JOHNSON, *supra* note 79, at 183–197.

THE OVERLOOKED UTILITY OF THE DEFENDANT CLASS ACTION

FRANCIS X. SHEN†

When and how can defendant class actions serve the goal of increasing social welfare? Existing literature on class actions has overlooked the utility of defendant class actions, and thus, has failed to answer this question. This Article presents a general theory of defendant class actions, and argues that three interrelated principles should guide the use and evaluation of defendant class actions. (1) Forward looking deterrence principle. The forward looking deterrence principle holds that the utility of defendant class actions should be measured by its contribution to future deterrence of harms by the proposed defendant class. (2) Dynamic effects principle. The dynamic effects principle holds that evaluation of a defendant class action should include all secondary effects such as feedbacks, price adjustments, new incentive structures, and changing group dynamics. (3) Aggregate analysis principle. Taking the dynamic effects principle one step further, the aggregate analysis principle holds that the evaluation of defendant class actions should ultimately rest on an aggregate, society-wide cost-benefit analysis. In developing its general theory, and synthesizing these three principles, this Article utilizes a newly constructed database of 177 cases considering defendant class action certification. This Article also spends significant time analyzing deficiencies in Hamdani and Klement's 2005 proposal for "the class defense." Three potential applications for defendant classes are considered at various points in the paper: (1) illegal file sharing on the Internet, (2) corporate fraud and illegal dealing, and (3) copyright infringement. In each context, this Article argues that existing literature and jurisprudence generally take a backwards looking approach, do not properly account for dynamic effects, and too often ignore aggregate analyses.

INTRODUCTION

You know what a class action lawsuit is. But what do you remember about *defendant* class actions from your civil procedure or torts class? The answer, most likely, is nothing. That is because defendant

† I owe a special debt of gratitude to David Rosenberg, who provided advice, critique, and feedback on this Article. Jack Goldsmith, Jed Sugarman, and fellow participants in the Harvard Law School Summer Academic Fellows program also provided very useful feedback. Sophia Beal provided her constant support and excellent editing skills. I note that my work is done A.M.D.G.

class actions are typically overlooked in both law school classes and legal scholarship.[1] This Article argues that upon closer examination the defendant class action can—in certain situations that may present themselves more frequently in coming years—serve the goal of maximizing social welfare.

To date, academic analysis of class action litigation has focused almost exclusively on plaintiff class actions.[2] Although there have been a handful of articles and notes concerned with the defendant class, they do not provide us with a comprehensive theory with which to understand and evaluate defendant class actions.[3] Recent proposals for expanding

1. *See infra* Figure 1.

2. *See* David L. Shapiro, *Class Actions: The Class as Party and Client*, 73 NOTRE DAME L. REV. 913, 914 n.2 (1998) ("A full bibliography of those publications devoted in whole or substantial part to the use of class actions in litigation would warrant a sizable appendix. But a listing of books and articles I have found helpful—some of which are long and detailed, while others, though short, are incisive and provocative—may serve a dual purpose: to provide a brief, accessible bibliography for those interested in further research and to furnish a single, easily consulted source of cross-reference for later citations in this essay.").

3. *See generally* Theodore W. Anderson & Harry J. Roper, *Limiting Relitigation by Defendant Class Actions from Defendant's Viewpoint*, 4 J. MARSHALL J. PRAC. & PROC. 200 (1971) (discussing the differences between plaintiff and defendant class actions under the then recently adopted Rule 23); Debra Lyn Bassett, *U.S. Class Actions Go Global: Transnational Class Actions and Personal Jurisdiction*, 72 FORDHAM L. REV. 41, 42–44 (2003) (discussing "the impact of the participation of other countries' citizens in U.S.-based class action litigation," specifically in regard to the issue of personal jurisdiction); Elizabeth Barker Brandt, *Fairness to the Absent Members of a Defendant Class: A Proposed Revision of Rule 23*, 1990 BYU L. REV. 909, 909–12 (1990) (discussing due process and general fairness concerns when Rule 32 is applied to defendant class actions); Vince Morabito, *Defendant Class Actions and the Right to Opt Out: Lessons for Canada from the United States*, 14 DUKE J. COMP. & INT'L L. 197, 197–202 (2004) ("[O]pt out regimes should not be employed in defendant class proceedings as they create serious obstacles to the fulfillment of the policy objectives of the class action device . . . and are not necessary to ensure that members of defendant classes are treated fairly."); A. Peter Parsons & Kenneth W. Starr, *Environmental Litigation and Defendant Class Actions: The Unrealized Viability of Rule 23*, 4 ECOLOGY L.Q. 881, 908–14 (1975) (arguing that the defendant class action can be used as "a constitutionally sound and highly practical vehicle for environmental litigation" in certain situations); Samuel M. Shafner, *The Juridical Links Exception to the Typicality Requirement in Multiple Defendant Class Actions: The Relationship Between Standing and Typicality*, 58 B.U. L. REV. 492, 492–93 (1978) (discussing how the standing doctrine and typicality requirement apply to class actions involving multiple defendants, specifically regarding the potential problems arising from the juridical links exception to typicality); Robert R. Simpson & Craig Lyle Perra, *Defendant Class Actions*, 32 CONN. L. REV. 1319, 1319 (2000) (exploring "the sparse law governing defendant class action lawsuits and its potential applicability to the recent wave of litigation against the firearms industry"); Barry M. Wolfson, *Defendant Class Actions*, 38 OHIO ST. L.J. 459, 459–61 (1977) (presenting defendant class action as a legitimate, useful and under-utilized, tool in litigating certain issues); Angelo N. Ancheta, Comment, *Defendant Class Actions and Federal Civil Rights Litigation*, 33 UCLA L. REV. 283, 283–89 (1985) (arguing that "the defendant class action is a powerful, albeit uncommon, procedure for vindicating constitutional and statutory civil rights"); Note, *Defendant Class Actions*, 91 HARV. L. REV. 630, 632–33 (1978) [hereinafter *The Harvard Note*] (presenting an overview of defendant class actions and discussing the potential due process and general fairness issues presented by this litigation device); Irving A. Gordon, Comment, *The Common Question Class Suit Under the Federal Rules and in Illinois*, 42 ILL. L. REV. 518, 528 (1948) ("[T]he defendant class suit presents both motive and opportunity for improper practice."); Debra J. Gross, Comment, *Mandatory Notice and Defendant Class Actions: Resolving the Paradox of Identity Between Plaintiffs and Defendants*, 40 EMORY L.J. 611, 611–13 (1991) (proposing a revision to Rule 23 which would mandate notice to all defendants in defendant class action suits); Robert E. Holo, Comment, *Defendant Class Actions: The Failure of Rule 23 and a Proposed Solution*, 38 UCLA L. REV. 223, 266–68 (1990) (arguing that Rule 23 ought not to govern defendant class actions and proposing the adoption of a new rule spe-

the use of defendant class action devices have focused primarily on issues arising out of internet and mass communication markets, without considering a more general application.[4] For example, these recent proposals have almost entirely missed the possibility of defendant class actions as a tool for improving responsible corporate decision-making.

In the standard treatment of class actions, commentators typically set aside analysis of defendant class actions altogether with an explanation such as, "today defendant class actions are rare and pose special problems of representation and due process that are beyond the scope of this paper."[5] The standard approach is correct in observing that defendant class actions are certainly more rare and, at present, more legally suspect in the eyes of courts than plaintiff class actions. But by stopping there,

---

cifically designed for this type of suit); Leighton Lee III, Comment, *Federal Rule of Civil Procedure 23: Class Actions in Patent Infringement Litigation*, 7 CREIGHTON L. REV. 50, 59–60 (1973) (noting the problem of adequate representation of defendants in patent infringement defendant class actions); Scott Douglas Miller, Note, *Certification of Defendant Classes Under Rule 23(b)(2)*, 84 COLUM. L. REV. 1371, 1371 (1984) (discussing how to determine when certification of a defendant class is appropriate and proposing a "test which minimizes heterogeneity by certifying only those classes whose members share a legal relationship that predates the litigation—a juridical link"); Note, *Statutes of Limitations and Defendant Class Actions*, 82 MICH. L. REV. 347, 347–50 (1983) ("[I]n defendant class actions the statute of limitations should be tolled as to all named and absent class members upon informal notice given by the plaintiff at the beginning of the suit."); Randy Clarke, A Defendant Class Action Lawsuit: One Option for the Recording Industry in the Face of Threats to Copyrights Posed by Internet Based File-sharing Systems (Spring 2001) (unpublished Honors Scholar Seminar Paper, Chicago Kent College of Law), http://www.kentlaw.edu/honorsscholars/2001students/writings/clarke.html (exploring the use of defendant class actions to litigate cases of peer-to-peer file-sharing copyright infringement).

4. *See generally* Nelson Rodrigues Netto, *The Optimal Law Enforcement with Mandatory Defendant Class Action*, 33 U. DAYTON L. REV. 59, 59–60 (2007) ("The objective of this article is to suggest an enhancement of law enforcement through mandatory aggregation of defendants and improvement of the defendant class action to incentivize the class lawyer."); Nicole L. Johnson, Comment, *BlackBerry Users Unite! Expanding the Consumer Class Action to Include a Class Defense*, 116 YALE L.J. 217, 217–18 (2006) ("This Comment takes the Hamdani and Klement proposal ["to allow certification of defense classes at the instigation of defendants"] a step further and suggests that the class defense has a more expansive applicability, not only for achieving economies of scale and overcoming collective action problems in litigation, but perhaps more importantly in obtaining settlements.").

5. Shapiro, *supra* note 2, at 919. Shapiro also notes that, "As Stephen Yeazell has shown in his informative history of the class action, defendant classes with a pre-existing coherence were often litigants in the early stages of class action development . . . ." *Id.* Nagareda, too, tables the question for another day:

> Though the Supreme Court has yet to speak definitively to the matter, federal appellate courts have proven relatively unreceptive to defendant classes under Rule 23(b)(2). Whether that chilly reception stands as either a proper reading of Rule 23 or otherwise a sensible conception of the class action is a question that I leave for another day.

Richard A. Nagareda, *The Preexistence Principle and the Structure of the Class Action*, 103 COLUM. L. REV. 149, 181 n.131 (2003) (citation omitted). Erichson makes the same move when he writes:

> Defendant class actions are permitted by Rule 23(a), and are certified on rare occasion. This paper, however, considers only plaintiff class actions, which are far more common and offer a better foil for understanding mass non-class litigation. Although mass litigation sometimes involves hundreds of defendants, and defense lawyers often coordinate their efforts through joint defense agreements, the mass collective representations that resemble class actions occur almost exclusively on the plaintiff side.

Howard M. Erichson, *Beyond the Class Action: Lawyer Loyalty and Client Autonomy in Non-Class Collective Representation*, 2003 U. CHI. LEGAL F. 519, 531 n.37 (2003) (citations omitted).

the standard mode of analysis gives us little insight into how we should evaluate this present state of affairs.

Why do defendant class actions receive such little treatment? If they are seen as theoretically untenable or unfair, then the theory needs to be examined. If we ignore defendant class actions because they are fewer in number than plaintiff class actions, the question to ask is whether they should be used more often. If the argument is that they are not feasible in practice, then system design issues come to the forefront. These issues—theory, frequency, and feasibility—are related, but distinct from one another. This Article will address each of them, focusing most of its attention on the fundamental principles that should motivate courts to certify defendant classes. The goal of this Article is thus to lay out a general theory of defendant class actions.

In developing its general theory, this Article argues that courts and commentators have recognized the benefits of aggregation, but have overlooked the informational advantages of the defendant class device. Specifically, this Article argues that the class action device can serve an auction-like function of producing information about defendants' relative contributions to harm. In situations where the market is unlikely to produce such information, the value of defendant class actions is greater. This Article delineates a series of real-world situations in which these informational benefits can be gained through a defendant class action.

In developing its theory, this Article argues that three interrelated principles should guide the use and evaluation of defendant class actions:[6]

(1) Forward looking deterrence principle. The forward-looking deterrence principle holds that the utility of a defendant class action should be measured by its contribution to future deterrence of harms by the proposed defendant class.[7] This principle stands in stark contrast to an existing strand of jurisprudence that looks backwards and attempts to determine pre-existing relationships (or "juridical links") between members of the proposed defendant class.[8]

---

6. To be sure, similar principles can be, and have been, applied to traditional class actions as well. *See* CHARLES FRIED & DAVID ROSENBERG, MAKING TORT LAW: WHAT SHOULD BE DONE AND WHO SHOULD DO IT 37 (2003) (identifying optimal precautions, optimal insurance, and redistribution of wealth as primary goals of the tort system).

7. As will be discussed subsequently, the premise is that individuals in the future, whether potential defendants or potential plaintiffs, will adjust their behavior according to the court's actions. Thus, the court is not constantly shifting, but rather making a clear statement about what individuals can expect if they act in certain ways, e.g., they might expect to be included in a defendant class and stuck with joint and several liability for the harm caused by their class. Courts can still be flexible in administering the rule in different contexts as changes occur (social, technological, etc.). *See infra* Part II.A.

8. *See generally* Shafner, *supra* note 3; Miller, *supra* note 3.

(2) Dynamic effects principle. The dynamic effects principle holds that our evaluation of defendant class actions should include all secondary effects such as information generation, feedbacks, price adjustments, new incentive structures, and changing group dynamics. This includes the standard law and economics approach to examine incentive structures, but "effects" here are broadly defined to also include group dynamics related to psychological mechanisms. This principle stands in opposition to the position that the court should focus solely on the immediate effects for the named plaintiff and defendants.

(3) Aggregate analysis principle. Taking the dynamic effects principle one step further, the aggregate analysis principle holds that our evaluation of defendant class actions should ultimately rest on an aggregate, society-wide cost-benefit analysis. In situations where deterrence of harm simultaneously involves deterrence of a good, the aggregate analysis principle instructs the legal analyst to consider multiple cross-cutting effects at high levels of aggregation.

With these three background principles laying the foundation, the Article makes a series of more specific arguments. Drawing on an analysis of 177 cases where defendant class actions were contemplated, the Article argues that courts have failed to see that plaintiff and defendant class actions should not be distinguished on conceptual grounds, but rather on the different group dynamics that are likely to exist in defendant, as opposed to plaintiff, classes. Specifically, the incentives for intra-class information sharing between plaintiff and defendant class members is likely to be quite different without the class device in place.

In developing its general theory, this Article analyzes Hamdani and Klement's proposal for "the class defense," a device that would allow defendants to class themselves with others similarly situated.[9] This Article argues that although Hamdani and Klement's analysis is more thorough than previous work on defendant class actions, it still fails to go far enough toward a general theory. The paper also examines Netto's recent argument for the use of defendant class actions in the case of illegal downloading. Netto provides a defense of aggregation, but like Hamdani and Klement, fails to recognize the informational benefits likely to arise out of some even small defendant classes.

In addition to a general discussion, two potential applications for defendant classes are considered at various points in the paper: (1) deterring illegal file sharing on the Internet, and (2) deterring corporate fraud and illegal dealing. In both contexts, this Article argues that existing literature and jurisprudence generally take a backwards-looking approach, do not properly account for dynamic effects, and too often ignore aggregate analysis. This Article argues that failure to follow these principles

9. Assaf Hamdani & Alon Klement, *The Class Defense*, 93 CALIF. L. REV. 685, 687 (2005).

makes it less likely that the existing solutions will achieve optimal deterrence. This Article also considers the hard case of copyright infringement, which challenges the feasibility of defendant class actions in cases where no group of defendants is readily identifiable as the group to lead the class defense.

This Article is organized into four sections. The first section of the paper reviews existing literature on defendant class actions. The second section develops a general theory, drawing in part on the psychology literature on group decision making. The third section then presents a system design based on the general theory, focusing in particular on the application of these systems to the case of illegal dealings by corporate executives, illegal file sharing on the Internet, and copyright infringement. The fourth section concludes with thoughts for future research directions in this area.

I. EXISTING LITERATURE

The existing literature on defendant class actions is comprised of a few journal articles, several Notes, and a handful of additional publications.[10] Much of the literature on defendant class actions has considered how Federal Rules of Civil Procedure 23 ("Rule 23" or "the Rule") can be applied to defendant class actions.[11] For example, Scott Douglas Miller, author of *Certification of Defendant Classes Under Rule 23(B)(2)*, discusses the "dispute over Rule 23's terminology" and provides an analysis of the text of the Rule.[12] Likewise, Randy Clarke moves through the language of the Rule in evaluating a potential defendant class action against music downloaders.[13] In his commentary on defendant class actions, Robert Holo also proceeds with a formalist analysis, considering how the language of the Rule applies: "Despite *Doss*, it is clear that (b)(2) certification of defendant classes is always inappropriate because of the express language of the rule. Courts should not ignore the clear language of the rule in order to better serve their perceptions of justice or fairness."[14] This Article does not focus on formalist

10. *See* sources cited *supra* note 3.

11. Fed. R. Civ. P. 23. Interpretation of Rule 23 has been a challenge for courts and academics alike because it is open to varying readings. As Judge Posner noted, "The question whether there can be a defendant class in a Rule 23(b)(2) suit cannot be answered by reference to authority." Henson v. E. Lincoln Twp., 814 F.2d 410, 413 (7th Cir. 1987). Because of this potential latitude, federal appeals courts have moved to reign in the class mechanism. The Court of Appeals for the Eleventh Circuit has written that a rule to the contrary would "enable any action, with the possibility that it might be one of multiple actions, to be certified pursuant to Rule 23(b)(1)(B)." Namoff v. Merrill Lynch, 829 F.2d 1539, 1546 (11th Cir. 1987).

12. The language in 23(b)(2) is his concern: "Few actions for equitable relief are based on plaintiffs' conduct; rather, plaintiffs initiate such suits in response to defendants' conduct." Miller, *supra* note 3, at 1375. Miller's analysis of court cases proceeds to consider how they look at the language of the Rule. "Thus, all federal courts that have considered defendant class certification under Rule 23(b)(2) have done little more than superficially reviewed the rule's terms." *Id.* at 1376.

13. Clarke, *supra* note 3.

14. Holo, *supra* note 3, at 264.

concerns such as the best interpretation of the language of Rule 23. Rather, this Article adopts a functionalist framework and theorizes about when defendant class actions will best serve the goals of maximizing social welfare.

Although articles by Netto (2007), Johnson (2006), and Hamdani and Klement (2005) have begun to address more functionalist concerns in the past few years, the literature remains limited.[15] My review of the literature argues that scholars have generally concentrated too much on proceduralist concerns (i.e., scrutiny of the language of Rule 23), and have failed to provide a thorough functionalist analysis. My purpose in reviewing this literature is to identify some of the most discussed market and incentive dynamics associated with defendant class actions. Once these dynamics are recognized, Section II of the paper develops a general theory to incorporate them.

A. All Defendant Classes are Not the Same

Defendant class actions originate out of the same legal history and Federal Rules of Civil Procedure as plaintiff class actions.[16] Like plaintiff class actions, defendant class actions became more feasible after the 1966 amendments to Rule 23.[17] Although defendant class actions are less frequent than plaintiff class actions, "[t]he use of a defendant class action is not a recent development."[18]

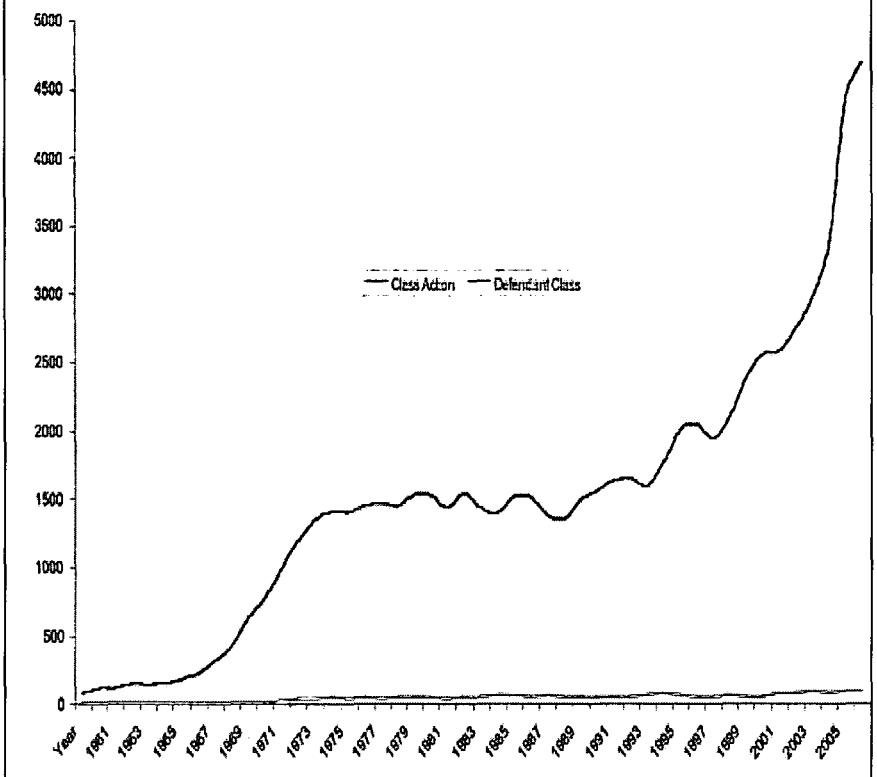
15. The literature also remains disconnected from previous studies. The literature, for instance, has yet to be synthesized in a single article. Even the more recent articles have not cited all previous works. In Hamdani & Klement's analysis of defendant classes, they fail to cite several works on defendant class actions, including a short piece from three years earlier that had considered defendant class actions in the similar context of file sharing. The uncited work was Clarke, *supra* note 3.

16. *See generally* Netto, *supra* note 4, at 76–87 (providing a history of the defendant class action, and its development in the United States).

17. Howard Downs notes that "[w]hereas original Rule 23 restricted binding class actions to cases involving 'joint or common rights' or actions affecting 'specific property,' amended Rule 23 relaxed these restrictions, which extended the social and economic uses of the class device." Howard M. Downs, *Federal Class Actions: Due Process by Adequacy of Representation (Identity of Claims) and the Impact of* General Telephone v. Falcon, 54 OHIO ST. L.J. 607, 608 (1993). "Although it appears that the modern-day class action was born probably some time during the Middle Ages, there are reports of ecclesiastical proceedings against numerous insects and animals dating as early as A.D. 824." Shaw v. Toshiba Am. Info. Sys., Inc., 91 F. Supp. 2d 942, 946–47 (E.D. Tex. 2000). "These early 'defendant class actions' date from a very early period: in A.D. 824, against moles in Aosta; in A.D. 864, bees in Worms; in A.D. 886 locusts of Romagna; and in the same century, serpents of Aux-les-Bains." *Id.* at 947.

18. Doss v. Long, 93 F.R.D. 112, 115 (N.D. Ga. 1981).

Figure 1. Number of "class action" and "defendant class" mentions in federal and state cases, 1960-2007



Notes: Graph notes the number of hits for "class action" vs. "defendant class" action cases in LexisNexis database of all state and federal cases. The number of hits does *not* represent the actual number of certified plaintiff or defendant classes.

Nonetheless, the explosion of class action litigation has been overwhelmingly on the plaintiff side. To gain some historical perspective, I conducted a LexisNexis search of all Federal and State cases from 1960-2007 using the phrase "class action" or "plaintiff class." I then ran the search again with the terms "defendant class action" or "defendant class." These searches, while not providing an accurate count of the actual number of cases contemplating class actions, nevertheless serve as a proxy for the popularity of the class device in the courts. The number of hits per year, presented graphically in Figure 1, gives us a sense of the disparity between defendant and plaintiff class actions. While discussion of class actions generally has risen steadily since 1966—growing very significantly in the last decade—contemplation of defendant class actions has remained quite low throughout the forty years. While this class action term search produced over 1,000 hits starting in the 1970s, over

2000 starting in the late 1990s, and over 4,000 in the most recent years, defendant class mentions have never risen over 100 hits. The number of plaintiff class actions clearly dwarfs the number of defendant class actions.

One straightforward reason for such little use of the defendant class device is current jurisprudence on Rule 23. Defendant class actions are governed by Rule 23, and thus, as a preliminary matter courts look for satisfaction of the four Rule 23(a) prerequisites: numerosity, commonality, typicality and adequacy of representation.[19] Analysis of these prerequisites explains much of the infrequency of defendant class actions, but tells us little about whether that infrequency is a useful (functional) outcome. I am not concerned with re-interpretation of Rule 23, but rather, I am primarily concerned with evaluating the outcome of its current interpretation, i.e., evaluating whether the defendant class action should be expanded on the grounds of improving social utility. For reasons to be discussed subsequently, I argue that in fact there should be such an expansion of defendant class action use.

Before moving to that argument, let us review the prerequisites that prevent many instances of efficient and socially desirable class certification. Courts currently do not depart radically from accepted views of Rule 23 jurisprudence. A recent 2003 decision from the District of New Jersey provides a concise summary of the state of the law:

> There is a significant split of opinion as to whether Rule 23(b)(2) ever permits injunctive relief against a defendant class. The Fourth and Seventh Circuits, together with the leading treatise on federal procedure, take the view that defendant classes are not authorized by Rule 23(b)(2). These authorities are generally of the view that the text of 23(b)(2) itself forbids defendant classes. . . .
>
> On the other hand, the Second Circuit, together with the leading class action treatise, take the view that defendant classes are permitted by Rule 23(b)(2). The Sixth Circuit appears to agree that defendant classes are permissible under Rule 23(b), but only if individual defendants are all acting to enforce a locally administered state statute or uniform administrative policy. The principal justification for permitting defendant classes under Rule 23(b)(2) seems to be that the device can be particularly useful to bind to a court decree a group of defendants who, out of recalcitrance or neglect, have refused to conform their conduct to settled substantive law or to eliminate the need for ancillary proceedings against a number of semi-autonomous defendants once the court has made a basic determination of legal issues applicable to all.[20]

19. *See* Fed. R. Civ. P. 23(a).
20. Clark v. McDonald's Corp., 213 F.R.D. 198, 217. (D.N.J. 2003) (citations omitted).

Ultimately, the court concluded that:

> A review of the foregoing district court decisions reveals that the certification of 23(b)(2) defendant classes has been implemented only tepidly in the Third Circuit, and has met success, if at all, only in cases where the individual defendants of the class are alleged to be acting in conformity with an illegal state statute, rule, or regulation.[21]

Commentary from other courts similarly note that "defendant classes seldom are certified," and if they are certified, "such certification most commonly occurs[:] (1) in patent infringement cases; (2) in suits against local public officials challenging the validity of state laws; or (3) in securities litigation."[22]

To gain a broader perspective on defendant class actions, I examined cases in which a defendant class action was contemplated.[23] Utilizing the LexisNexis database of all federal and state cases, as well as previous academic and court citations, I identified 177 cases in which a defendant class was contemplated.[24] These cases, listed in the online appendix, were coded for subject. Table 1 provides a summary of the subject matter.

The analysis of these cases is consistent with the courts' observations that defendant class actions have been used frequently for securities cases and for constitutional challenges. These two categories alone account for fifty three percent of the defendant class action cases. There are, however, more extensive uses of class actions than typically ac-

---

21. *Id.* at 220. In this particular case, plaintiffs sought to certify as a class all McDonald's under Title III of the Americans with Disabilities Act. The court did not certify a defendant class because "the individual members of the defendant class have been non-uniform in their non-compliance with such policies." *Id.* The court speculated that:

> Had plaintiffs alleged, for example, that McDonald's and its franchisees adhered to a company-wide policy of providing just one handicapped parking space in restaurant parking lots, or of installing no "grab bars" in restaurant toilet stalls, then one could imagine why injunctive relief—against the defendants as a class—might be appropriate to redress such violations.

Id. at 220–21.

22. Thillens, Inc. v. Cmty. Currency Exch. Ass'n of Ill., Inc., 97 F.R.D. 668, 674 (N.D. Ill. 1983) (citations omitted). *Thillens* went on to read:

> Several rules, useful in unilateral as well as bilateral defendant class actions, emerge from *In re Gap* and similar cases: (1) A defendant class will not be certified unless each named plaintiff has a colorable claim against each defendant class member; (2) A defendant class will not be certified under Fed. R. Civ. P. 23(b)(3) without a clear showing that common questions do *in fact* predominate over individual issues; (3) The requirement that each named plaintiff must have a claim against each defendant may be waived where the defendant members are related by a conspiracy or "juridical link."

Id. at 675–76. Netto notes that the defendant class action device "is more frequent in lawsuits involving civil rights, disputes challenging constitutionality of state and local law and practices enforced by public officials, and suits against unincorporated associations, e.g., labor unions. Defendants' classes have also been certified in other contexts, such as patent infringement, antitrust, securities, and environmental law." Netto, *supra* note 4, at 87.

23. Defendant classes sometimes emerge out of counter-claims in plaintiff class actions. I have excluded them from this analysis, as they are not the focus of the paper.

24. The search was conducted in February 2008.

knowledged. While declaratory judgments on property rights and benefits are similar to the constitutional challenge and security cases, ten percent of the cases concerned damages.

Table 1. Summary of selected cases in which defendant class action was proposed

Case Subject Matter	Number	Pct.
Constitutional Challenge	63	35.6%
Securities	31	17.5%
Damages	18	10.2%
Property Rights	17	9.8%
Benefits - Insurance or Retirement	14	7.9%
Monopoly / Anti-Trust	7	4.0%
Taxes / Fees	6	3.4%
Patent	7	4.0%
Contractual	4	2.3%
Bankruptcy	4	2.3%
Other	4	2.3%
Copyright	2	1.1%

NOTES: Defendant classes were not certified in all cases. The 177 cases coded here were identified through searches in the LexisNexis database of All Federal and State cases. See text for details of search procedures.

Both academics and judges have paid close attention to the nature of the potential defendant class. Over twenty-five years ago, *The Harvard Note* recognized the functional nature of many defendant class actions: "The structure of certain types of defendant class actions virtually guarantees adequate representation. Suits against the members of a labor union or other unincorporated association, naming the officers as representative of the class, provide one example."[25]

When the relationship between defendants is clearly demarcated, the courts see fewer barriers to certifying defendant classes. Analyzing when courts are likely to certify defendant classes, Miller finds that "[c]orrectional institutions, county magistrates, county sheriffs, local prosecutors, and voting officials have all been certified and bound as defendant classes."[26] Courts have developed the juridical links exception

25. *The Harvard Note*, *supra* note 3, at 642.
26. Miller, *supra* note 3, at 1379.

to understand connections between defendant members of the class. Courts have defined a juridical link "as 'some [independent] legal relationship which relates all defendants in a way such that single resolution of the dispute is preferred to a multiplicity of similar actions.'"[27] Examples of such links include partnerships or joint enterprises, conspiracy, and aiding and abetting. These terms denote some form of relationship or activity on the part of the members of the proposed defendant class "that warrants imposition of joint liability against the group even though the plaintiff may have dealt primarily with a single member."[28]

In a similar vein, Holo sees defendant class actions as more likely when the defendants are connected through some superior authority.[29] Holo provides examples of courts certifying defendant classes in securities fraud cases, and suits against groups of state/local officials.[30] It is important to note here that in these cases, the courts look backwards to see if a relationship existed between the potential defendant class members prior to the allegations. When a juridical link already exists, courts are willing to see the group dynamics. But they do not see how they could actually *create* such links in the future via their judgments in the present case; there is no forward looking jurisprudence.

The courts' analyses in these cases bear some resemblance to the search for a conspiracy or coordinated action. In a 1990 opinion, Federal District Judge D. Brock Hornby recognized this connection in a footnote, in which he quotes Holo and states that:

---

27. Follete v. Vitanza, 658 F. Supp. 492, 507 (N.D.N.Y. 1987) (alteration in original) (quoting *Thillens*, 97 F.R.D. at 676).

28. *Id.* at 508 (quoting Akerman v. Oryx Commc'ns, Inc., 609 F. Supp. 363, 375 (S.D.N.Y. 1984)). In his Note, Miller stated: "The test suggested by this Note minimizes the dangers inherent in class heterogeneity by certifying only those defendant classes whose members share a relationship predating the litigation, and whose role in the litigation derives from their membership in the preexisting group. Courts have characterized such classes as 'juridically linked.'" Miller, *supra* note 3, at 1394–95. "When the defendant class is juridically linked these courts miss the mark. In such cases individual relief is subordinate to class relief. Traditional party relationships should be far less significant than the general nature of the interclass dispute." *Id.* at 1400–01. Courts do not always agree on whether sufficient juridical links exist. In *Funliner of Alabama, L.L.C. v. Pickard*, the Alabama Supreme Court focused on a lack of written agreement as determinative:

> In *In Re Activision Securities* . . . the Court found that the defendants, who were all underwriters and members in a securities syndicate, had entered into a written agreement . . . . We do not find the facts of *Activision* analogous to those of the instant case. There has been no finding that the defendants in this case entered into a written agreement or that they agreed to be bound to a common course of conduct; the trial court did not even note that the plaintiffs alleged a conspiracy among the defendants. Thus, the juridical-link exception found in *Activision* is missing here.

873 So. 2d 198, 215–16 (Ala. 2003).

29. Holo, *supra* note 3, at 239 ("All the defendants are bound together because of their common obligation to adhere to a particular state law or policy.").

30. "For example, modern securities fraud litigation often involves a plaintiff class of investors suing a defendant class of securities underwriters." *Id.* at 227. Holo also notes the usefulness of defendant class action in the context of state/local officials who are illegally discriminating. "By binding all members of a defendant class to a single judgment, widespread discriminatory practices can be brought to a halt more quickly and efficiently." *Id.* at 228.

> I leave to the plaintiffs determination of how properly to join the dealers as named defendants. I recognize the complexities in joining a large number of defendants or, as suggested at oral argument, creating a defendant class. Commentators have wondered: "Can the existence of a conspiracy be proven in a single proceeding representing the entire defendant class, or does proof of a conspiracy depend upon proving each defendant's participation in the alleged conspiracy, an inherently individual question that must be answered separately for each defendant?"[31]

One additional rule courts have introduced in analyzing the relationship of potential members of a defendant class is a membership ratification theory.[32] "Under [the membership ratification] theory dealing with individual proof of illegal conduct becomes unnecessary. Rather, a presumption arises that all members of the association joined in the alleged conspiracy."[33] It is essentially a 'guilty by association with the Association' rule. Functionally, this is telling individual defendants that they should have asked questions up front and should have monitored their association, or otherwise contracted *ex ante* to avoid this liability.[34] Like the juridical links rule, however, the membership ratification rule looks back to earlier relationships between potential class members. But even though the courts are backwards-looking here, we can see in their jurisprudence the roots for more functionally effective legal rules. For instance, laying down a 'guilt by association with the Association' rule would likely have a strong future deterrent effect on the behavior of individuals in that Association.

B. Financial Incentives & Free Riding

The most frequently noted motivational problem with defendant class actions is the lack of adequate incentive for defendant class representatives to fully litigate. This basic insight was offered over two decades ago:

> Defendants generally oppose motions to certify them as class representatives. The major reason for their opposition presumably is a desire to avoid a possible increase in litigation expenses if they represent a class, in light of the fact that no source of funds is available to pay for any additional costs.[35]

---

31. *In re* New Motor Vehicles Can. Exp., 307 F. Supp. 2d 136, 141 n.7 (D. Me. 2004) (quoting Holo, *supra* note 3, at 258).

32. Holo, *supra* note 3, at 259.

33. *Id.*

34. *See* Phelps Dodge Ref. Corp. v. Fed. Trade Comm'n, 139 F.2d 393, 396–97 (2d Cir. 1943).

35. *The Harvard Note*, *supra* note 3, at 648.

This point has been reiterated since then in most discussions of defendant class actions.[36] As discussed by Hamdani and Klement, when the defendant class wins, "the defendants owe nothing to the plaintiff—no money changes hands."[37] Thus, there is no money to pay counsel for the class representative because no single member of the defendant class has the proper financial incentives to litigate the defense fully.[38]

The incentive problem is connected to a free-rider problem: defendant class members who are not litigating stand to benefit without cost from a successful class defense.[39] Unlike plaintiff classes, where litigation costs can be subtracted out of a settlement, it is more difficult to extract money from passive defendants in the class.[40] Analysts have been grappling with this problem—and how to correct it—for many years.[41] Dwelling on the comment that it would be difficult, if not impossible, to "devise a method to tax such 'free riders,'" the 1978 *The Harvard Note* observes that:

> [A]ssuming that all the class members or collateral estoppel beneficiaries could be identified, there would still be problems of determining how much to charge each individual. Only the common issues will have been litigated if the defendant class prevails, and the court will therefore have no knowledge of the magnitude of total liability avoided or of the proportion attributable to each class member. . . .

---

36. *See, e.g.*, Netto, *supra* note 4, at 92 ("There are three foremost concerns related to the choice of adequate representation in defendant class actions: (i) the choice of the representative is made by the plaintiff; (ii) the absence of incentive for any defendant to bear the expenses of defending a lawsuit on behalf of the entire class when the costs of litigation are disproportionate to the representative party's stake; and (iii) the difficulty of compensating class counsel for the benefits conferred upon the class."); *see also* Brandt, *supra* note 3, at 919–20 ("In comparison, a defendant class representative will seldom be able to take advantage of the same fee incentives as a plaintiff representative. . . . Consequently, the defendant representative must be prepared to assume some, if not all, of the economic burden of the litigation.").

37. Hamdani & Klement, *supra* note 9, at 691.

38. An important exception, discussed *infra* Part II.B.1, is when there are "dominant players" in the class.

39. *See The Harvard Note*, *supra* note 3, at 648 ("[T]here might seem to be a certain unfairness to the defendant class representative even if his defense of the class entails no extra costs; if the common question is resolved in favor of the defendant class, absentee members will have received a benefit at the representative's expense without having to compensate him for it."); *see also* Miller, *supra* note 3, at 1385 ("Further, party heterogeneity increases the legal fees and administrative costs associated with coordinating a defense. The defendant class representative cannot expect to recoup these additional costs . . . .").

40. The ability to correct for the free-rider problem, as discussed in Part III.D, depends heavily on the nature of the group dynamics within the defendant class.

41. *See The Harvard Note*, *supra* note 3, at 652–53 (The Note argues for expanded use of what it terms "expanded common question defendant class action." They suggest that courts frame the question "not in terms of what each individual class member owes but rather in terms of what formula should be used to allocate the total liability." Unfortunately, after this interesting discussion, the Note suggests that, "[o]f course, in any of these 'fully litigated' defendant class actions a final stage of individualized hearings is needed—whether conducted along with the class suit or entirely separately from it.").

> [T]he class members would need to be taxed according to their potential liability, a figure difficult if not impossible to determine.[42]

Aware of the incentive problems with defendant class actions, some courts have refused to certify defendant classes on the grounds that the parties representing the class do not have the proper incentives to litigate fully.[43] This issue of free riding and funding optimal defendant class representation is a topic I take up at length in the system design section of this Article.

C. Funding Defendant Class Actions

Recognizing the free-rider problem, several funding schemes have been proposed. Some of these proposals involve a tax-like levy on defendant class members. To fund the defendant class action, the court could choose "to tax the expenses attributable to the class action to the plaintiff, to tax them to the absentee defendants, or to refuse to certify the class on any questions not perfectly common to the class members."[44] This proposed solution is to tax the absentees "with a proportionate share of at least the class-action-related expenses of the named defendant."[45]

A common alternative is to find some organization with deep pockets and make them a party as well.[46] Plaintiffs bringing the suit are typically in a position to identify the deep pocket class members on the other side. In the securities case *Northwestern National Bank of Minneapolis v. Fox & Co.*,[47] the plaintiffs sued the class of Fox partners in addition to Fox itself "in order to assure recovery of the substantial judgment likely to issue if plaintiffs succeed in proving their claims."[48] Courts have recognized that financial stakes will motivate defendants to mount adequate defenses. In *Consolidated Rail Corp. v. Town of Hyde Park*,[49] the court held that "[b]y including as [defendant] class representatives the 10 highest tax collectors from Conrail . . . the district judge created a fair group

---

42. *Id.* at 648–49 & n.96.

43. *See, e.g.*, Nat'l Ass'n for Mental Health, Inc. v. Califano, 717 F.2d 1451, 1458 (D.C. Cir. 1983) (The defendant university, U.S.C., said explicitly in testimony that "it was 'unwilling to expend the effort and funds necessary to defend itself in this action, let alone represent the interests of a large group.' . . . The school's position was supported by the affidavit of one of its administrators, who stated: 'Due to the minimal amount of its alleged liability in this action, the University of Southern California does not intend to defend this action on behalf of itself or any others.'").

44. *The Harvard Note*, *supra* note 3, at 656.

45. *Id.* at 657.

46. *See, e.g.*, Holo, *supra* note 3, at 271 (Holo's solution is for the judge to bring in some defendant with the money: "Nevertheless, the judge may, in her discretion, assign additional defendants to act as corepresentatives, thus lessening the financial burden on any one defendant and at the same time preventing any defendant from shirking his duties."); *see also The Harvard Note*, *supra* note 3, at 656 (arguing that adequate representation (aligning incentives) might be accomplished in some instances by requiring "the plaintiff to name as an additional defendant a trade organization whose membership coincided with that of the class").

47. 102 F.R.D. 507 (S.D.N.Y. 1984).

48. *Id.* at 510.

49. 47 F.3d 473 (2d Cir. 1995).

of representative parties who presumably have the greatest financial motivation to defeat Conrail's case."[50] Where plaintiffs do not already include deep pocket defendants, the court can also find the necessary parties. Holo suggests that, "a court can require a plaintiff to join additional defendants as class representatives and can also permit associations or other institutional representatives to join as representative defendants."[51] In other words, the court can look to kick a private market into motion to fund the class defense. In *In re Integra Realty Resources, Inc.*,[52] the Tenth Circuit did just this, naming Fidelity as the defendant class representative against Integra because the Fidelity Capital Appreciation Fund was the largest Integra shareholder when Integra spun off (thus bringing on the litigation).[53]

D. Aggregation, Opt-Out, and Deterrence

The benefits of aggregating claims in order to enjoy economies of scale is discussed at several points in previously published literature. Netto writes that amongst courts and academics today, "[i]t is a general consensus that the primary advantage of class actions is to override the transactional cost of low stake claims, which would not be individually prosecuted because the costs of litigation . . . supersede the expected utility from the adjudication."[54] The primary point, as noted by Holo, is that aggregation "allows the defendants to pool their resources, decide who among them would be the most fit representative, and present a strong, united front against their opponents."[55] The spirit of these comments, in favor of aggregation, is the same spirit animating David Rosenberg's arguments in *Mass Tort Class Actions: What Defendants Have and Plaintiffs Don't*.[56] In Rosenberg's analysis of plaintiff class actions he recognizes that defendants are able to enjoy the benefits of scale in defending themselves, while plaintiffs—unless they have a class device—cannot.[57] Here, in the case of defendant class actions, plaintiffs start with pooled resources that defendants do not have. The defendant class action serves as a tool to address this imbalance.

The majority of analyses on defendant class actions have argued for an opt-out option based on fairness and due process concerns.[58] But the

50. *Id.* at 484.
51. Holo, *supra* note 3, at 234.
52. 262 F.3d 1089 (10th Cir. 2001), *aff'd*, 354 F.3d 1246 (10th Cir. 2004).
53. *See id.* at 1096.
54. Netto, *supra* note 4, at 98; *see also* Hamdani & Klement, *supra* note 9, at 711–13.
55. Holo, *supra* note 3, at 268.
56. David Rosenberg, *Mass Tort Class Actions: What Defendants Have and Plaintiffs Don't*, 37 HARV. J. ON LEGIS. 393, 394–97, 399–402 (2000).
57. *See id.* at 393–94, 400–02, 407–08, 412.
58. *See, e.g.*, Brandt, *supra* note 3, at 911–13; *see also* Netto, *supra* note 4, at 98 ("In fact, some circumstances will actually create incentives not to opt out of a defendant class. For example, a plaintiff who commences a defendant class against a group of underwriters of a new stock offering may also threaten and be able to commence litigation against each of the underwriters individually. Given the certainty of having to make a choice between remaining in a defendant class or defending

cost of opt-out (and unraveling the defendant class) is significant. Simpson and Perra explain the rationale for not allowing opt out:

> [O]rdinarily no one wants to be a defendant, so that defendant class members who have an opportunity to opt out can be expected to do so . . . . Massive opt-out undermines the breadth and finality of judgments, increases the possibility of duplicative litigation, and lessens the probability of giving plaintiffs full relief.[59]

The empirical data on how the opt-out option is used in practice is lacking. As observed by Morabito, "there is little information available concerning the percentage of class members who have opted out of defendant class proceedings, after being offered the opportunity to do so following the certification of defendant classes."[60] Morabito found only three U.S. cases in which opt-out rates were discernible: "3 defendants opted out of a class of 91; no one exited another defendant class; and in a third proceeding, 115 defendants opted out."[61]

The deterrence objectives of class action litigation have also been raised several times in the existing literature.[62] In the deterrence view, "the primary purpose of class litigation is not so much to redress injured plaintiffs as to deter wrongful conduct on the defendant's part by forcing him to disgorge his unlawful gains or by restructuring his behavior through the use of injunctions."[63] Hamdani and Klement have focused extensively on deterrence, and their proposal will be discussed more in depth in the next few sections.

---

individual litigation, the economics of a joint defense considerably outweighs those of defending an individual action, and defendant class members would have an incentive to remain in the class.").

59. Simpson & Perra, *supra* note 3, at 1334 (alteration in original); *see also* Holo, *supra* note 3, at 266 (considering proposal of a no-opt-out rule before moving away from the suggestion). Although Holo doesn't stick with it, he actually considers proposing a no-opt-out rule as well:

> One more possible modification would be to eliminate the 23(c)(2) opt-out provision for proposed members of a defendant class. Some courts have worried that any defendant named in a 23(b)(3) defendant class action would promptly opt out, thus rendering the class action device useless, but this modification would successfully resolve that problem.

Id. In the next line, he moves away from this suggestion, but the logic was there. *Id.* ("In the final analysis, however, these measures also would be inadequate.").

60. Morabito, *supra* note 3, at 226 (footnotes omitted).

61. *Id.*

62. *See* Simpson & Perra, *supra* note 3, at 1319 (suggesting the possibility of using defendant class actions to solve the problem of holding the firearms market liable). As they ask at the outset:

> [H]ow can municipalities and other "representative organizations" summon each allegedly culpable firearms industry player to the table? How can these suits be structured to ensure that each participant in the manufacturing, advertising and distribution channels is held accountable for its tortious behavior? How can a plaintiff, who has suffered damages potentially caused by 191 different firearms manufacturers, hundreds of wholesalers and over 80,000 retailers nationwide, join these potential defendants in a manner that ensures that each suffers its proportional share of damages caused?

Id. Simpson & Perra structure their article, however, around the language of Rule 23, demonstrating how the four requirements can be met—not discussing why it would be a good thing to have defendant class actions. *Id.* at 1324–29.

63. *The Harvard Note*, *supra* note 3, at 654.

E. Formalist and Proceduralist Concerns

While there are strains of functionalist thinking throughout the literature on defendant class actions, the bulk of the literature still grounds itself in proceduralist concerns. Although some of these authors acknowledge deterrence objectives, they fall back on a position articulated by Miller, that the "usual incentive for defendant class certification rather is not economic utility but social justice."[64] This focus on social justice is accompanied by a focus on due process and fairness.

In the context of defendant class actions, the concerns of commentators and courts are often those of due process.[65] The court in *Thillens, Inc. v. Community Currency Exchange Association of Illinois, Inc.*[66] noted, "Fundamental fairness to absentee members must be balanced against judicial savings. Where representative adjudication occurs pursuant to a defendant class, due process concerns not inherent in plaintiff class actions arise. The crux of the distinction is: the unnamed plaintiff stands to gain while the unnamed defendant stands to lose."[67] The court in *Gaffney v. Shell Oil Co.*[68] arrived at the same point, arguing that "[i]n the final analysis, the propriety of a class action—plaintiff, defendant or both—depends upon a finding that due process will be accorded the members of the class who are not before the court."[69] This Article argues in the next Part that courts' analysis of gain and loss should include not only unnamed parties, but also future potential parties with deterrence in mind.

Exceptions, such as the juridical links exception just discussed, are used by courts to address due process concerns.[70] Fairness, usually to

64. Miller, *supra* note 3, at 1387 (footnote omitted).

65. *See In re* the Gap Stores Sec. Litig., 79 F.R.D. 283, 291 (N.D. Cal. 1978) ("[Parsons and Starr have] reviewed the use of defendant class actions in environmental litigation and . . . carefully explored the due process problems posed by defendant class adjudications" and have observed, "The basic constitutional dilemma of defendant class actions arises out of the due process rights of absent members of the defendant class. Fundamental to due process is the notion that the authoritative determination of a personal liability, obligation or right of a defendant requires the court's in personam jurisdiction over that party.") (quoting Parsons & Starr, *supra* note 3, at 888); *see also* Netto, *supra* note 4, at 105–06 ("[M]andatory class actions aggregating damages claims implicate the due process principle . . . [and] deep-rooted historic tradition that everyone should have his own day in court.") (quoting Ortiz v. Fibreboard Corp., 527 U.S. 815, 846 (1999) (internal quotation marks omitted)). *See generally* Downs, *supra* note 17, at 627–30 (discussing due process in class actions).

66. 97 F.R.D. 668 (N.D. Ill. 1983).

67. *Id.* at 674 (citation omitted).

68. 312 N.E.2d 753 (Ill. App. Ct. 1974).

69. *Id.* at 991.

70. Scott Douglas Miller states:

> A defendant class member would consent to representative adjudication only if he perceived, or might reasonably be expected to perceive, that the savings resulting from another party's representation would exceed any liabilities—monetary or otherwise—resulting from the representation. An absent defendant would only prefer representative action where he perceived himself as adequately represented. The perceived probability of loss would then be no greater in representative than in individual adjudication, but there would be a net savings of litigation costs. Only if the defendant class is juridically linked would absent members be so confident.

absentee defendants, is another way of discussing due process.[71] These fairness concerns are rooted in the belief that we should treat class actions in the same way we would treat one-to-one litigation. Bassett argues that:

> [T]here is no reason to believe that a court has the power to issue a binding judgment upon a defendant—even if that defendant is part of a defendant class—where that defendant has no nexus to the forum and her purported consent to suit is based on her failure to opt out of the class. Accordingly, there is no reason to treat members of a defendant class any differently than a defendant in a non-class lawsuit . . . .[72]

Due process issues can be summarized in what one court has labeled the 'Rubik Cube' puzzle: "[E]ach plaintiff does not have a cause of action against each defendant."[73] When faced with this situation, courts may be hesitant to certify the defendant class because they look backwards for pre-existing connections.[74] I argue that this view, articulated in different forms by most courts, fails to recognize the intra-group dynamics that a class device introduces. I therefore argue that in cases where there are sizeable enough informational and incentive benefits to be gained from classing a group of defendants, there is *every* reason to treat members of the class differently—although collectively the same—than we would treat them if they were a stand-alone defendant.

Miller, *supra* note 3, at 1399. Brandt, trying to reconcile defendant class actions with due process concerns, proposes a complicated measure:

> In order to protect the due process rights of absent defendant class members, Rule 23 should be revised in two respects. First Rule 23 should ensure that absent defendants will not be bound by a class judgment unless they receive actual notice of the pendency of the action. This protection should be extended so that it applies not only to actions under 23(b)(3) but also to defendant class actions maintained under 23(b)(1) and 23(b)(2).

Brandt, *supra* note 3, at 944–45.

71. The Harvard Note, for instance, introduces the subject by arguing that defendant class actions should not be "purchased at the expense of fundamental unfairness to persons who are not before the court that binds them." *The Harvard Note*, *supra* note 3, at 632.

72. Bassett, *supra* note 3, at 76. Bassett continues that this means "that minimum contacts with the forum state would be necessary in order to bind the defendant class member to the judgment, and if minimum contacts were not established, the class judgment would be unenforceable with respect to that defendant." *Id.*

73. Doss v. Long, 93 F.R.D. 112, 119 (N.D. Ga. 1981). The Rubik Cube problem can be considered "in terms of standing, typicality, or commonality," but underlying it is concern with due process. *Id.* at 120. The *Thillens* court noted the same thing: "There is great judicial reluctance to certify a defendant class *when the action is brought by a plaintiff class*. The primary concern with bilateral actions, antitrust or other types, is a fear that each plaintiff member has not been injured by each defendant member." Thillens, Inc. v. Cmty. Currency Exch. Ass'n of Ill., Inc., 97 F.R.D. 668, 675 (N.D. Ill. 1983).

74. In *LaMar v. H.B. Novelty & Loan Co.* in 1973, the Ninth Circuit considered the issue:

> [W]hether a plaintiff having a cause of action against a single defendant can institute a class action against the single defendant and an unrelated group of defendants who have engaged in conduct closely similar to that of the single defendant on behalf of all those injured by all the defendants sought to be included in the defendant class.

489 F.2d 461, 462 (9th Cir. 1973).

II. DEVELOPING A GENERAL THEORY OF DEFENDANT CLASS ACTIONS

With the groundwork now laid, the Article picks up on its three central principles, and develops a general theory of defendant class actions.

A. Forward Looking Deterrence

To build a general theory of defendant class actions, a preliminary question about the purpose of tort law must be addressed. This Article, like Netto (2007), adopts the initial position taken by Fried and Rosenberg, that "tort liability should be seen as part of the imperfect and partial system serving the goals of compensation and deterrence."[75] This approach follows a line of scholarship that focuses on maximization of social welfare as the goal of law generally, and of tort law specifically.[76] In the context of mass torts and collectivized adjudication, the Article follows Rosenberg's (2002) premise that when government and first-party insurance are not adequate, a "need exists for 'optimal tort deterrence' to prevent unreasonable risk of accident and for 'optimal tort insurance' to cover residual reasonable risk."[77] This position has not gone uncontested; scholars such as Richard Epstein and Richard Nagareda have criticized this approach in exchanges with Rosenberg and others.[78]

The Fried and Rosenberg approach rests on an appreciation of the *ex ante* perspective.[79] The *ex ante* perspective is one which seeks to un-

---

75. FRIED & ROSENBERG, *supra* note 6, at 2. The authors discuss these three functions at length in Chapter 3, and justify them in Chapter 2. In addition to deterrence, Fried and Rosenberg identify "optimal insurance, and related appropriate redistribution of wealth" as goals of the tort system. *Id.* at 37. I consider redistribution and insurance issues in Part III.A.1.

76. *See generally* David Rosenberg, *Mandatory-Litigation Class Action: The Only Option For Mass Tort Cases*, 115 HARV. L. REV. 831 (2002) [hereinafter Rosenberg 2002]. Rosenberg relies on several works for "theories of deterrence, insurance, law enforcement, rational choice analysis, and welfare economics." *Id.* at 831 n.1; *see generally* GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS (1970); A. MITCHELL POLINSKY, AN INTRODUCTION TO LAW AND ECONOMICS (2d ed. 1989); RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW (5th ed. 1998); STEVEN SHAVELL, ECONOMIC ANALYSIS OF ACCIDENT LAW (1987); Louis Kaplow & Steven Shavell, *Fairness Versus Welfare*, 114 HARV. L. REV. 961 (2001).

77. Rosenberg 2002, *supra* note 76, at 832.

78. *See, e.g.*, Richard A. Epstein, *The Consolidation of Complex Litigation: A Critical Evaluation of the ALI Proposal*, 10 J.L. & COM. 1, 2–5, 49–50 (1990); Richard A. Epstein, *Class Actions: Aggregation, Amplification, and Distortion*, 2003 U. CHI. LEGAL F. 475 (2003) [hereinafter Epstein, *Class Actions*]; Richard A. Nagareda, *Autonomy, Peace, and Put Options in the Mass Tort Class Action*, 115 HARV. L. REV. 747 (2002). In criticizing Rosenberg's position, Epstein argues that:

> Even if we reject (as current law manifestly does) the view that ex post compensation is irrelevant, powerful implications still flow for the governance of class action litigation. This position presupposes that the judgment should be collective and not individual, such that a person who objected to the strategies pursued by the class would be required to remain a class member on the ground that the economies of scale in running the class action would leave him better off than before. There is obviously a powerful paternalistic streak in this argument.

Epstein, *Class Actions*, *supra* note 78, at 494. Because the larger debate has been carried out elsewhere, this article will not review it in detail here.

79. Aside from this paragraph's brief discussion, this Article does not elaborate on the details of the Fried & Rosenberg framework. Those details can be found in Chapter 2 of their book. FRIED & ROSENBERG, *supra* note 6, at 13–36.

derstand an individual's preferences "under conditions of uncertainty, at a point in time before the person knows which of possible alternative fates will come to pass."[80] In this *ex ante* state, "each individual internalizes all possible fates of all possible people."[81] Because the individual internalizes all possible states of the world, the individual rationally desires a legal system that maximizes welfare in all possible situations the individual may find himself. In a 2002 article, Rosenberg emphasized the importance of the *ex ante* perspective as central to the argument for mandatory class action for mass torts:

> Essentially, this argument addresses the fundamental disjuncture between an individual's preferences ex ante—that is, before knowing whether one will suffer tortious injury, and if so, how strong the related claim will be—and ex post—after learning the "luck of the draw." Understanding how individual preferences change over time, particularly as individuals acquire knowledge, is central to the argument for mandatory mass tort class action.[82]

In the context of defendant class actions, the starting point for an *ex ante* approach is recognizing that *ex ante*, an individual does not know whether he/she will be on the plaintiff or defendant side, or whether he/she will be part of a large firm or in a large class of individuals. Thus, in the *ex ante* world, a rational, social-utility maximizing individual would have no reason to favor either 'plaintiff' or 'defendant' classes. In the context of music downloading, for example, an individual does not know if they will be an RIAA employee, a musician, a downloader of copyrighted music, a non-downloading user of the Internet, or some other individual that might be affected by a class action against those who download copyrighted music. In the context of corporate fraud, an individual does not know if they will be on the corporate board, working in the corporation's mailroom, holding stock in the corporation, or purchasing services produced by the firm. In the context of mass copyright violation (e.g., hundreds of thousands of pirated DVDs being sold across the globe), one does not know where in the supply chain they will be located.

Hamdani and Klement's analysis fails to consider this *ex ante* position. As a result, Hamdani and Klement's core thesis does not plant its roots as deeply as it could. Hamdani and Klement's "core thesis is that the fundamental justification for consolidating plaintiff claims applies with equal force to defendants."[83] Their fundamental justification is, "[i]n the plaintiff case, the cost of bringing a suit might dissuade victims from suing wrongdoers . . . [and this] failure to litigate undermines justice and

80. *Id.* at 14.
81. *Id.* at 15.
82. Rosenberg 2002, *supra* note 76, at 831.
83. Hamdani & Klement, *supra* note 9, at 689.

deterrence."[84] This fundamental justification, however, is not adequate. Justice and deterrence *may* be undermined if plaintiffs cannot bring their case, but it may also be that something other than a plaintiff class action will generate optimal deterrence for similarly-situated defendants. We need a more general theory to understand in what contexts the defendant class action is likely to be effective for achieving optimal deterrence.

The lack of a general theory is evident in Hamdani and Klement's choice to ground their analysis in the "standard justification for class actions."[85] The authors implicitly acknowledge their choice of the standard justification in a footnote. Citing the work of Rosenberg, they note, "The standard justification for class actions focuses on claims for insignificant amounts that would not be filed individually. However, that class actions are desirable even for larger claims as long as the common defendants enjoy economies of scale . . . ."[86] Beyond this citation, however, the authors do not discuss the Rosenberg position and why even large claim class actions may be desirable.

Nicole Johnson's recent extension of the Hamdani and Klement argument also fails to adequately consider fundamental principles. Johnson "takes the Hamdani and Klement proposal a step further and suggests that the class defense has a more expansive applicability, not only for achieving economies of scale and overcoming collective action problems in litigation, but perhaps more importantly in obtaining settlements."[87] The new settlement possibilities produced by aggregation of claims are important, but we need more general discussion of when such possibilities are likely to occur, and thus, when courts should look toward defendant class certification.

Nelson Netto has advanced the defendant class argument on the basis of Rosenberg and Fried's theory of collectivizing claims. Netto argues that "the optimal economy of scale for investment in litigation requires the compulsory reunion of the defendants and their defenses."[88] Similar in spirit to Netto's argument, I start from the *ex ante* perspective and build a series of propositions about what defendant class actions should seek to do.

---

84. *Id.* at 689–90.

85. *Id.* at 689 n.14.

86. *Id.* The authors cite David Rosenberg, *Mandatory-Litigation Class Action: The Only Option for Mass Tort Cases*, 115 HARV. L. REV. 831 (2002).

87. Johnson, *supra* note 4, at 218. Additionally, Johnson notes:

> In the recently settled suit between NTP and RIM, a consumer class defense would have allowed consumers, including large corporate firms that rely on BlackBerry devices for critical communication, to protect their interests and take action in their own defense. BlackBerry users might have obtained an earlier settlement or might have been assured that they could reach a settlement regardless of a standoff between the parties.

Id. at 224.

88. Netto, *supra* note 4, at 98.

In light of great uncertainty in the *ex ante* world, what can we say about individual preferences for design of a legal system? First, we can say that an individual will desire to maximize his utility "*across all possible states of the world.*"[89] Since an individual could end up either as defendant or plaintiff, this leads to the corollary that in the context of class actions, the individual will seek to maximize utility by maximizing *total utility of defendant and plaintiff.* In the case of traditional plaintiff class actions, this means that we are not only concerned with the reduction in harm to the plaintiff class, but also the *cost of reducing harm* as paid by the defendant. In the case of defendant class actions, the same logic is applicable; we should consider not only the harm/risk-reduction to the plaintiff, but also the cost of precautions to the defendant class.[90]

Second, we can say that defendant class actions should be considered in light of their future deterrent effect. I label this "forward looking" in order to distinguish it from jurisprudence and commentary that looks "backward" at pre-existing links between potential defendant class members. My position can also be seen, however, as going "all the way back" to the *ex ante* position. Regardless of which conception one uses—forward looking or a return to the *ex ante* world—the important point for defendant class actions is that we are not concerned primarily with *existing or previous* relationships between individuals/firms, but rather with the *likely future relationships between similarly situated individuals/firms that will result from a particular legal ruling.*[91]

A corollary of this second point is that courts should ask the following question: Will classing this group of individuals/firms be more effective for optimal deterrence than would the alternatives of individual proceedings or joining under Rule 20 of the Federal Rules of Civil Procedure? If the answer is "yes," then the court should certify the defendant class. If the answer is "no," then the court should deny certification.

By posing the question this way, the analysis invites a comparison to joinder. Rule 20 of the Federal Rules of Civil Procedure says that defendants can be joined if "any right to relief is asserted against them jointly, severally, or in the alternative with respect to or arising out of the same transaction, occurrence, or series of transactions or occurrences; and any question of law or fact common to all defendants will arise in the action."[92] Courts look to the number of defendants to determine whether

89. FRIED & ROSENBERG, *supra* note 6, at 17 (emphasis added).

90. In principle, this bears some resemblance to Learned Hand's famous negligence calculus: finding someone negligent when $B < PL$, where B = the "burden of precautions," P is the "probability of harm" and L is the "gravity of harm." Both formulas emphasize a type of cost-benefit analysis.

91. Current or previous relationships between individuals and firms would be important to the extent that they help us predict what would happen in the future. But, they should not be, in and of themselves, the standard for evaluating a defendant class.

92. FED. R. CIV. P. 20(a)(2)(A)–(B).

joinder is impracticable.[93] Presently, if the number of defendants is greater than forty, then joinder will generally be presumed to be impracticable.[94] Courts often look to class devices as an alternative if joinder is not possible.[95] As the U.S. District Court reasoned in *Flying Tiger Line, Inc. v. Central States*,[96] "[b]efore the Court takes the drastic step of certifying a defendant class; however, the joinder alternative should be investigated more thoroughly."[97] While courts have made the focus of their joinder analysis the number of defendants, I argue in the next section that we should compare the two options on the basis not only of numerosity, but also on the basis of more general group dynamics.[98]

My proposed approach also makes clear that the defendant class action is not necessarily—as seems to be suggested by Hamdani and Klement—a device to go after the "little guy."[99] In deciding whether or not to class the corporate executives of a failed financial firm, for instance, the approach advocated in this Article might well lead to the conclusion that they too should be classed. The reason would *not* be that they are too numerous or incapable of being joined under other rules, but rather that treating them as a class would better deter similarly-positioned executives in the future. If the executives know they will sink or swim as a class, then they have greater incentive to internally check up on one another. This would create a mechanism of self-governance that should improve deterrence. Forward-looking deterrence is not concerned with parceling out causation within the group. Critics might argue at this point that such an approach will fail to make the proper causal connections between harm-causing parties' actions and sanctions. To see why this will not be the case, we need to consider the second guiding principle, which is dynamic effects.

---

93. *See* Monaco v. Stone, 187 F.R.D. 50, 64 (E.D.N.Y. 1999).

94. *Id.* at 66 ("[A] class of more than forty members raises a presumption that joinder is impracticable.").

95. Stewart v. Abraham, 275 F.3d 220, 226 (3d Cir. 2001) (citing FED R. CIV. P. 23(a)).

96. No. 86-304 CMW, 1986 U.S. Dist. LEXIS 17409 (D. Del. Nov. 20, 1986).

97. *Id.* at *16.

98. This distinguishes my analysis from Netto, who writes that "[t]he defendant class action, certified on behalf of the defendant requirement, is superior to mandatory joinder, because mandatory joinder is impractical when the number of defendants is extent and a defendant class action does not need any drastic modification of actual statutes." Netto, *supra* note 4, at 105 n.186 (citing Charles Silver, *Comparing Class Action and Consolidations*, 10 REV. LITIG. 495 (1991) and Edward Hsieh, Note, *Mandatory Joinder: An Indirect Method for Improving Patent Quality*, 77 S. CAL. L. REV. 683 (2004)).

99. While not explicit on this point, the tone of Hamdani & Klement's article is that small, dispersed defendants need a device to help them fight a potentially over-bearing plaintiff. The language of their article is not neutral in this respect. For instance, in building what seems to the specter of groups like the RIAA, the authors write that "[m]ost alarmingly, plaintiffs can act strategically to exacerbate the problem confronting each defendant, further diminishing the incentives to go to trial." Hamdani & Klement, *supra* note 9, at 698.

B. Dynamic Effects

The principle of dynamic effects holds that we should consider all likely effects of the court's legal ruling. In the context of defendant class actions, this means that we should pay particularly close attention to the *group dynamics* that would operate if a court decided to class a group of defendant individuals/firms. One of the most important, but overlooked, dynamic effects of the class device is the creation of a new market for information generation. Drawing on the work of Michael Abramowicz, who has reviewed and made the normative case for integrating market mechanisms into legal proceedings, this section focuses on how incentive structures change when individuals are made members of a class. [100]

1. Group Dynamics

In determining whether it is marginally beneficial to class a group of individuals/firms as a defendant class, we have to know what the "baseline" group dynamics are; i.e., if the court did nothing to class the defendants, how would they likely act in the face of individual lawsuits? Up to this point in the Article, the proposed theory has laid out only similarities between plaintiff and defendant class actions. This is consistent with the argument that at a conceptual level, there is little to distinguish plaintiff and defendant classes. In contrast to the conceptual and theoretical similarities, however, at the level of group dynamics, defendant and plaintiff class actions have markedly different baselines. Specifically, I argue that individual plaintiffs are (without any judicial intervention) less likely than individual defendants to establish a "market relationship" with others in their group.

I define "market relationship" as broadly as possible. I take market relationship to mean any sort of relationship in which individuals/firms act (or react) either directly or indirectly in response to actions (or reactions) by other individuals/firms. This concept of market relationship considers not only traditional market elements such as collective action and price adjustments, but also social psychological elements such as herd mentality and the fundamental attribution error (where we fail to recognize the effects of situation in determining human behavior). It also emphasizes the ability of the market to produce information, and most importantly, information on relative contributions to harm by defendants or relative harm experienced by plaintiffs.

Defendant class actions have been promoted in the past few years as a solution to dispersed defendants each generating a small amount of damage through new technological means. Netto argues, for instance, in favor of mandatory defendant class actions as a response to mass produc-

100. *See generally* Michael Abramowicz, *The Law-and-Markets Movement*, 49 AM. U. L. REV. 327, 408–30 (1999).

tion and a "technologically savvy society with the propensity for massive unlawful behavior."[101] While defendant class actions may be useful in this context, it is important not to view the defendant class device narrowly as only a response to technological innovation. A defendant class may be useful more generally as an auction-like mechanism to produce information about relative contributions to harm.

Auction mechanisms are already used in a variety of legal contexts.[102] Auctions and exchange can serve important informational purposes. For instance, if plaintiffs are allowed to sell their claims to bidders, "[t]he price at which such shares trade in the secondary market provides an indication of the plaintiff's expected recovery at trial and thus may dampen parties' abilities to puff in pretrial settlement bargaining."[103] In the context of patent buy-outs, Michael Kremer has proposed that an auction be used to determine the value of the patent.[104] Applying similar reasoning to defendant class actions, the class action device may be useful as a means of generating information about relative harms. [105] That information can then be used for settlement purposes.

To illustrate how this information production might play out, consider a simple case in which two firms, A and B, are both defendants in a case where negligence has caused 100 units of damage. The plaintiff firm knows that it experienced damage of 100, but it does not know that Firm A caused thirty percent of the damage and Firm B is responsible for seventy percent of the damage. To see how collectivization can be useful even with just two firms as defendants, examine the pay-off matrices with and without the defendant class device that are presented in Table 2.

Without knowing relative contributions to harm, and without joint and several liability, Firm B will have an incentive in the settlement stage to settle for fifty percent of the damage because Firm B knows that if it goes to trial, it will be shown liable for seventy percent of the harm. Firm A, however, faces a different incentive structure. Firm A would rather litigate than settle for 50 because litigation will lead to liability for only 30 units of the harm. If Firm A and Firm B are treated separately, then the plaintiff (who we assume here knows nothing of the actual relative contributions) will likely settle with Firm B and proceed to litigate

---

101. Netto, *supra* note 4, at 59.

102. *See generally* Abramowicz, *supra* note 100, at 335–52. Over twenty years ago, Marc Shukaitis proposed a market for personal injury tort claims. Marc J. Shukaitis, *A Market in Personal Injury Tort Claims*, 16 J. LEGAL STUD. 329, 329–30 (1987).

103. Abramowicz, *supra* note 100, at 359–60.

104. Michael Kremer, *Patent Buyouts: A Mechanism for Encouraging Innovation*, 113 Q.J. ECON. 1137, 1146–47 (1998).

105. The informational benefits of defendant class actions were recognized by the Ohio Supreme Court in 1990, which noted that "[a] class suit may be especially useful in a case where putative class members refuse to identify themselves or deliberately act to avoid being controlled in law." Planned Parenthood Ass'n of Cincinnati, Inc. v. Project Jericho, 556 N.E.2d 157, 164 (Ohio 1990).

with Firm A. The non-class result, shaded in gray in Table 2, is thus a transfer of 80 from the two firms to plaintiff.

Now consider what happens if the two firms are considered a single class, held jointly and severally liable for the damage. Knowing that they face a total payout of 100 if they litigate or settle, the choice will be to settle. But now in the settlement stage, Firm A has an incentive to make clear its contribution to harm, either through proceedings against Firm B (more likely) or through negotiations with Firm B. Whichever route is taken, information will be generated about relative contributions to harm; information that would not have been generated in the world without class certification.

Table 2. Pay Off Matrices With and Without Defendant Class Action

NO CLASS ACTION

		Firm A (30%)	
		Litigates	Settles
Firm B (70%)	Litigates	-30, -70	-50, -70
	Settles	**-30, -50**	-50, -50

WITH DEFENDANT CLASS ACTION

		Defendant Class: A + B (100%)	
		Litigates	Settles
		-100	**-100**

The 2 x 2 matrix is admittedly greatly over-simplified, but it suggests a general point that incentives *amongst the defendants* change when they are held liable *as a class*, and not just as individuals. For a group of N defendants, the N defendants in the class have an incentive to work out their proportional liability to the plaintiff. Of course, enforcement remains a challenge, and I will address that challenge in Part III of the paper on system design.

I should emphasize that my suggested approach does not always lead to defendant class action certification. Rather, it looks for the marginal value that the class device potentially offers. In cases where all defendants are jointly and severally liable, the class device will not significantly change the incentive structure already in place. A defendant class device may also not be useful if all defendants are already bound

through a single party. For instance, in *Gaunt v. Brown*,[106] the U.S. District Court correctly concluded that since the case was being brought against the Attorney General, there was not a need to certify a class of local boards of elections (in a case challenging the age requirement for elections).[107] In this case, since the entirety of the defendant class was bound by law to follow the Ohio Attorney General's directives, the marginal value of the class device was zero.[108]

2. Comparing Defendant and Plaintiff Group Dynamics

I now consider a broader set of possible plaintiff and defendant dynamics. As a basis for discussion, Table 3 considers the possible combinations that might occur in a world where there are four types of groups: (1) single firms/individuals, (2) a single dominant firm/individual, (3) an intermediate number of firms/individuals, and (4) a large number of firms/individuals. I assume that each group could find themselves either on the plaintiff (harm bearing) or defendant (harm causing) side. This generates 16 scenarios to consider. I sketch out what I believe would be the "baseline" result; the likely result if there was no judicial certification of a class on either side. I then offer my suggested "class outcome." In other words, what would likely happen if the court decided to certify a defendant class, a plaintiff class, or both?

The table reinforces that our focus should be on the *difference between the baseline and class outcome* columns. This is the marginal value added by class certification. I have arranged the table so that every other row flips the defendant and plaintiff sides. To make the table easier to read, and to isolate the differences between defendant and plaintiff class actions, I have highlighted the rows where defendant class actions would be a possibility.

106. 341 F. Supp. 1187 (S.D. Ohio 1972).

107. *Id.* at 1193. The court argued:

> We agree that if plaintiffs prevail this would be an appropriate case to designate as a plaintiff class action. However, we are not persuaded that it should be designated as a defendant class action if plaintiffs prevail, inasmuch as the Secretary of State of the State of Ohio is a party-defendant, and his duties are to advise members of local boards of elections as to proper methods of conducting elections. Also, the Secretary of State has the further duty to "compel the observance by election of officers in the several counties of the requirements of the election laws." Since the Secretary has the duty and power over all the members whom plaintiffs would have us include in a defendant class action, the need for a defendant class action is not apparent.

Id. at 1192–93 (citations omitted).

108. *See id.*

Table 3. Comparison of baseline and predicted class-outcomes for selected configurations of plaintiff and defendant groups

No.	PLAINTIFF SIDE	DEFENDANT SIDE	BASELINE	CLASS OUTCOME
	Who is suffering the harm/damage ?	**Who is causing the harm/damage ?**	**With no judicial class certification, what do we expect to see?**	**What changes if the court classes defendants, plaintiffs, or both?**
1	Single firm / individual	Single firm / individual	Traditional tort outcome: single party vs. single party	Optimal deterrence achieved at baseline; class certification provides no additional value
2	Single firm / individual	Dominant firm, controlling >50% market share	Case against the dominant firm, dominant firm will litigate fully	Optimal deterrence achieved at baseline; class certification provides no additional value
3	Dominant firm, controlling >50% market share	Single firm / individual	Dominant firm will prosecute fully, defendant the same	Optimal deterrence achieved at baseline; class certification provides no additional value
4	Single firm / individual	Intermediate number of firms, controlling <50% market share	Plaintiff will enjoy economies of scale; Defendants likely to bind together because they will be joined as named defendants (maybe conspiracy alleged), and will see benefits of collective defense	If court certifies defendant class action, optimal deterrence will be achieved; But even without court certification, defendants may bind together when they are sued

5	Intermediate number of firms, controlling <50% market share	Single firm / individual	If firms cannot overcome collective action problem, they will either drop the case, or will bring a case without enough resources to litigate fully; Defendant will be able to enjoy economies of scale	If court certifies plaintiff class action, optimal deterrence will be achieved [a]
6	Single firm / individual	Large number of firms / individuals, each controlling very small market share	If plaintiff cannot figure out who is causing the harm, may not be able to bring enough suits; Defendants, when sued, will not be able to match resources with the plaintiff [b]	If court certifies defendant class action, optimal deterrence achieved (so long as economic incentive issues are corrected for) [b]
7	Large number of firms / individuals, each controlling very small market share	Single firm / individual	Plaintiffs will not be able to overcome collective action problem (the traditional plaintiff class action); Defendant will enjoy economies of scale	If court certifies plaintiff class action, optimal deterrence will be achieved
8	Dominant firm, controlling >50% market share	Dominant firm, controlling >50% market share	Dominant firm on both sides should be able to kick their market into gear	Optimal deterrence achieved at baseline; class certification not necessary
9	Dominant firm, controlling >50% market share	Intermediate number of firms, controlling <50% market share	Dominant plaintiff has resources and incentive to bring suit; defendant firms	If court certifies defendant class action, optimal deterrence will be achieved; But

			will likely find it beneficial to work together as named defendants in the same suit	even without court certification, defendants may bind together when they are sued
10	Intermediate number of firms, controlling <50% market share	Dominant firm, controlling >50% market share	If firms cannot overcome collective action problem, they will either drop the case, or will bring a case without enough resources to litigate fully; Defendant firm will kick market into gear	If court certifies plaintiff class action, optimal deterrence will be achieved; but class certification may not be necessary
11	Dominant firm, controlling >50% market share	Large number of firms / individuals, each controlling very small market share	If plaintiff cannot figure out who is causing the harm, may not be able to bring enough suits; Defendants, when sued, will not be able to match resources with the plaintiff [b]	If court certifies defendant class action, optimal deterrence achieved (so long as economic incentive issues are corrected for)[b]
12	Large number of firms / individuals, each controlling very small market share	Dominant firm, controlling >50% market share	Plaintiffs will not be able to overcome collective action problem (the traditional plaintiff class action); Defendant will be able to defend itself and generate informational market	If court certifies plaintiff class action, optimal deterrence will be achieved

13	Intermediate number of firms, controlling <50% market share	Intermediate number of firms, controlling <50% market share	Indeterminate. Both sides may face collective action problems, but both have a chance to overcome them.	Class action certification, on either side, will promote optimal deterrence if collective action problems are serious.
14	Intermediate number of firms, controlling <50% market share	Large number of firms / individuals, each controlling very small market share	Large number of defendants makes it more difficult for plaintiffs to overcome collective action problems	Certification of both defendant and plaintiff classes may lead to optimal deterrence
15	Large number of firms / individuals, each controlling very small market share	Intermediate number of firms, controlling <50% market share	Plaintiffs are not likely to overcome collective action problems	Certification of plaintiff class would promote optimal deterrence; Defendants may need class certification as well
16	Large number of firms / individuals, each controlling very small market share	Large number of firms / individuals, each controlling very small market share	Neither side will be able to overcome collective action problems	Certification of both classes is required to obtain optimal deterrence

NOTES: a. *See* Rosenberg (2000), *supra* note 32. b. *See* Hamdani & Klement (2005), *supra* note 6.

Perhaps the most interesting (and contentious) action in Table 3 occurs when we compare rows 4-5 and rows 9-10. In each case, we are flipping the "intermediate" number of firms from the defendant to the plaintiff side. The crux of my argument is that it is more likely for this mid-size group to overcome collective action problems when they are on the defendant side. The reason for this logic is straightforward; on the defendant side, parties do not have to initiate the proceedings. In fact, if the plaintiffs name them all as defendants in a suit, they have had much of the informational work of identification done for them. To the extent that this happens, defendants already take concerted actions when sued by a plaintiff. The court's class certification would be functionally redundant, and the marginal value of class certification would be minimal.

Defendant class actions are likely to have more value when defendants are less capable of somehow binding themselves together. This failure is most likely to happen in two scenarios: (1) when identification and monitoring is not possible or practical, or (2) when enforcement of group "rules" is not possible or practical. Both challenges open the door for significant free riding. In Section III, on system design, I consider both of these issues and possible legal remedies to correct for them.

3. A Closer Look at What Binds Individuals in a Potential Defendant Class

Another way to think about the difference between plaintiff and defendant class actions is to see that in the plaintiff class action case, the individual plaintiffs are passive harm-takers. In the defendant class action cases, the individual defendants are active harm-makers. This distinction leads to important differences between plaintiff and defendant classes in terms of the *ex ante* market relationships that may develop. Three types of relationships are likely to exist between individual defendants: (1) they are all conducting market transactions with a single (or small set of closely related) firms; (2) they are all legally bound in a government organization; (3) they are all voluntarily bound in an organization of their own making. In the first case, adjustments can be made via price levels. In the second and third cases, contracting can be worked out through the governing organizations. It is only when none of these relationships exist that we see a need for defendant class actions.

What distinguishes the harm caused by small defendants, as opposed to the harm caused by large firm defendants, is the indirect nature of the small defendants' action. In almost every case where defendant class actions seem apt, there is a "market" intermediary. For instance, in the context of securities fraud, the defendant security underwriters were not hired by individual plaintiffs, but were working through some firm. In the context of other corporate fraud, middle managers and others in the firm who acted wrongly were all bound via contract to the same employer. In the context of state/local officials, they are causing harm by virtue of their role within the state government/legal system. In the context of music downloading, individuals are working with the help of several intermediaries: their Internet Service Provider, their software maker, etc.

To make this argument clearer, consider these two contrasting hypotheticals. First, consider a standard plaintiff class action in which a firm has a poorly constructed factory, which sits on the corner of a busy intersection. Everyday bricks fall off the building and cause damage to passing cars. Because the damage is always minor, the cars never stop, and no potential plaintiff ever brings a case. A plaintiff class action would be necessary here because there is likely no *ex ante* market relationship between those who have been harmed. They were each harmed

directly by the firm, with no intermediary; the brick fell directly on their car.

Now consider a second hypothetical. A big firm has an old factory that they no longer use. The factory, however, has bricks that are very valuable if taken and re-sold. Imagine that individuals go up to this factory and remove one brick at a time. No single person takes more than one brick. Setting aside, for now, the question of what precautions the firm could take to stop this, let's consider the relationship between these individual brick-stealers. It could be that each brick-stealer randomly wandered up to the factory, in the same way that the car drivers randomly drove past the brick-drop intersection. But it is more plausible that the brick-stealers share common traits; common traits that make them more likely to belong to one of the three types of *ex ante* markets laid out above. In this case, they are probably all selling their bricks on similar markets. They could also belong to a brick collector's society.[109]

When relationships such as these exist between defendants, the defendant firm can find convenient entry points for litigation. It need not necessarily resort to a defendant class action because it can go after the agency, organization, or other binding agent between the defendants. For example, when authors and publishers of the American Society of Composers tried to move against the Girl Scouts for copyright infringement—for singing copyrighted songs around the campfire—the plaintiff authors and publishers did not have to go after thousands of young Girl Scout members nationwide.[110] Instead, the plaintiffs went directly to the national organization that binds the girl scouts together.[111] The push toward potentially making Internet Service Providers ("ISPs") liable in illegal music downloading can be understood in a similar vein.[112]

C. Aggregate Analysis

The aggregate analysis principle adds an extra layer to both the forward-looking deterrence and the dynamic effects principles. The aggregate analysis principle holds that we should look at deterrence and dynamic effects *at an aggregate, system-wide level.* In this section I will show how the analysts discussing both the Internet and corporate fraud examples have missed this aggregate picture.

---

109. The hard case, a version of which I consider in the system design section, would be if they each found the brick valuable for some reason that didn't require re-sale, e.g., as a mantle piece.

110. *See* Ken Ringle, *ASCAP Changes its Tune; Never Intended to Collect Fees for Scouts' Campfire Songs, Group Says*, WASH. POST, Aug. 28, 1996, at C03.

111. *See id.*

112. See, for example, *In re* Aimster Copyright Litig., 334 F.3d 643, 652 (7th Cir. 2003), where Judge Posner upheld the grant of a preliminary injunction against a website that provided file-sharing services that were ultimately used for violation of federal copyright law by individual members of the public.

1. Cost Benefit Considerations

At the outset, a distinction should be made between (1) an overall cost-benefit valuation; i.e., do we want to reduce the activity or care levels of downloading of copyrighted material? And, (2) a comparison of the cost of precautions versus the benefit of harm reduction associated with that precaution. Conflating these two distinct evaluations may lead to some confusion. To make each stage clear, I will label the first cost-benefit analysis process "valuation" and the second (applying the Fried/Rosenberg framework), "determining optimal precautions."

In both stages, an aggregate perspective is important. At the valuation stage, aggregation means we must determine overall how much utility is being lost, and how much utility is being gained from a particular activity which individuals or firms are engaging in. Class actions factor into this analysis in a preliminary way; it is more likely that we will have aggregate analysis when there is a class action then when there is not. The reason is that courts will have to consider welfare/utility across *all* members of the class, not just the ones listed on the court documents as representatives. When adjudicating, courts will weigh both sides at the aggregate level.

2. Aggregate Analysis of Internet Governance

When discussions of Internet governance are raised, popular (and to a large extent academic) discussion has focused on illegal file sharing.[113] Jonathan Zittrain argues that such a narrow focus is greatly misguided: "Current scholarship about 'Internet governance' largely fails to appreciate this larger picture, rendering most of its deliberations absurdly narrow, with public policy recommendations that have a near-uselessly short shelf life . . . ."[114] Zittrain is announcing an aggregate analysis principle, suggesting that analysts should be considering more than simply the issue immediately before them.

The aggregate analysis principle has great bite in the Internet context because of the Internet's great "generativity." Zittrain defines "generativity" as a function of (1) how deeply a technology leverages a set of possible tasks; (2) its "adaptability to a range of different tasks"; (3) its "ease of mastery"; and (4) its "accessibility."[115] The Internet provides a

113. For a summary of this literature, see generally JONATHAN ZITTRAIN, THE FUTURE OF THE INTERNET AND HOW TO STOP IT (2008). The three cases in this area cited most often are: MGM Studios Inc. v. Grokster, Ltd., 545 U.S. 913 (2005); *Aimster*, 334 F.3d 634; A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004 (9th Cir. 2001).

114. Jonathan Zittrain, The Future of the Internet—And How to Save It 30 (Feb. 2005) (unpublished manuscript) (on file with the Denver University Law Review) (draft version 1.6 of the Zittrain's book published in 2008).

115. *Id.* at 1981.

new "generative grid" whose potential is still being realized.[116] What does talk of a generative grid, or of the Internet so generally, mean for defendant class actions? It means that our analysis of defendant class actions in the Internet context cannot rest solely on the costs of file-sharing. Rather, it must also consider how file sharing may contribute to generative goods.

Hamdani and Klement provide an example of analysis that stops short. They introduce considerations of overall social welfare in their analysis of a proposed class defense mechanism, but do not carry out an aggregate analysis. In considering a hypothetical lawsuit from the RIAA against an individual, for instance, Hamdani and Klement detail how an individual's incentive will be to settle for $3,000, even when they have done nothing illegal.[117] They argue, correctly, that "the ex post settlement decisions of defendants impact the ex ante decisions of other Internet users" whether to download music.[118] But it does not necessarily follow, as they argue in the next sentence, that "[w]hen defendants settle even when they may have a good defense, there is a considerable risk of excessively deterring music downloads by . . . Internet users."[119] The reason it does not necessarily follow is that optimal deterrence must be determined at an aggregate level. In other words, we may want to deter perfectly legitimate uses of file-sharing (and therefore make some innocents pay $3,000) if we believe that it will benefit society overall (by keeping the bad guys out of the game). By the same logic, we may want to allow illegal file-sharing by some crooks, if we believe that it will benefit society overall (by letting the good guys stay in the game).

This Article takes no substantive position on what the legal rule should be about file sharing; i.e., whether we should hold Internet Service Providers (ISPs) liable, or whether the Digital Millennium Copyright Act ("DMCA") properly assigns liability.[120] This Article does, however,

---

116. The generative grid phrase is Zittrain's. *See id.* at 1975. Scholarship is emerging to try and assess the myriad of effects the Internet has had on our lives. *See, e.g.*, Eugene Borgida & Emily N. Stark, *New Media and Politics: Some Insights From Social and Political Psychology*, 48 AM. BEHAV. SCIENTIST 467, 467 (2004) (investigating "the extent to which the Internet is providing . . . an important and increasingly influential forum for acquiring politically relevant information").

117. Hamdani & Klement, *supra* note 9, at 701. The reason is that they face a decision between settling for $3,000 or going through a lawsuit for $50,000 just to avoid payment.

118. *Id.*

119. *Id.*

120. The DMCA was signed into law in 1998, and among other things, holds ISPs liable for their users' illegal actions if the ISPs do not follow guidelines laid out by the Act (e.g. removing offensive material, reporting violations, etc.). *See* 17 U.S.C. § 1201 (2006). I also take no view here as to whether it is in the Record Company's best long term interest to prosecute file-swappers. Some have suggested that alternative strategies may be better suited:

> Coverage of the lawsuits could hurt as much as help the anti-piracy crusade. Anthony Prapkanis, a University of California-Santa Cruz professor of social psychology, says that while people may be sympathetic to the music industry's plight, "the image is out there of the bully ganging up on people with the least amount of money, the rich taking from the poor."

Jefferson Graham, *RIAA Lawsuits Bring Consternation, Chaos*, USA TODAY, Sept. 10, 2003, at 4D.

argue that we should assess the DMCA, and related decisions such as *In re Aimster Copyright Litigation*,[121] *A&M Records, Inc. v. Napster, Inc.*,[122] and *MGM Studios Inc. v. Grokster, Ltd.*,[123] under the aggregate analysis principle. At a minimum, this will involve incorporation of several strands of literature such as economic analysis of the effects of individuals' copyright infringement,[124] as well as analysis of actual usage of a file-sharing program, especially estimates of usage for illegal versus legal purposes.[125] More importantly, such aggregate analysis also demands that courts take seriously the technological aspects of the cases they are dealing with. In the context of file-sharing, for instance, the future is not in limiting the ability to trade, but in limiting the ability to *play*, via Digital Rights Management ("DRM").[126] Mark Stefik has observed that despite the fact that "[e]veryday experience with computers has led many to believe that anything digital is ripe for copying . . . [b]ehind the scenes . . . technology is altering the balance once again."[127]

---

121. 334 F.3d 643 (7th Cir. 2003).

122. 239 F.3d 1004 (9th Cir. 2001).

123. 545 U.S. 913 (2005).

124. In the context of music file-sharing, there remains an empirical debate over the effect of illegal file sharing on music sales. *See, e.g.*, Kai-Lung Hui & Ivan Png, *Piracy and the Legitimate Demand for Recorded Music*, 2 CONTRIBUTIONS TO ECON. ANALYSIS & POL'Y (2003), *available at* http://www.bepress.com/bejeap/contributions/vol2/iss1/art11/ ("[T]he demand for music CDs decreased with piracy, suggesting that 'theft' outweighed the 'positive' effects of piracy. However, the impact of piracy on CD sales was considerably less than estimated by industry."); Stan J. Liebowitz, *File-Sharing: Creative Destruction or Just Plain Destruction?* 32 (Ctr. for the Analysis of Prop. Rights, Working Paper No. 04-03, 2004), *available at* http://som.utdallas.edu/centers/capri/documents/destruction.pdf (finding that the evidence seems compelling that file-sharing is responsible for the recent large decline in CD sales for which it has been blamed); Felix Oberholzer & Koleman Strumpf, The Effect of File Sharing on Record Sales: An Empirical Analysis (Mar. 2004), *available at* http://www.unc.edu/~cigar/papers/FileSharing_March2004.pdf (unpublished working paper) ("Downloads have an effect on sales which is statistically indistinguishable from zero, despite rather precise estimates. [But], these estimates are of moderate economic significance and are inconsistent with claims that file sharing is the primary reason for the recent decline in music sales."); Rafael Rob & Joel Waldfogel, *Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students* (Nat'l Bureau of Econ. Research, Working Paper No. 10874, Oct. 2004), *available at* http://www.nber.org/papers/w10874.pdf ("[D]ownloading reduces their per capita expenditure (on hit albums released 1999-2003) from $126 to $100 but raises per capita consumer welfare by $70."); Alejandro Zentner, Measuring the Effect of Online Music Piracy on Music Sales (2004), *available at* http://economics.uchicago.edu/download/musicindustryoct12.pdf (unpublished working paper) (finding that peer-to-peer usage reduces the probability of buying music by an average of 30 percent, and that without file sharing, sales in 2002 would have been around 7.8 percent higher).

125. Some evidence from Russia suggests that even amongst young people, use of the Internet for illegal file-sharing is not a common activity. Oxana Palesh, Kasey Saltzman & Cheryl Koopman, *Internet Use and Attitudes Towards Illicit Internet Use Behavior in a Sample of Russian College Students*, 7 CYBERPSYCHOL. & BEHAV. 553, 553 (2004) ("Among Internet users, most reported having Internet access either at home or at a friends' home, and 16 % reported having Internet access from work, school, or a computer center. Among Internet users, the main purpose was for school-related activities (60%), followed by e-mail (55%), entertainment (50%), chatting (24%), and searching for pornography (6%).").

126. For an introduction, see ROB FRIEDEN & CHRISTY CARPENTER, DIGITAL RIGHTS MANAGEMENT 2, 6 (2004).

127. Mark Stefik, *Trusted Systems*, SCI. AM., Mar. 1997, at 78, 78.

If courts are not aware, or deliberately choose to avoid discussion of what's going on "behind the scenes," then their rulings and analyses are not only likely to be out-dated, but also could be seriously flawed. For instance, what if DRM technology had already advanced to a stage where recording artists could protect (with great assurance) everything they wanted to, but courts (unaware of this development) went ahead with a legal regime that severely limited file sharing? The result would be over-deterrence. On the other hand, if courts errantly believed that DRM had reached a point where the state-of-the-art was to produce files incapable of being pirated (when in fact this was not the case), they would under-deter file-swapping. The substantive analysis is beyond the scope of this Article, but I have aimed to demonstrate that if courts do not take technological considerations into account, they violate the aggregate analysis principle, and likely produce sub-optimal outcomes as a result.

3. Aggregate Analysis of Corporate Wrongdoing

At first glance, the high-profile corporate wrongdoing over the past few years may seem an odd place to think about defendant class actions. The defendants are not numerous, hard to identify, or judgment proof. Why, then, should we consider defendant class actions a potentially useful tool? The answer, as it did in the Internet context, centers on the realization that there is something more going on here than simply the actions of the named defendants. In the Internet case, that "something more" is more readily identifiable: complex and changing technologies are clearly tied into the cases at bar. In the corporate fraud cases, the "something more" is subtler.

Drawing on social psychology and research on the corporate environment, the "something more" that a defendant class action can aim its reach at is the "situation" or "corporate climate" that may contribute mightily to fraud and wrongdoing.[128] There is a longstanding consensus amongst social psychologists that we commit a "fundamental attribution error" in attributing actions to individual choices, rather than to situational pressures. As articulated by Phillip Zimbardo and Michael Leippe, "we tend to look for the person in the situation more than we search for the situation that makes the person."[129] The value of a defendant class action is that it has the potential to get at the "situation" because it will implicate virtually everyone working in the office.

---

128. *See* Jon Hanson & David Yosifon, *The Situation: An Introduction To The Situational Character, Critical Realism, Power Economics, And Deep Capture*, 152 U. PA. L. REV. 129, 201, 221–30 (2003) (providing an introduction to social psychology literature in the corporate law context); *see also* Christopher W. Williams, Paul R. Lees-Haley & J. Randall Price, *The Role of Counterfactual Thinking and Causal Attribution in Accident-Related Judgments*, 26 J. APPLIED SOC. PSYCHOL. 2100, 2109–10 (1996) (providing an introduction to the implications of the attribution theory in the legal arena).

129. PHILIP G. ZIMBARDO & MICHAEL LEIPPE, THE PSYCHOLOGY OF ATTITUDE CHANGE AND SOCIAL INFLUENCE 93 (1991).

Research and expert commentary on the corporate environment suggests that situational pressures to commit wrongs are indeed intense. When he talked about the "numbers game" that corporate executives sometimes play, former SEC chairman Arthur Levitt suggested "that almost everyone in the financial community shares responsibility . . . [and] [c]orporate management isn't operating in a vacuum. In fact, the different pressures and expectations placed by, and on, various participants in the financial community appear to be almost self-perpetuating."[130] One of the most comprehensive studies of moral action in the workplace is Robert Jackall's study, *Moral Mazes*.[131] Jackall engaged in extensive case studies of two firms, and found that most middle managers would sacrifice their own morals in order to fit in: "Team play also means . . . 'aligning oneself with the dominant ideology of the moment,' or . . . 'bowing to whichever god currently holds sway.'"[132]

If it is the case that it is not just a few top executives that are contributing to the harm caused by the firm, then a legal regime which points liability solely toward those CEOs is not likely to achieve optimal deterrence. Consider *Federal Insurance Co. v. Tyco International Ltd.*,[133] where separate actions were brought against former Tyco CEO Dennis Kozlowski, former chief lawyer Mark Belnick, and former CFO Mark Swartz.[134] From a deterrence perspective, members of society (and most especially Tyco shareholders) did not care who actually cooked the books. What society wants is for this sort of firm behavior not to happen again in the future, by Tyco, *or by any other firm.* In order to achieve that deterrence objective, we must have an understanding of the causal factors for the fraud. To the extent that it was not just a few "bad apples," but instead is in part driven systematically by certain kinds of corporate cultures, we want a legal device that can possibly change those cultures. A defendant class action might do that. In operation, if future members of a firm knew that they could be held liable (as a defendant class member) for any harm caused by the firm, it seems more likely that they would stand up to their bosses when asked to do illegal tasks.

4. Additional Comments on Aggregate Analysis

In response to likely concerns, two additional comments should be made in regards to aggregate analysis. First, is aggregate analysis too much for the courts to handle? I believe not, as courts (themselves and in conjunction with administrative agencies) already engage in substantial,

130. Arthur Levitt, Chairman, Sec. & Exch. Comm'n, Remarks at the New York University Center for Law and Business (Sept. 28, 1998).

131. SIMON JACKALL, MORAL MAZES: THE WORLD OF CORPORATE MANAGERS (1988).

132. *Id.* at 52.

133. 422 F. Supp. 2d 357 (S.D.N.Y. 2006).

134. *Id.* at 360.

aggregate cost-benefit analysis.[135] It can also be said that even if courts at present are not well-equipped to handle these sorts of analyses, there may be other parts of the system that courts can out-source to carry out the analysis. The Government Accountability Office ("GAO") frequently engages in these sorts of analyses.

My final comment is one that bends in a normative direction. When taken together as a pair, the Internet and corporate fraud examples make it clear that defendant class actions are not designed to go after a particular kind of group, i.e., "the little guy" or the "big, bad corporation." Rather, the defendant class action is a neutral tool that can be employed whenever it is needed to kick-start informational markets into gear.

III. SYSTEM DESIGN

This section of the paper identifies the major challenges courts face in implementing defendant class actions. Although the challenges are significant, I build on the proposals made by Netto and put forth a number of system design elements which may make defendant class actions more feasible and more capable of achieving the objective of optimal deterrence.

In addition to the Internet and corporate fraud examples which I have already discussed, this section will also address a third, more difficult, type of case: the case where there are defendants who appear to have no connection to each other. To make this hard case concrete, let's consider the following scenario. One-hundred thousand individuals across the world illegally sneak a camera into their local movie theatre and digitally record a blockbuster movie. They then show this digital movie to their friends and family, who consequently do not pay for either movie admission or for the DVD when it is released. This is a case where there is no discernible "market" relationship between any of the defendants. Note that it is not the size of the class that matters, but the relationship between them. There could be one million illegal tapers of the movie, but if they all acted independently there would still be no easy way to tie them together as a class. As I proceed with my discussion of system design, I will return to this hard case and how the general theory of defendant class actions should be applied to it.

A. Preliminary Considerations

A discussion of defendant class action implementation must begin with a discussion of principles. In this Section I lay out several funda-

135. *See* Robert W. Hahn, *Policy Watch: Government Analysis of the Benefits and Costs of Regulation*, 12 J. ECON. PERSP. 201, 201–10 (1998) (discussing the cost-benefit analysis in the government context); *see also* David Whiteman, *The Fate of Policy Analysis in Congressional Decision Making: Three Types of Use in Committees*, 38 W. POL. Q. 294, 297 (1985) (discussing aggregate principles used in the legislative decision-making process).

mental functions of tort law that the defendant class action must serve. To be sure, these are not the only functions of tort law. Nevertheless, they serve as a useful starting point for constructing, from the bottom up, a system in which the defendant class action can play a central role.

1. Insurance and Redistributive Functions of Tort Law

The tort system serves an insurance and redistributive function as well as a deterrence function.[136] In the context of defendant class actions, if the harm to the plaintiff can be identified, it does not seem that having a large number of small harms (as opposed to a single large harm) should affect insurance availability or premiums. If there were a market for these insurance claims, this situation might be different because having a larger number of smaller claims would make it more difficult for insurers to get paid.[137] Questions of redistribution are taken up again under the issue of fee-shifting and making sure that class defendants have proper economic incentives to fully litigate a defense for the entire class.

2. Deference to the Market and Legislative Bodies

I adopt the position that as a guiding principle, courts should be deferential to the market they find in operation. As Fried and Rosenberg observe, "[n]o logical impediment exists to the market's serving as a full substitute for legal intervention to achieve the social objective of ensuring optimal precautions."[138] Because the cost-benefit calculations, especially at the aggregate level, can be quite complicated, I also take the position that courts should be deferential to legislatures and administrative bodies that have conducted research on particular issues. Where courts see that legislatures are captured by interest groups whose goals may not be based on objective analysis and research, and when legislatures are not adhering to the principle of aggregate analysis, then they should take more independent actions.

3. Activity and Care Levels

Throughout considerations of system design, it is important to keep in mind the distinction between activity levels and care levels. This is a point seemingly missed by Hamdani and Klement. Using the Hamdani and Klement example, individuals may react to RIAA litigation in one of two general ways. First, they may simply reduce their activity level. This is the only possibility that the authors consider.[139] But second, individuals may react to RIAA litigation by increasing their care level. They may

136. *See* FRIED & ROSENBERG, *supra* note 6, at 55, 69–70.

137. David Rosenberg, *Deregulating Insurance Subrogation: Towards an Ex Ante Market in Tort Claims* 1 (Harvard John M. Olin Ctr. for Law, Econ., & Bus. Discussion Paper No. 395, 2002), *available at* http://www.law.harvard.edu/programs/olin_center/. I leave this question for another day, as currently such a market does not exist.

138. *See* FRIED & ROSENBERG, *supra* note 6, at 45.

139. *See* Hamdani & Klement, *supra* note 9, at 724.

take extra measures to insure that they are not found liable. This care level adjustment may take one of two forms. It may take the traditional form of trying to avoid the harm; i.e., downloading only with approved programs. But it may also take another form of trying to avoid detection. The distinct possibility of this "circumvention care" is particularly important to consider in the context of defendant class actions involving technology.

B. Identification & Monitoring

Identifying exactly who is generating the harm/risk may arise as an acute problem in the defendant class action context. Initially, there are several distinctions to make. First, in order to work effectively, class members need to identify not only who is causing the harm, but also how much marginal contribution is being made by each member. Monitoring can be introduced here as a form of "repeated identification"—re-evaluating on a regular basis who is in the market and what their market share is. Depending on the stability and fluidity of the market, this monitoring may be more or less costly.

A second distinction to make is between "ability" and "feasibility" to identify harm/risk producers. Prohibitively high identification costs may make it infeasible for identification to occur in some situations when it is theoretically possible (in a costless world). Putting these two concepts together, the identification problem can be considered along a continuum. Table 4 provides a rough outline of the scope of this problem.

Table 4. The Scope of the Identification Problem

Perfect ID	Strong ID	Mid-Strong ID	Mid-Weak ID	Weak ID	No ID
Know who caused the harm and each party's marginal contributions	Know who caused the harm, a little less sure of marginal contributions	Not entirely sure who caused harm, but can narrow it down, and can do the same for marginal contributions	Know the general "group" of people who caused the harm, but not the specific individuals in the "group," and know nothing of marginal contributions	Not entirely sure which "groups" are responsible, and have no idea of marginal contributions to harm	Do not know who caused the harm

1. Legal Tools to Address the Problem of Identification

Looking at Table 4, the goal of the legal system should be to enable parties to move as far as possible to the left, toward the ideal of perfect identification. There are three plausible ways that the legal system might improve identification of defendants and their relative contributions to harm. As a first cut, the legal system can use sub-classes to reduce its workload. Sub-classes will be most beneficial when it is easier to identify the marginal causal contribution of some members of the defendant class, relative to others. In practice, courts have carved out sub-classes in larger defendant class actions since at least 1968.[140] By breaking up the larger class, the court reduces the number of individuals on the right hand side of the table. In an Alabama case, where all state registrars of voters were made into a defendant class, the court administered its ruling on the basis of different sub-classes.[141] The "harm" in this Alabama case was to convicted felons who were thrown off the voter rolls.[142] The court found that some counties had done more harm than others, and appropriately tailored their remedy.[143] The same logic can be applied by courts in other defendant class action contexts.

The second option courts can undertake is to create incentives for self-identification by adjusting presumptions on marginal contribution to harm, and then allowing for rebuttal of that presumption with sufficient evidence. To flesh this out, it is helpful to consider the numerical example in Table 5. Let's assume that a plaintiff has experienced total harm of 500, and has won in court. The defendant class is composed of 100 individuals, and neither the plaintiff nor the court knows which defendants produced what amount of harm. Each individual defendant does not know the other defendant's contribution to harm, but he knows his own. He knows how many people are in the class, so he knows that the average harm is 5. Let's say that the distribution is as presented in Table 5.

140. *See* Technograph Printed Circuits, Ltd. v. Methode Elecs., Inc., 285 F. Supp. 714, 725–27 (N.D. Ill. 1968) (finding a defendant class appropriate in a patent infringement case, and in administering it, creating several sub-classes).

141. Hobson v. Pow, 434 F. Supp. 362, 365 (N.D. Ala. 1977).

142. *See id.*

143. *See id.* at 367–68.

Table 5. Hypothetical: Contributions to Harm								
[A] Individual contribution to harm:	2	3	4	5	6	8	10	*TOTAL*
[B] Number of individuals:	5	25	25	15	10	10	10	*100*
[A] x [B] = [C] Sub-Total of Harm:	10	75	100	75	60	80	100	*500*

The puzzle is this: if we cannot directly observe their contributions to harm, how do we get the various sub-groups to volunteer, *ex post*, information about their contributions to harm? Courts may be able to use damage assignments as a carrot-and-stick. In this example, instead of setting the average damage payment for defendants at 5, courts could set it at 8. Initially such a move would strike off over-deterrence because total damages would equal 800. But courts could, at the same time they set damages to 8, offer defendant class members a chance to reduce their liability to 6 if they can show that they contributed less than 8 to harm. In this example, 80 people would rationally come forward to get their liability reduced by 2. The 10 in the "8 category" would break even, and the 10 in the "10" category would get away with 2. Total damages would thus be 800 – 160 = 640. There is likely some over-deterrence here, but it should be noted that the over-deterrence is the cost of identification. Over time, courts could calibrate their carrot-and-stick game.

A third option, which is probably quite costly and therefore not as practical, is for the court to appoint a guardian or special master specifically for the purpose of determining marginal contributions to risk. Guardians have been a frequent topic of discussion in the class action context.[144] Here, "special master" may be a more appropriate title, but the person charged with the responsibility of looking at contributions to harm will also likely be faced with questions of settlement and infighting as well.

In addition to these mechanisms, courts must also recognize that their choice of representative can affect information production.[145] A

144. *See* Alon Klement, *Who Should Guard the Guardians? A New Approach for Monitoring Class Action Lawyers*, 21 REV. LITIG. 25, 26–28 (2002) (discussing private lawyers' roles in the "guardian" class-action context and who should supervise those lawyers); *see also* Edward Brunet, *Class Action Objectors: Extortionist Free Riders or Fairness Guarantors*, 2003 U. CHI. LEGAL F. 403, 466 (2003) ("The theory of appointing a guardian ad litem is deceptively simple. The guardian will represent the interests of the absent class members and thereby monitor the behavior of class and defense counsel during settlement negotiations.").

145. Courts have long recognized the problem of defendant class representation, but not usually through the lens of information production. In *In re the Gap Stores Sec. Litig.*, the court noted as follows:

common problem with large defendant classes is how many and which defendants should be assigned as class representatives. Courts have encountered what I call the "red rover" problem; just as in the game red rover, where kids seek to run through the weakest link on the other side, plaintiffs want to pick the weakest link as the defendant class representative. Such was the case in *Leon N. Weiner & Assocs., Inc. v. Krapf*,[146] where a corporate lot owner sought to name just one lot owner as representative of a class of 203 lot owners.[147] The corporate owner sought declaratory judgment that its property was not subject to restrictions, and the alleged "harm" the corporate owner experienced was the potential restrictions on the land as carried by other lot owners.[148]

Faced with this situation, the court recognized that the single named defendant was not in a position to adequately produce information on the many possible restrictions that might arise from the deeds of other lot owners.[149] The court concluded that the plaintiff corporate owner "selected one neighbor to represent the property interests of 203 lot owners, many of whom will likely have different interests and views. The effect of Weiner's motion is to place the costs of notice, discovery and litigation on the shoulders of the Krapfs."[150] Such costs would make it virtually impossible for the defendant representative to engage his peers and kick-start the information market.[151]

C. Enforcement

Even in some situations in which identification and monitoring are practical, enforcement may not be. Here, "enforcement" means getting other defendant class members to contribute to (i) the litigation costs, and then (ii) if necessary, the damage costs as well. Enforcement is difficult because without a well-working network between defendants (i.e., without an umbrella organization), no single class member (or even a small pocket of class members who may be strongly connected) can achieve

---

> Commentators have frequently criticized the potential for inadequate representation of defendant classes. Because the named defendant generally does not seek his representative status and often vehemently opposes it, a court may fear that an unwilling representative will necessarily be a poor one. Related to this concern is the fear that the plaintiff will exercise his power of selection to appoint a weak, ineffective opponent as class representative. "It is a strange situation where one side picks out the generals for the enemy's army."

79 F.R.D. 283, 290 (N.D. Cal. 1978) (citations omitted).

146. No. 8938, 1988 Del. Ch. LEXIS 8 (Del. Ch. Jan. 19, 1988).

147. *Id.* at *1.

148. *Id.* at *4.

149. *Id.* at *8–9.

150. *Id.* at *9.

151. A similar analysis was made in a case in Illinois where a single owner of a Shell gas station was proposed as representative of a class of all Shell gas owners in the state. In rejecting this proposed representative, the court reasoned, "The entire economic burden of defending the present suit was thrust upon one man, Razowsky. His financial stake in the outcome of the suit was not shown to be greater than that of any other of the hundreds of Shell dealers in Illinois." Gaffney v. Shell Oil Co., 312 N.E.2d 753, 759 (Ill. App. Ct. 1974).

the economies of scale required to effectively enforce group-wide policies. In the face of such an enforcement problem, individuals *ex ante* would look to the legal system to provide mechanisms for making enforcement feasible.

On this enforcement point Hamdani and Klement offer a useful analysis with their class defense proposal. They propose a "class defense" mechanism, a "defendant-initiated procedure designed to create parity between a single plaintiff and a group of similarly positioned defendants."[152] With help from the court (via fee-shifting), a defendant could use Hamdani and Klement's class defense procedure to reach out and essentially force contributions from the entire class.[153] Hamdani and Klement's proposal allows defendants to class themselves with less judicial intervention than would currently be required. Legal tools that make it easier for aggregation of claims promote the aggregate analysis principle, and thus the forward-looking deterrence principle as well.

Hamdani and Klement's proposal runs into trouble, however, when we reach the hard hypothetical case of the blockbuster movie DVDs. Suppose that through some investigation, the movie's production studio is able to identify 100 of the 100,000 people who illegally taped the movie. Suppose too that the movie studio then asks for certification of a defendant class for all illegal recorders (which they have estimated at 100,000 based on lost revenue from movie tickets and DVD sales). The problem at this point is that even if the defendants "class" themselves, no single defendant is in a position to serve as a representative for the entire class. Even when the hundred identified defendants put their resources together, it is not going to scale up enough to match the movie studio's legal resources. This is a problem because the issues may not be fully litigated. For instance, perhaps the movie studio made some contribution to the harm, which would not come out unless the defendant class had better representation.

To deal with this hard case, it is helpful to recall that a forward-looking court hopes to minimize similar harms like these from arising in the future.[154] In order to arrive at optimal deterrence, we need to conduct aggregate analysis. In this hard case, the only way to achieve fully litigated aggregate analysis would be for the court to incur tremendous costs and essentially fund a legal team for the defendant class. The great majority of the defendant class remains anonymous, and thus would not contribute to a pool to fund the legal fees. Given these prohibitive costs, and the requirement of aggregate analysis which fails in this hard case,

152. Hamdani & Klement, *supra* note 9, at 710.

153. The procedure assumes that identification and monitoring are possible. If a defendant has no idea who else is in his class, he will not obtain maximum benefits from classing himself.

154. "Fairness" to this particular group of movie tapers or to the production company is not, in the general welfare framework of this article, at issue.

the general theory of this Article suggests that here, defendant class actions will not be an optimal legal tool.[155]

While defendant class actions are not optimal in these hard cases, it should be emphasized that such cases are very rare in practice. The Judicial Panel on Multidistrict Litigation, for instance, usually seems to find a few "big players" or some other market mechanisms by which to identify representatives for the diverse parties involved in litigation. The actual cases where defendant class actions have been certified also point to consistent findings of links between defendants.[156] More generally, it is difficult to find frequently occurring instances in which there are no market relationships between defendants in a potential defendant class.

D. Free Riding

The free riding problem is the result of identification, monitoring, and enforcement failures. Examining the defendant class action, there are two types of free rider problems we need to consider. The first is most analogous to the hypothetical case already discussed in Section I. It is where a class-wide defense would be beneficial to all defendants, but no single defendant can fund the defense adequately because they cannot extract payments from the free-riders in their class. As Netto has pointed out, "Only economy of scale in investment in the lawsuit can overcome the problem of the reluctance of defendants to assume the litigation as class representative. This objective is achieved with incentives for the class counsel through an optimal mechanism of compensation for his performance."[157] In these cases, an individual *ex ante* would desire that the legal system provide a means by which the defense can be properly funded. Individual defendants would desire a mast-tying device.

But a second sort of free-riding problem may also exist. This second type of free-riding problem arises when the defendants would not be better off if they all stopped causing the harm. Rather, it is society that would be better off because the utility that the defendants are deriving

155. Although they arrived at the conclusion by different means, the court in *Angel Music, Inc. v. ABC Sports, Inc.*, a copyright case with a large potential class of copyright infringers, denied class certification. 112 F.R.D. 70, 71 (S.D.N.Y. 1986). The *Angel* court based its ruling on the lack of connections between defendants, and issues of standing. *Id.* at 77. The plaintiff, Angel Music, argued that "the members of the defendant class have engaged in a common violation of the Copyright Act which places their actions within the juridical link exception to *LaMar*," but the court recognized that no such relationship existed. *Id.* at 75–76. What the court could have also said was that when confronted with this hard case of copying infringement, a class device was not likely to create links between future defendants in similar situations.

156. In a defendant class action brought under the Sherman anti-trust laws, for instance, Columbia Broadcasting System, Inc. (CBS) targeted the American Society of Composers, Authors and Publishers (ASCAP) and Broadcast Music, Inc. (BMI) in order to get at the numerous defendant musicians and performers in the proposed class. Broad. Music, Inc., v. Columbia Broad. Sys., Inc., 441 U.S. 1, 4 (1979).

157. Netto, *supra* note 4, at 93.

from their harmful behavior is not equal to the dis-utility they cause others (the plaintiff).

The *ex ante* perspective is crucial for understanding this second free riding problem. Once an individual knows if he will be in the defendant class, it is no longer in the individual's interest to maximize utility over both states (plaintiff and defendant) of the world. To see how these interests can diverge once we move out of the *ex ante* world, consider this numerical example.

One hundred defendants each cause 5 in harm to the plaintiff when they steal a brick, for a total of 500 harm. They derive only 3 in utility from each harm, by selling the brick, for a total of 300 utility. Plaintiff cannot identify every member of the class, but when they are identified, plaintiff knows that the marginal contribution to harm is 5. Overall, we want to deter the defendants if we can do it for less than 200. Let's say plaintiff can find 20 of the wrongdoers, and every time wins against them for their marginal contribution, 5. When we look overall at defendant and plaintiff (Table 6), we see that optimal deterrence is not achieved because the defendants wind up 200 better off, while the plaintiff is 400 worse off. The harm producers do not bear the loss.

Table 6. Social outcomes *without* a Certified Defendant Class

	Defendant Class	**Plaintiff**	**Society Overall**
Initial gain / loss	100 people gaining 3 utility each from their wrong-doing = + **300**	100 wrong-doers each causing plaintiff 5 in: **-500**	**-200**
Subsequent legal gain / loss	20 people get sued and lose 5 each: **-100**	Successfully suing 20 people for 5 gain each: **+100**	No change
Final Result	**+200**	**-400**	**-200 with harm producers NOT bearing the loss, so sub-optimal deterrence**

Table 7. Social outcomes *with* a Certified Defendant Class			
	Defendant Class	**Plaintiff**	**Society Overall**
Initial gain / loss	100 people gaining 3 utility each from their wrong-doing = + **300**	100 wrong-doers causing 5 in harm a piece: - **500**	-200
Subsequent legal gain / loss	20 people get sued, and get certified as Defendant Class, so must pay the entire harm, lose 500 total: **-500**	Successfully suing the defendant class: **+500**	No change
Final Result	**-200**	**0**	**-200 with harm producers bearing the loss, so optimal deterrence**

Now consider how a defendant class action would change the final results (Table 7). If the 20 defendants were certified as a defendant class, they would be liable not only for their marginal contribution (the 100), but for the entire 500 in harm. This would benefit society overall because it would create the proper deterrent effect, but it would not benefit the defendant class. Thus, one's desire for a defendant class would depend on whether one knows if they will be in the class or not.

Courts encountering this issue—making some defendants liable for the harms of the entire class—have been wary of pushing forward. In *In re the Gap Stores Securities Litigation*,[158] the U.S. District Court for the Northern District of California suggested:

> [A] defendant class action may be simply an inappropriate method of adjudicating any case where the combination of punitive damages and joint and several liability threaten to transform a statutory scheme for personal accountability into ready martyrdom for the unlucky defendant whose deep pocket will pay for the sins of the multitude.[159]

158. 79 F.R.D. 283 (N.D. Cal. 1978).

159. *Id.* at 295. A New Jersey court echoed a similar sentiment in a defendant class action case: [I]t is noted that the New Jersey Antitrust Act, under which relief is requested, contemplates joint and several liability. The accumulated damages, trebled pursuant to statute,

The court's focus solely on the "unlucky defendant" is misplaced, for there is also an "unlucky" plaintiff who has experienced harm. The court should look to the good of both plaintiff and defendant, using aggregate analysis to consider the overall social welfare implications of its legal rule.

My argument for aggregate analysis is distinct from Hamdani and Klement's approach. When they propose the "class defense" mechanism, they fail to recognize that whether it is a plaintiff who wants to certify a defendant class or defendants who want to certify themselves, our evaluation of the merits of that class certification should rest upon the determination of overall benefit to society. If one's primary social objective is maximizing overall utility, then focusing solely on maximizing plaintiffs' or defendants' utility is misguided.

1. Solving the Free Riding Problem with Fee Shifting

The free rider problem is one of the most difficult challenges to overcome in successfully carrying out a defendant class action. The problem, however, has been addressed through various fee-shifting proposals. Most on point is Netto's proposed solution, drawing on the English rule for attorney fees:

> Defendant-favoring fee shifting is considered fee-shifting on a one-way (or one-side) basis, granting fees only to the defendant's attorney when the defendants prevail in the lawsuit, but not awarding fees to the plaintiff's lawyer even if he wins the case. . . .
>
> . . . .
>
> . . . The advantages of the defendant-favoring fee-shifting system include: (i) overcoming the asymmetric costs between separate litigation and collective suit, aggregating the multitude of defendants, (ii) compensating the class counsel by equalizing his investment in the litigation with the amount of the fees award; and (iii) precluding nuisance value suits.[160]

The fee-shifting literature also provides other solutions relevant to defendant class actions. Particularly useful is Joseph Miller's work on the free rider problem faced by those who challenge the validity of a

---

> recoverable by the entire class of mortgagors from the entire class of mortgagees, may aggregate many millions of dollars. Yet, if the class recovery were allowed, each member of defendant class, no matter how minor its participation in the scheme, would be individually answerable for the full amount of the judgment. We conclude that such a result would constitute a major alteration in the substantive legal relations between the parties and goes beyond the intent of class action policy.

Kronisch v. Howard Sav. Inst., 335 A.2d 587, 598 (N.J. Super. Ct. Ch. Div. 1975) (citing 3B JAMES WM. MOORE ET AL., MOORE'S FEDERAL PRACTICE ¶ 23.45[3] (2d ed. 1996)).

160. Netto, *supra* note 4, at 114–16.

patent.[161] Miller begins his analysis by stating, "A court judgment that a patent claim is invalid is a public good. And obtaining such a judgment requires the expensive, up-front cost of patent litigation. These facts suggest that profit-maximizing firms will supply definitive patent challenges at a less-than-optimal rate."[162] A similar fact pattern is found in our context. When a defendant class is engaged in activity that (as determined by aggregate analysis) is good for society, that is a public good. Fully litigating such a stance and showing that you are engaged in a public good also involves substantial costs. Just as invalidating a patent invalidates it for everyone, so winning the right to continue engaging in your actions (i.e., file-swapping) allows everyone else to do the same. Given these similarities, what can we learn from Miller's analysis? The first lesson is one about theoretical approach. Miller sets the stage in this way:

> Any bounty mechanism—in the patent context or elsewhere—depends for its success upon when the bounty is awarded (or, put another way, what one must do to earn it), and of what the bounty consists (e.g., cash payment of $X, or enough money to cover expense Y). A poor choice as to either feature reduces a bounty's effectiveness at encouraging the desired result, making these features the best focus in assessing whether a proposed bounty is likely to succeed.[163]

Miller's analysis, not detailed further here, considers two existing bounty and fee-shifting proposals in the patent context.[164] Like Miller, I believe that, "[p]aying a successful patent challenger a cash bounty that need not be shared with others who benefit from the patent's invalidation directly counteracts the free rider problem . . . ."[165] The questions then become: (i) *When* should the bounty be awarded, and (ii) *How much* should the bounty be?

For defendant class actions, the timing question is somewhat easier than the parallel question in patent law.[166] The bounty should be awarded at the litigation stage. A litigation stage bounty should be awarded to those defendants who step up to defend on behalf of the entire class. If too many lawyers step forward, the court can adjudicate between them, either on the merits or via a lottery. The timing of this bounty would encourage full litigation of the issues. To pay for the bounty, the court could mix-and-match between (i) fee shifting provisions in the event of a win by the defense, (ii) a mandatory "litigation tax" imposed on all members of the defendant class, and (iii) a sliding "litigation investment" in which defendant class members could contribute to the class defense,

---

161. Joseph Scott Miller, *Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents*, 19 BERKELEY TECH. L.J. 667 (2004).

162. *Id.* at 688 (footnote omitted).

163. *Id.* at 695–96 (footnote omitted).

164. *Id.* at 696–704.

165. *Id.* at 704.

166. The reason for this is that the challengers to patent infringers must initiate the lawsuits.

with a promise that they would receive their investment plus a percentage of the second-stage bounty. The amount of the bounty is something that courts would have to determine based on the size of the class and the issues involved.

Legislatures can be a partner in establishing and revising fee-shifting programs. In Colorado in 1990, for instance, the company Terrestrial Systems sought to bring a class action suit against a class of television owners that they alleged were using certain equipment to gain unauthorized access to broadcasts.[167] Fee-shifting in the case was guided by legislative mandate.[168] Under Colorado Revised Statutes § 18-4-702(3), "in any action for civil theft of cable television service the prevailing party shall be entitled to an award for his reasonable attorney fees."[169] The case illustrates the possibilities of fee-shifting to be designed for specific situations and to be put into practice. Legislatures thinking about social good for the state or country can produce background fee-shifting rules that address the free-rider concerns inherent in defendant class actions.

All of these funding options still leave open the possibility that lead defense lawyers might be quick to settle, or might work out a sweetheart settlement for themselves. Because they might be representing defendants who are not even known to the plaintiffs (thinking back to the identification problems), there seems a distinct possibility that whatever the bounty or fee shifting regime, settlement incentives will remain askew. To counter this, I propose making representation of defendant classes a *repeat game* by looking favorably upon legal defense teams that have successfully litigated in the past, and looking unfavorably upon those who have lost, and especially unfavorably at those who have struck deals that seemed to be of the sweetheart variety. Such repeat games are similar in spirit to proposals to use repeated auctions for informational purposes.[170] If law firms in these cases are one-time players, then this solution will do little. But in a world of consolidated firms, I suspect that we would see many repeat players. Because they are now maximizing revenue not just in this particular case, but across all future cases, firms will be less likely to engage in behavior that is not in keeping with the class as a whole.

2. Solving the Free Riding Problem with Command-and-Control

If all else fails, full-blown government regulation in the form of command-and-control may be necessary. This approach is likely to be incredibly expensive. Terry Fisher has proposed such an approach for

167. Terrestrial Sys., Inc. v. Fenstemaker, 132 F.R.D. 71, 73 (D. Colo. 1990).
168. *Id.* at 73–74.
169. *Id.* at 73 (citing Colo. Rev. Stat. § 18-4-702(3) (2004)).
170. *See* Abramowicz, *supra* note 100, at 351–52, 378–79.

copyright.[171] In his proposal, Fisher suggests that ISPs pay royalties, based on the level of downloads of particular pieces, into a government-run fund, which would then disperse those royalties to individual artists.[172]

It should be observed that many potential defendant class actions have already been addressed by government regulation. The "tragedy of the commons" cases preempt class action lawsuits by using regulatory agencies (fines, taxes, etc.) to deter socially detrimental conduct such as littering. The government may be in the best position to identify, monitor, and deter the risk-creation of the large number of defendants. Where the legislature has not already stepped in, however, courts may be more hesitant to push for such regulation.[173]

E. Liability rules

While the legal rule may vary in some situations, the default rule should be strict liability for the defendant class, with contributory negligence. Strict liability would have the benefit of eliminating in-fighting within the defendant class. For instance, none of the members of the brick re-sellers association could show that they had *not* stolen bricks from this particular factory. This should theoretically create very strong self-monitoring and self-policing incentives. The logic is that if you do something illegal, we all pay for it, so we're going to try and make sure that you do not do anything illegal. Or, perhaps more realistically, we are going to take more care and screen our members to make sure that we reduce our risk.

The tool of vicarious liability could also be used to bring in an existing organization that has been standing on the sidelines or to generate the creation of a new organization that no one had the incentive to start yet.

---

171. WILLIAM W. FISHER III, PROMISES TO KEEP: TECHNOLOGY, LAW, AND THE FUTURE OF ENTERTAINMENT (2004).

172. *Id.* at 202–03. Fisher provides an overview of his compensation alternative and then uses the rest of the chapter to elaborate in detail how such a system would operate.

173. A moderate, and potentially more cost-effective path for legislatures to take is to mandate defendant class actions in certain circumstances. An illustration is a Missouri law that required certain annexation proceedings to proceed via a class action. MO. REV. STAT. § 71.015(1)(5)(c) (2010). In *City of St. Ann v. Buschard*, the court explained this law as follows:

> [The] Sawyer Act passed by the 67th General Assembly . . . provides that before a city may proceed to annex any area otherwise authorized by law, it must file an action in the Circuit Court of the County in which such unincorporated area is situated praying for a declaratory judgment authorizing such annexation. According to the Sawyer Act: "The petition in such action shall state facts showing:
>
> 1. The area to be annexed;
>
> 2. That such annexation is reasonable and necessary to the proper development of said city; and
>
> 3. The ability of said city to furnish normal municipal services of said city to said unincorporated area within a reasonable time after said annexation is to become effective. Such action shall be a class action against the inhabitants of such unincorporated area under the provisions of Section 507.070 RS Mo."

299 S.W.2d 546, 547 (Mo. Ct. App. 1957) (citation omitted).

In thinking at the aggregate level about deterrence—even though it might be too late for the particular group on trial to create an organization that could have better protected their interests—future groups in similar situations will look to this court's ruling and realize that the threat of individual liability is so great that they are not going to even enter the market (i.e., not going to take a single brick) unless they are sure that there is some sort of organization/agency/binding agreement that they can become a party to.

Allowing for contributory negligence makes sure that plaintiffs do not get off the hook. It might be the case, for instance, that recording artists made their work too easy to illegally obtain. Contributory negligence could be assessed to the extent that a firm is not up to the state-of-the-art with certain technological precautions.

IV. CONCLUSION

This Article has synthesized existing knowledge about defendant class actions and proposed a general theory of defendant class actions. The argument of the paper rests on three principles: (1) forward looking deterrence; (2) dynamic effects; and (3) aggregate analysis. Of these three, it is the aggregate analysis principle that overshadows the other two in importance. The Article provides some illustrations of these principles, and sketches out some ways in which these principles can be applied in system design. The proposals made in this Article challenge courts and legislatures to broaden the scope of their legal reasoning beyond purely formalist concerns about the language of Rule 23.

There is much more to be considered in the defendant class action context. It remains to be seen, for instance, how the proposed tools of system design will hold up in practice. Because of the aggregate analysis principle, more work needs to be done on bringing in additional data and perspectives on the substantive issues at hand.

Despite these unanswered questions, it is my hope that this Article has contributed to the literature by calling for scholars to frame their discussion of defendant class actions within a broader theoretical framework. What is it that one wants a defendant class action to do? Which parties should we think about when adjudicating defendant class actions? How much marginal value do we expect defendant class actions to have in particular situations? Continuing to answer these questions in more detail will enable courts to feel more confident in their ability to certify defendant classes. That, ultimately, will lead to greater social welfare.

APPENDIX

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Broad. Music Inc. v. Columbia Broad. Sys., Inc.	441 U.S. 1	1979	CBS	American Society of Composers, Authors and Publishers (ASCAP) and Broadcast Music, Inc. (BMI), and their members and affiliates	Decided in favor of defendant class on basis of Sherman Act
Zablocki v. Redhail	434 U.S. 374	1978	All Wisconsin residents who have minor issue not in their custody and who are under an obligation to support such minor issue by any court order or judgment and to whom the county clerk has	All county clerks in Wisconsin	Civil rights suit, statute struck down on equal protection grounds

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			refused to issue a marriage license without a court order		
Trainor v. Hernandez	431 U.S. 434	1977	Those "who have had or may have their property attached without notice or hearing upon the creditor's mere allegation of fraudulent conduct pursuant to the Illinois Attachment Act."	Various Illinois state officials	Civil rights suit, the Court held that the case ought to have been dismissed for other reasons and the question of constitutionality never addressed by the district court
In re Integra Realty Res., Inc.	354 F.3d 1246 (10th Cir.)	2004	Creditors' trust	Seven large shareholders in bankrupt Integra company (instead of 6,000 individual shareholders)	The case was reviewed on appeal, and the district court settlement agreement was upheld.
Tilley v. TJX Cos.	345 F.3d 34 (1st Cir.)	2003	Graphic artist	557 retailers who sold artist's work	Appeals Court overturned District Court's

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
					certification of defendant class on the grounds that the court erred in certifying a defendant class under Rule 23(b)(2)
S. Ute Indian Tribe v. Amoco Prod. Co.	2 F.3d 1023 (10th Cir.)	1993	Indian tribe	Oil companies, lessees, and well operators.	Dispute concerning the ownership of coalbed methane
Henson v. E. Lincoln Twp.	814 F.2d 410 (7th Cir.)	1987	People denied due process when applying for welfare	770 Illinois local welfare departments not receiving state aid	Certification of defendant class not allowed under 23(b)(2)
AFP Imaging Corp. v. Ross	780 F.2d 202 (2d Cir.)	1985	AFP Imaging Corporation	Twenty-nine shareholders of Xenon	Securities fraud case under 10(b) of the Securities Exchange Act of 1934
Baker v. Wade	769 F.2d 289 (5th Cir.)	1985	Donald F. Baker, a homosexual	"[A]ll district, county and city attorneys in the State of Texas responsible for the	Challenged constitutionality of TX law that proscribes "engag[ing] in deviate

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				enforcement of Texas Penal Code Ann. § 21.06 . . . ."	sexual intercourse with another individual of the same sex" as unlawful
Thompson v. Bd. of Educ. of the Romeo Cmty. Sch.	709 F.2d 1200 (6th Cir.)	1983	"All female teachers of such school boards who have been since March 24, 1972 or will be in the future, denied the benefits of a sick leave policy which treats pregnancy related disabilities the same as other temporary disabilities."	"All school boards in the State of Michigan which, since March 24, 1972, have treated or now treat pregnancy related disabilities differently than other temporary disabilities, limited to the school boards in districts wherein the MEA has female members who have been or will be subject to such policies or practices."	District Court certified, but Appeals Court reversed

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Kerney v. Fort Griffin Fandangle Ass'n	624 F.2d 717 (5th Cir.)	1980	Individual injured during a theatrical performance	Officers and members of unincorporated theatrical association	Suing individuals and the class for injuries
Greenhouse v. Greco	617 F.2d 408 (5th Cir.)	1980	Black children attending parochial schools in defendant diocese	Parish corporations holding title to parochial schools in the diocese; Bishop of the diocese, the superintendant of diocese schools, and the diocese itself	Appeals Court affirmed the District Court's ruling that the lack of appropriate representatives precludes the suit from moving forward as a defendant class action
Marcera v. Chinlund	595 F.2d 1231 (2d Cir.)	1979	Pre-trial detainees in county jails	42 county sheriffs who denied contact visits in their jails	Detainees wanted more contact visits allowed
Brooks v. Flagg Bros.	553 F.2d 764 (2d Cir.)	1977	Individuals who had storage in defendant's warehouse	"[W]arehousemen doing business in the State of New York and who impose liens and subject goods to sale pur-	Due process related to liens on stored materials

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				suant to New York Uniform Commercial Code ss 209-210 (sic) without affording the owner of the goods a prior opportunity to be heard."	
Anastasia v. Cosmopolitan Nat'l Bank of Chi.	527 F.2d 150 (7th Cir.)	1975	"Those persons in Chicago, Illinois, except for the owners, managers and operators of hotels, whose personal property is now detained by a hotel pursuant to the Illinois Innkeepers' Lien Law."	"Those owners, managers, and operators of hotels in Chicago, Illinois, who now have the personal property of the class of plaintiffs detained pursuant to the Illinois Innkeepers' Lien Law."	Plaintiffs challenged the seizures as violations of both the Fourteenth and Fourth Amendments.
Appleton Elec. Co. v. Advance-United Expressways	494 F.2d 126 (7th Cir.)	1974	"Appleton Electric Company and all other persons similarly situated who . . . have	1,400 carriers who "shipped goods in interstate commerce between May 20,	Shippers wanted refunds from carriers under Interstate Commerce Commis-

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			shipped goods through interstate commerce via the facilities of the defendants . . ."	1968 through August 29, 1969 within a region covered by tariff rates involved in Interstate Commerce Commission Docket No. 34971"	sion rules
Brown v. Kelly	244 F.R.D. 222 (S.D.N.Y. 2007), *aff'd in part, vacated in part,* 609 F.3d 467 (2d Cir. 2010)	2007	All persons arrested, charged or prosecuted for a violation of loitering for the purpose of begging in the State of New York from October 7, 1992 onward	"[A]ll political subdivisions and all law enforcement/prosecutorial policy-making officials in the State of New York with authority to arrest, charge or prosecute a person with a violation under New York Penal Law."	Case on 1st, 4th, and 14th amendment grounds; bilateral statewide classes and city subclass certified
Baksalary v. Smith	No. Civ.A. 76-429, 2005	2005	Plaintiff employees,	"[A]ll insurance carriers	Court granted the insur-

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
	WL 1941319 (E.D. Pa.)		"whose workers' compensation benefits had been terminated under the § 413(a) automatic supersedeas provision"	and self-insured employers who had invoked, or would in the future invoke, the automatic supersedeas procedure of section 413(a)."	ers' motion to vacate the consent decree.
Pension Transfer Corp. v. Benef. Under Third Amend. to Fruehauf Trailer Corp. Ret. Fund No. 003	319 B.R. 76 (D. Del.)	2005	Pension Transfer Corporation	All the beneficiaries of an amendment to a bankruptcy by Fruehauf Trailer Corporation	Bankruptcy; alleged that amendment to the plan was a fraudulent transfer
Matte v. Sunshine Mobile Homes, Inc.	270 F. Supp. 2d 805 (W.D. La.)	2003	Owners of mobile homes alleged to be inherently defective	282 manufactured home builders	Certification not allowed; plaintiff claims dismissed
Wyandotte Nation v. City of Kansas City	214 F.R.D. 656 (D. Kan.)	2003	Indian tribe	All individuals and entities ... who claimed an interest in any portion of those sections of land at	Seeking declaratory judgment, recovery of possession of real property

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				issue in Wyandotte County, Kansas	
Clark v. McDonald's Corp.	213 F.R.D. 198 (D.N.J.)	2003	Disabled individuals who were "denied the full and equal enjoyment of the goods, services, programs, facilities, privileges, advantages, or accommodations of any of the McDonald's"	"[A]ll owner/operators of McDonald's brand restaurants throughout the United States."	Certification not allowed
Mayo v. Hartford Life Ins. Co.	193 F. Supp 2d 927 (S.D. Tex.), *aff'd,* 354 F.3d 400 (5th Cir. 2004)	2002	Employees (and former employees) living in Texas who are (or were) insured under the COLI policies owned by any of the defendant employers	"Companies that bought insurance policies written by AIG Life Insurance Company, Mutual Benefit Life Insurance Company or Hartford Life Insurance Company, that insure or insured	Argument was that company-owned life insurance policy in the employee's name violated the Texas insurable interest doctrine

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				the lives of Texas employees."	
Silva v. County of Los Angeles	215 F. Supp. 2d 1079 (C.D. Cal.)	2002	Attorney	All Los Angeles Superior Court judges, court commissioners and California Court of Appeal justices	Dismissed; alleged that the County's payment of benefits to judges and the judges' failure to disclose such payment deprived Silva of his rights to due process, equal protection
Monument Builders of Pa., Inc. v. Am. Cemetery Ass'n	206 F.R.D. 113 (E.D. Pa.)	2002	Monument Builders of Pennsylvania, Inc., a trade association	All cemeteries and cemetery associations throughout the Commonwealth of Pennsylvania	Alleged that cemeteries inflated the prices of monuments
Canadian St. Regis Band of Mohawk Indians v. New York	146 F. Supp. 2d 170 (N.D.N.Y.)	2001	Descendants of the Village of St. Regis	New York State defendants, St. Lawrence and Franklin Counties, Village of Massena, Town of	Dismissed (Land claim case for 200 years of dispossession)

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				Massena, Town of Bombay, Town and Village of Fort Covington, Key Bank of Northern New York, the Nationwide Mutual Insurance Company, Niagara Mohawk Power Corporation and Canadian National Railways, individually and on behalf of all other persons who claimed an interest in any portion of the subject lands	
Oneida Indian Nation of New York State v. County of Oneida	199 F.R.D. 61 (N.D.N.Y.)	2000	Oneida Indian Nation	Proposed a class of "approximately 20,000 or more persons" who "occupy or have or claim an	Class not certified; amendment to add 20,000 or more individual defendants

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				interest in any of the subject lands . . . and their successors and assigns"	denied (Land claim case for 200 years of dispossession)
E.R. Squibb & Sons, Inc. v. Accident & Cas. Ins. Co.	82 Civ. 7327 (JSM), 1999 U.S. Dist. LEXIS 8333 (S.D.N.Y.), *aff'd in part, rev'd in part sub nom.* E.R. Squibb & Sons, Inc. v. Lloyd's & Cos., 241 F.3d 154 (2d Cir. 2001)	1999	Squibb	Lloyd's Underwriters and Lloyd's Underwriting Syndicate # 210 (and the thousands of individuals underwriting)	Insurance case, seeking declaratory judgment; Appeals court recommended defendant class action, saw a 23.2 unincorporated association
Monaco v. Stone	187 F.R.D. 50 (E.D.N.Y.)	1999	Gregory Monaco	All local criminal court judges; "all sheriffs or other individuals who transport incompetent defendants from jail to OMH psychiatric institutions"	Constitutional challenge of NY law under which defendants found incompetent to stand trial for minor felonies and misdemeanors in New York State are

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
					involuntarily committed to state-operated psychiatric hospitals.
United Cos. Lending Corp. v. Sargeant	20 F. Supp. 2d 192 (D. Mass.)	1998	United Companies Lending Corporation	Defaulting borrowers	Sought declaratory judgment that a Mass. law was void and unenforceable
K. Bell & Assocs. v. Lloyd's Underwriters	92 Civ. 5249 (AJP)(KTD), 1998 U.S. Dist. LEXIS 7798 (S.D.N.Y.)	1998	K. Bell & Associates	"[C]onsortium of individual investors, known as 'Names,' that are severally, but not jointly, liable for their fraction of the risk on [Lloyd's] insurance policy."	Dismissed for lack of subject matter jurisdiction
Leer v. Wash. Educ. Ass'n	172 F.R.D. 439 (W.D. Wash.)	1997	All non-member public school district employees who were or are "represented exclusively for purposes	"[A]ll UniServ Councils and local associations affiliated with Defendant WEA . . . which collect agency	Court found certification inappropriate under Fed. R. Civ. P. 23(b)(1)(A) because there was "no contract,

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			of collective bargaining by Defendants and were subject to demands for or collections of agency fees for the WEA and any of its affiliates."	fees from nonmembers." (at least 100 such local affiliates)	agreement, or enforced system between" the defendants that would support a finding of a juridical relationship.
Coal. for Econ. Equity v. Wilson	122 F.3d 692 (9th Cir.)	1997	The Coalition For Economic Equity, et. al.	All state officials, local government entities or other governmental instrumentalities bound by Proposition 209.	Challenging the constitutionality of Proposition 209.
Capital Cities/ABC, Inc. v. Ratcliff	No. 94-2488-GTV, 1996 U.S. Dist. LEXIS 14798 (D. Kan.), *aff'd*, 141 F.3d 1405 (10th Cir. 1998)	1996	"All persons who have been home delivery or single copy agents for The Kansas City Star Company pursuant to an agency agreement since the delivery agent	"All persons engaged . . . in the delivery in the states of Kansas and/or Missouri of The Kansas City Star and/or The Kansas City Times pursuant to an	Underlying question: whether carriers were entitled to employee benefits under the Employee Retirement Income Security Act of 1974 (ERISA),

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			system implemented by The Star" after November 30, 1984, and on or prior to November 30, 1994.	agency agreement with the Kansas City Star Company that provides or provided that the person is a self-employed independent contractor and not an employee or servant of The Star."	
Nat'l Union Fire Ins. Co. v. Midland Bancor, Inc.	158 F.R.D. 681 (D. Kan.)	1994	National Union Fire Insurance Company of Pittsburgh, Pa.	Class I: "all persons who are or ever were directors or officers of the Institutions, who have or may in the future make claims which might fall within the scope of coverage of the Policy." Class II: "all person or entities who have	Defendant class not certified; wanted declaratory judgment on liability policy

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				made or may in the future make claims against the directors and officers of the Institutions which might fall within the scope of coverage of the Policy."	
United States v. Local 1804-1, Int'l Longshoremen's Ass'n.	90 Civ. 0963 (LBS), 1993 U.S. Dist. LEXIS 15083 (S.D.N.Y.), *aff'd in part, vacated in part*, 44 F.3d 1091 (2d Cir. 1995)	1993	United States	Waterfront Employer Class (associations and 18 companies which specialize primarily in the maintenance and repair of marine containers and chassis and securing and unsecuring of cargo aboard ships)	RICO action
Amnesty Am. v. County of Allegheny	822 F. Supp. 297 (W.D. Pa.), *aff'd,*	1993	Amnesty Intl., on behalf of Jane Does	Two classes: (1) all employees	Certification not granted; Alleged

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
	172 F.3d 40 (3d Cir. 1998)			(approximately 25 employees) of the County and City who were assigned to take custody of and process protestors; (2) employees (approximately 10 employees) who were assigned to take into custody and process female protestors	violations of 4th, 13th, and 14th amendments; related to anti-abortion protesting
Heffler v. U. S. Fid. & Guar. Ins. Co.	No. 90-7126, 1992 U.S. Dist. LEXIS 3090 (E.D. Pa.)	1992	All Pennsylvania residents who purchased from ISO members, subscribers and service purchasers motor vehicle bodily injury liability insurance with provisions at issue	All ISO members, subscribers and service purchasers that issued motor vehicle insurance policies in PA with certain provisions at issue	Motion denied; alleged that companies unlawfully restricted the availability of intra-family coverage by introducing a family limitation
Terrestrial Sys., Inc.	132 F.R.D. 71	1990	Terrestrial Systems	Television owners	Case dismissed in

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
v. Fenstemaker	(D. Colo.)			using unauthorized equipment	favor of defendants in pre-trial
Hammond v. Hendrickson	No. 85 C 09829, 1990 U.S. Dist. LEXIS 11071 (N.D. Ill.)	1990	James Hammond	Underwriters represented by Kidder Peabody & Co	Securities case, alleged misrepresentations contained in a prospectus
Alvarado Partners, L.P. v. Mehta	130 F.R.D. 673 (D. Colo.)	1990	Alvarado Partners, L.P.	34 Underwriters	Securities fraud case
Winder Licensing, Inc. v. King Instrument Corp.	130 F.R.D. 392 (N.D. Ill.)	1990	Winder Licensing Inc.	Manufacturers of the patented product at issue	Patent infringement case; class certification denied
Luyando v. Bowen	124 F.R.D. 52 (S.D.N.Y.), *rev'd sub nom.* Luyando v. Grinker, 8 F.3d 948 (2d Cir. 1993)	1989	NY Individuals receiving benefits from Aid to Families with Dependent Children, but have not received first $ 50 of each month's support payment collected periodically by HRA	"All persons who are commissioners of social services districts in New York State."	Challenging $50 pass through law
Williams v. State Bd. of Elections	696 F. Supp. 1574 (N.D. Ill.)	1988	Illinois voters	Five classes: All persons	Voters' suit about judicial elections

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				elected (1) in county at-large elections, (2) elected in suburban-wide at-large elections, (3) in the city-wide at-large elections to seats on the Circuit Court of Cook County; (4) All persons elected or appointed to the Appellate Court of Illinois, First District; (5) All candidates for judicial vacancies in Cook County on the November, 1988 ballot.	
United States v. Rainbow Family	695 F. Supp. 294 (E.D. Tex.)	1988	United States	Rainbow Family and its members	Seeking preliminary injunction against congregation without permit in national

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
					forest
Follette v. Vitanza	658 F. Supp. 492 (N.D.N.Y.), *vacated in part sub nom.* Follette v. Cooper, 671 F. Supp. 1362 (N.D.N.Y. 1987)	1987	Debtors whose wages were garnished	Enforcement officer class: all sheriffs, marshals, constables, or others empowered to enforce income executions upon the wages or other earnings of judgment debtors in NY	Suit based on Consumer Credit Protection Act
Flying Tiger Line, Inc. v. Cent. States	No. 86-304 CMW, 1986 U.S. Dist. LEXIS 17409 (D. Del.)	1986	Flying Tiger Line, Inc.	Pension funds to whom Tiger may be liable because of its prior ownership of Hall's Motor Company (up to 26 class members)	Tiger sought declaratory and injunctive relief that Tiger is not an "employer" and is therefore not subject to the MPPAA
In re Activision Sec. Litig.	No. C-83-4639(A) MHP, 1986 U.S. Dist. LEXIS 18834 (N.D. Cal.)	1986	Stockholders	Underwriters	Securities action

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Davis v. Crush	646 F. Supp. 1192 (S.D. Ohio), *rev'd*, 862 F.2d 84 (6th Cir. 1988)	1986	Owner of medical clinic where abortions are performed	"Persons picketing [the area at issue] who have been personally served with this order as well as their officers, agents, servants, employees, attorneys and those persons in active concert or participation with them who receive personal service of the order."	Preliminary injunction overturned, no class certified
Angel Music, Inc. v. ABC Sports, Inc.	112 F.R.D. 70 (S.D.N.Y.)	1986	Class of music publishers and music copyright owners	"[T]elevision networks, television stations, syndications such as motion picture studios and their television production affiliates, independent television program	Class certification denied; Alleged violation of the Copyright Act, 17 U.S.C. § 101

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				producers and others in the business of creating and selling television programs."	
Emp'rs. Ins. of Wausau v. FDIC	112 F.R.D. 52 (E.D. Tenn.)	1986	Insurers	Federal Deposit Insurance Corporation (FDIC) and individual directors and officers against whom claims were made by the FDIC	Class action motion denied
Akron Ctr. for Reprod. Health v. Rosen	110 F.R.D. 576 (N.D. Ohio), *rev'd sub nom.* Ohio v. Akron Ctr. for Reprod. Health 497 U.S. 502 (1990)	1986	Reproductive health clinic	City prosecutors throughout the state of Ohio	Constitutional challenge to parental notification by physicians who intended to perform certain abortions
Vargas v. Calabrese	634 F. Supp. 910 (D.N.J.)	1986	All qualified black and Hispanic voters in Jersey City who were re-	All district board workers employed throughout Jersey City on June 11,	Defendant class certification denied

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			quired to produce several different forms of identification in order to vote.	1985 and all McCann challengers employed then.	
Mechigian v. Art Capital Corp.	612 F. Supp. 1421 (S.D.N.Y.)	1985	Purchaser of original artwork	"[A]ll former and present investors in the investment plan or plans conducted by defendants."	Brought as a securities case; claims dismissed
Green v. Harbin	615 F. Supp. 719 (N.D. Ala.)	1985	Judgment debtor	Alabama's circuit and district court clerks	Dismissed as to some defendants, but not all; alleged that the state's garnishment laws violated the Due Process Clause
In re Consumers Power Co. Sec. Litig.	105 F.R.D. 583 (E.D. Mich.)	1985	Investors who purchased common stock in Consumers Power Company	Three classes of underwriters (approximately 300 total)	Securities case
Rodriquez v. Twp. of DeKalb	No. 80 C 1509, 1984 U.S. Dist. LEXIS	1984	Applicants for welfare in DeKalb and Joliet	27 Illinois townships	Due process claim related to receiving general

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
	22302 (N.D. Ill.)		townships		assistance, defendant class not certified
Akerman v. Oryx Commc'ns, Inc.	609 F. Supp. 363 (S.D.N.Y.), *aff'd*, 810 F.2d 336 (2d Cir. 1987)	1984	Investors in Oryx Communications	All underwriters of Oryx	Securities case
Harris v. Graddick	593 F. Supp. 128 (M.D. Ala.)	1984	All black citizens of Alabama	All officials responsible for the appointment of poll officials (approximately 198 members)	Alleged that state appointed disproportionately too few black persons as poll officials, in violation of § 2 of the Voting Rights Act of 1965
Nw. Nat'l Bank of Minneapolis v. Fox & Co.	102 F.R.D. 507 (S.D.N.Y.)	1984	Banks making loans to Saxon Securities	All partners in auditing firm Fox & Company in designated time period	Alleged misrepresentations about Saxon (which went bankrupt)
In re Victor Techs. Sec. Litig.	102 F.R.D. 53 (N.D. Cal.), *aff'd*, 792 F.2d 862 (9th Cir. 1986)	1984	Purchasers of stock in question	Securities underwriters	Securities case
Klein v. Council of Chem. Ass'ns	587 F. Supp. 213 (E.D. Pa.)	1984	Individual exposed to chemical used in	"[M]anufacturers, distributors, and	Defendant class not certified

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			printing process	suppliers of the offending carcinogens, acting individually or in concert or conspiracy with one another as members of various trade associations and lobbying groups."	
In re Fortune Sys. Sec. Litig.	No. C 83-3348(A), 1984 U.S. Dist. LEXIS 18090 (N.D. Cal.)	1984	Investors in Fortune Systems	Underwriters of Fortune Systems	Securities case
O'Connell v. David	35 B.R. 146 (E.D. Pa.), *aff'd*, 740 F.2d 958 (3d Cir. 1984)	1983	Chapter 13 trustee	"[I]ndividuals or business entities not licensed to practice law who were alleged to be counseling or advising debtors on the preparation and filing of bankruptcy petitions."	Bankruptcy action
Coleman v. McLaren	98 F.R.D. 638 (N.D. Ill.)	1983	Illinois Taxpayers	All Illinois counties other	Due process and equal

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				than Cook County; all Illinois judges who do not sit in Cook County	protection claims
Cayuga Indian Nation of N.Y. v. Cuomo	565 F. Supp. 1297 (N.D.N.Y.)	1983	Cuyaga Indian Nation	"[A]ll other persons who assert an interest in any portion of the Original Reservation lands" (approximately 7,000 individuals and entities)	Land rights case
Thillens, Inc. v. Cmty. Currency Exch. Ass'n of Ill., Inc.	97 F.R.D. 668 (N.D. Ill.)	1983	Check cashing service	"17 named individual defendants, approximately 350 unnamed individual past and current members of the [Community Currency Exchange Association of Illinois, Inc.,] and	Charged conspiracy under anti-trust and RICO

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				the more than 500 community currency exchanges owned by those members and represented by the Association."	
Wiggins v. Enserch Exploration, Inc.	743 S.W.2d 332 (Tex. Ct. App.)	1987	Corporation and company	The royalty interest owners in the Opelika Gas Unit	Seeking declaratory judgment on meaning of a royalty provision
Wash. Educ. Ass'n v. Shelton Sch. Dist. No. 309	613 P.2d 769 (Wash.)	1980	Statewide organization representing teachers in the K-12 public school system, local education associations, public school women coaches, several parents of school-age daughters.	Statewide organization of junior and senior high schools for interscholastic athletic competition, and 14 local school districts	Sex discrimination suit
Gellantly v. Chelan Cnty.	534 P.2d 1027 (Wash.)	1975	Washington taxpayers	Six WA counties and their officers	Class not certified, challenging levy

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
					limit of taxes
State *ex rel.* Erie Fire Ins. Co. v. Madden	515 S.E.2d 351 (W. Va.)	1998	Individuals who signed releases with the other insurance companies without court approval.	Insurance companies in West Virginia (approximately 300 companies)	Dismissed, no juridicial links
Planned Parenthood Ass'n of Cincinnati, Inc. v. Project Jericho	556 N.E.2d 157 (Ohio)	1990	Health clinic offering abortion services	"[P]ersons picketing [area at issue] who have been personally served with this order as well as their officers, agents, servants, employees, attorneys and those persons in active concert or participation with them who receive personal service of the order."	Seek to enjoin class from interfering with clinic's work
Dayton Women's Health Ctr. v. Enix	555 N.E.2d 956 (Ohio)	1990	Health clinic offering abortion services	Protestors of clinic	Seeking permanent injunction and damages assessed

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
					against the protestors
Carlson v. Indep. Sch. Dist. No. 283	370 N.W.2d 51 (Minn. Ct. App.)	1985	Female teachers who have been denied use of accumulated sick leave during periods of disability relating to pregnancy or child birth	"[A]ll school districts in Minnesota who are or were the employer of the plaintiff class."	Alleged discrimination by denying the use of sick pay for pregnancy and child-birth related disability during maternity leaves
Excelsior Springs v. Elms Redevelopment Corp.	18 S.W.3d 53 (Mo. Ct. App.)	2000	City and redevelopers	Timesharers in a local hotel	Suit to eliminate timesharers' property rights in hotel
State *ex rel.* Ashcroft v. Kansas City Firefighters Local No. 42	672 S.W.2d 99 (Mo. Ct. App.)	1984	State of Missouri	Members of Kansas City Firefighters Local No. 42. (approximately 700 members)	Seeking civil remedy for damages against a labor union to redress a strike by public employees, only applied to 12 named defendants not class
Exxon Corp. v. East Brunswick	470 A.2d 5 (N.J. Super. Ct. App. Div.)	1983	Exxon	"[A]ll taxing jurisdictions within the State in which	Tax court, related to taxing on storage tanks

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				plaintiff owns or leases service stations having underground fuel storage tanks."	
Kronisch v. Howard Savs. Inst.	335 A.2d 587 (N.J. Super. Ct. Ch. Div.)	1975	Mortgagors	Mortgagees under federally insured residential mortgages	Class certification denied, seeking treble damages arising from an alleged conspiracy in restraint of trade under the New Jersey Antitrust Act
Rochester v. Chiarella	448 N.E.2d 98 (N.Y.)	1983	City of Rochester	Real property taxpayers in Rochester	Suit to prevent a multiplicity of lawsuits concerning its prior levy of taxes in excess of constitutional limits
Leon N. Weiner & Assocs., Inc. v. Krapf	No. 8938, 1988 Del. Ch. LEXIS 8 (Del. Ch.)	1988	Corporate lot owner	203 lot owners of North Hills Subdivision,	Class not certified, seeking a declaratory

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				New Castle County	judgment that the property was not subject to any restrictions
Regal Entm't Grp. v. Amaranth LLC	894 A.2d 1104 (Del. Ch.)	2006	Regal Entertainment Group, the issuer of a series of convertible notes	Noteholders (approximately 90 persons), represented by hedge fund	Dispute over method for calculating the number of shares of common stock
Glosser v. Cellcor Inc.	No. 12725, 1995 Del. Ch. LEXIS 16 (Del. Ch.)	1995	Investors	Underwriters	Securities case
In re Broadhollow Funding Corp.	66 B.R. 1005 (Bankr. E.D.N.Y.)	1986	Debtor (Broadhollow Funding Corp)	Investor-creditors of the brokerage business	To determine ownership
Funliner of Ala., LLC v. Pickard	873 So. 2d 198 (Ala.)	2003	All persons who played the video gaming machines over a period of time	Two classes: (1) owners of arcades in which there are 20 or more video-gaming machines for the public's use; (2) entities that lease the video-gaming machines to certain	Class not certified

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				businesses throughout Alabama	
Kadish v. Ariz. State Land Dep't	747 P.2d 1183 (Ariz.)	1987	Arizona taxpayers	"All present and future mineral lessees of state lands."	Seeking declaratory judgment related to revenues from royalty on minerals
In re Dehon, Inc.	298 B.R. 206 (Bankr. D. Mass.)	2003	Plan administrator	All current and former direct or indirect holders of shares of common and/or preferred stock of Dehon, Inc. (1,000+ members).	Bankruptcy proceedings to subordinate the Stock Repurchase Claims to the claims of general unsecured creditors
In re Rusty Jones, Inc.	128 B.R. 1001 (Bankr. N.D. Ill.)	1991	Insurance company	Wisconsin auto rustproofing warranty holders	Certification denied for lack of standing and adequate representation, sought declaratory judgment.
In re Cardinal Indus., Inc.	105 B.R. 834 (Bankr. S.D. Ohio)	1989	Debtors	"[A]ll persons and entities who have or obtain a	Certified defendant class for the sole issue of declara-

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				mortgage or other security interest in property of a limited partnership in which CII or a wholly-owned subsidiary is a general partner."	tory relief
Mojica v. Automatic Emps. Credit Union	363 F. Supp. 143 (N.D. Ill.)	1973	Debtors	Creditors	Dismissed; sought declaratory judgment on automobile repossession and resale provisions in Illinois Commercial Code
Samuel v. Univ. of Pittsburgh	56 F.R.D. 435 (W.D. Pa.)	1972	Two female graduate students	21 named University defendants plus all other state and state-related universities and colleges in PA (71 members)	Challenging financial aid rulings when husband is deemed to be out-of-state
Hodgson v. Hamilton Mun.	349 F. Supp. 1125	1972	United States Depart-	Ohio courts, judges,	Related to OH law on gar-

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Court	(S.D. Ohio)		ment of Labor	and clerks of court	nishment
Smith v. United Bhd. of Carpenters and Joiners of Am.	No. C 75-177, 1976 U.S. Dist. LEXIS 15980 (N.D. Ohio)	1976	Black citizens who were denied employment opportunities within the carpenter construction industry	All employers and/or contractors within the territorial jurisdiction of the union defendants	Alleged discriminatory employment practices
In re Bourns Patent Litig.	385 F. Supp. 1260 (J.P.M.L.)	1974	Patent owner	Companies accused of patent infringement	Certification not granted
Pennsylvania v. Local Union 542, Int'l Union of Operating Eng'rs	469 F. Supp. 329 (E.D. Pa.)	1978	"[M]inority workers involved in or desiring admittance to the operating engineer trade in Eastern Pennsylvania and Delaware."	Local union, 1400 construction contractors and employers receiving referrals through union; construction trade associations for the induction of new operating engineers.	Employment discrimination suit
In re the Gap Stores Sec. Litig.	79 F.R.D. 283 (N.D. Cal.)	1978	Investors	Underwriters	13 cases in multidistrict securities litigation
Institu-	78 F.R.D.	1978	Institu-	"[D]irecto	Class

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
tionalized Juveniles v. Sec'y of Pub. Welfare	413 (E.D. Pa.)		tionalized juveniles in Pennsylvania hospitals	rs of all mental health and mental retardation facilities in Pennsylvania which are subject to regulation by the defendant, Secretary of Public Welfare."	certified
United States v. Trucking Emp'rs, Inc.	75 F.R.D. 682 (D.D.C.)	1977	United States	"[C]omm on carriers of general commodity freight by motor vehicle that employed over-the-road drivers, were parties to or were bound by the national master freight agreement . . . employed at least 100 persons, and had annual gross revenues of at least $1,000,00	Employment discrimination case

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				0."	
Mashpee Tribe v. New Seabury Corp.	427 F. Supp. 899 (D. Mass.)	1977	Mashpee Tribe	Landowners in the town of Mashpee	Property rights
Continental Ill. Nat'l Bank v. Mohr & Sons	No. 80 C 2642, 1980 U.S. Dist. LEXIS 13040 (N.D. Ill.)	1980	Debtors	Creditors	After bankruptcy proceedings
Marchwinski v. Oliver Tyrone Corp.	83 F.R.D. 606 (W.D. Pa.)	1979	Women cleaning personnel	Named defendants plus owners throughout the city who have employed members of the putative plaintiff class.	Class certification denied; Title VII claim and the Labor Management Relations Act claim
Payton v. Abbott Labs	83 F.R.D. 382 (D. Mass.)	1979	All women exposed in utero to a chemical supplied by defendants	All companies that manufactured DES.	Defendant class certification denied on typicality and representativeness
Lynch Corp. v. MII Liquidating Co.	82 F.R.D. 478 (D.S.D.)	1979	Lynch Corporation	MII shareholders of M-Tron Industries (290 members)	Defendant class certified
Mississippi v. United States	490 F. Supp. 569 (D.D.C.)	1979	State of Mississippi	U.S. and all black citizens and black registered	Preclearance for statutory reapportion-

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				voters in Mississippi qualified to vote in state legislative elections	ment
Marcera v. Chinlund	91 F.R.D. 579 (W.D.N.Y)	1981	Detainees held in county jails that did not have a contact visitation program	The sheriffs in charge of the jails	Defendant class certified.
Joseph L. v. Office of Judicial Support of the Court of Common Pleas of Del. Cnty.	516 F. Supp. 1345 (E.D. Pa.)	1981	"[A]ll individuals whose real property has been sold pursuant to the Act at a Delaware County tax sale."	All purchasers, heirs, and assigns, of lands sold at Delaware County Treasurer's tax sales pursuant to 72 P.S. §§ 5971a ff., who had not consummated a quiet title action against the property owners at the time that this action was instituted	Class not certified
In re Itel Sec. Litig.	89 F.R.D. 104 (N.D. Cal.)	1981	Purchasers of securities	Underwriters	Securities case
Stewart v. Winter	87 F.R.D. 760 (N.D. Miss.)	1980	All county jail prisoners statewide	MS sheriffs, boards of supervi-	Classes not certified

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			in MS	sors, and county health officers	
McFarland v. Memorex Corp.	96 F.R.D. 357 (N.D. Cal.)	1982	Purchasers of common stock of Memorex Corporation	Underwriters	Securities case
In re Arthur Treacher's Franchise Litig.	93 F.R.D. 590 (E.D. Pa.)	1982	Franchisor	All franchisees who had executed written contracts with the company and failed to make royalty payments pursuant to the terms of the written contracts	Refused to certify under 23(b)(2)
Doss v. Long	93 F.R.D. 112 (N.D. Ga.)	1981	"[A]ll those who are now or will in the future be civil defendants in Georgia courts operating under the fee system, and also those threatened with actions in those courts."	State judges (1,000+)	Challenging fee system of paying judges by the case

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Diamond v. Charles	476 U.S. 54	1986	Physicians who performed abortions	State's Attorneys in all the counties of the State of Illinois	Challenging abortion law
Bank of N.Y. v. Janowick	470 F.3d 264 (6th Cir.)	2006	Bank of New York	Employee-claimants of stock	To settle claims
Robinson v. Tex. Auto. Dealers Ass'n	387 F.3d 416 (5th Cir.)	2004	Car buyers	Texas car dealerships	Class not certified
S. Ute Indian Tribe v. Amoco Prod. Co.	151 F.3d 1251 (10th Cir.)	1998	Ute Indian Tribe	All persons, except the Tribe and governmental entities, who claim an ownership interest in coalbed methane	Resolve ownership of coalbed mine
Socialist Workers Party v. Leahy	145 F.3d 1240 (11th Cir.)	1998	Socialist Workers Party and Florida Green Party	All sixty-seven Florida county supervisors of elections	Constitutional challenge
Consol. Rail Corp. v. Hyde Park	47 F.3d 473 (2d Cir.)	1995	Interstate railroad, Conrail	Assessing jurisdictions and taxing districts around the state	Alleging violations of the Railroad Revitalization Act and Regulatory Reform Act
League of	999 F.2d	1993	Voters	Texas	Voting

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
United Latin Am. Citizens, Council No. 4434 v. Clements	831 (5th Cir.)		and citizens' league	officials responsible for enforcing a statute	Rights Act
Bachelier v. Hamilton Cnty., Ohio	No. 90-3725, 1991 U.S. App. LEXIS 8829 (6th Cir.)	1991	Planned Parenthood	Abortion protestors outside clinic	Seek to enjoin class from interfering with clinic's work
Real Estate Alliance, Ltd. v. Sarkisian	No. 05-cv-3573, 2007 U.S. Dist. LEXIS 70339 (E.D. Pa.)	2007	Patent holder	Realtors accused of patent infringement	Class certification denied
Albrecht v. Treon	No. 1:06cv274, 2007 U.S. Dist. LEXIS 18613 (S.D. Ohio)	2007	Next of kin of deceased	"[A]ll county coroners and/or medical examiners in the State of Ohio that have removed, retained, and disposed of body parts without prior notice to next of kin." (87 counties)	Class certified
Moffat v. Unicare Midwest Plan Grp. 314541	No. 04 C 5685, 2006 U.S. Dist. LEXIS	2006	Individuals insured under UniCare Midwest	"UniCare Defendants and ERISA Plans that	Class not certified

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
	16348 (N.D. Ill.)		Plan Group	received claims from participants and beneficiaries for Infusion Therapy from Participating Providers and who did not pay that part that exceeded Covered Expenses."	
Iowa Ass'n of Bus. & Indus. v. Efco Corp.	No. 4:04-cv-40270, 2005 U.S. Dist. LEXIS 2382 (S.D. Iowa)	2005	Trade association	Underwriters	Dismissed, claim was under Employee Retirement Income Security Act (ERISA)
Aid for Women v. Foulston	327 F. Supp. 2d 1273 (D. Kan.)	2004	Health care workers	All Kansas county and district attorneys	Seeking declaratory judgment on reporting requirements
Doe v. Miller	216 F.R.D. 462 (S.D. Iowa)	2003	Sex offenders currently living in Iowa	All county attorneys in Iowa	Constitutional challenge to Iowa statute
Forbes v. Woods	71 F. Supp. 2d 1015 (D.	1999	Arizona physicians	All County Attorneys	Constitutional challenge

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
	Ariz.)			who have the authority to enforce A.R.S. § 36-2303	of law related to fetal research
Sebo v. Rubenstein	188 F.R.D. 310 (N.D. Ill.)	1999	Patients	Urologists and shareholders	Conspiracy
N.C. Right To Life, Inc. v. Bartlett	No. 5:96-CV-835-BO(1), 1998 U.S. Dist. LEXIS 6443 (E.D.N.C.)	1998	Political organizations and officers	District attorneys from all 39 state prosecutorial districts	Certification denied
In re Chambers Dev. Sec. Litig.	912 F. Supp. 822 (W.D. Pa.)	1995	"All persons who purchased or acquired Chambers Development Company, Inc., securities from March 18, 1988, through October 20, 1992, inclusive."	All persons who are or were partners of Grant Thornton during the class period.	Securities case
In re Marion Merrell Dow Inc., Sec. Litig.	No. 92-0609-CV-W-6, 1994 U.S. Dist. LEXIS 10053 (W.D. Mo.)	1994	Purchasers of securities of Marion Merrell Dow Inc	Underwriters	Securities case
Deloitte Noraudit A/S v.	148 F.R.D. 523	1993	Deloitte Noraudit	All member firms of Deloitte	Contractual case after

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Deloitte Haskins & Sells	(S.D.N.Y.)			Haskins & Sells International; Deloitte Ross Tohmatsu International	merger
Endo v. Albertine	147 F.R.D. 164 (N.D. Ill.)	1993	Stock purchasers	Underwriters	Securities case
Resolution Trust Corp. v. KMPG Peat Marwick	No. 92-1373, 1992 U.S. Dist. LEXIS 16670 (E.D. Pa.)	1992	Resolution Trust Corporation	All persons who were partners of either Main Hurdman or KMG Main Hurdman in designated time periods	Suit alleging wrongful preparation of financial documents
In re Phar-Mor, Inc. Sec. Litig.	875 F. Supp. 277 (W.D. Pa.)	1994	Equitable Life Assurance Society, et al	Coopers & Lybrand partners and principals	Fraudulent financial activities
Pabst Brewing Co. v. Corrao	161 F.3d 434 (7th Cir.)	1998	Pabst Brewing Company	Retirees	Declaratory judgment on pensions under ERISA
Dale Elecs., Inc. v. R.C.L. Elecs., Inc.	53 F.R.D. 531 (D.N.H.)	1971	Electronics business	Electronics corporations	Patent infringement
LaMar v. H&B	489 F.2d 461 (9th	1973	Customers of pawn	All the pawn	Certification de-

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Novelty & Loan Co.	Cir.)		brokers in Oregon	brokers licensed to conduct business under the laws of Oregon	nied
In re Catawba Indian Tribe Of S.C.	973 F.2d 1133 (4th Cir.)	1992	Catawba Indian Tribe	Occupants and holders of disputed land	Certification denied, sought writ of mandamus
Technograph Printed Circuits, Ltd. v. Methode Elecs., Inc.	285 F. Supp. 714 (N.D. Ill.)	1968	Holder of patents in dispute	"Six subclasses of potential patent infringers; Class 1: All parties who have been or are manufacturing printed circuits by any process claimed in United States Patent No. 2,441,960 or who have been hereafter and before September 5, 1967, notified that they have infringed."	Patent infringement
Weiner v. Bank of King of	358 F. Supp. 684 (E.D. Pa.)	1973	Customers and/or borrowers	Named banks plus all other	Motion dismissed

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Prussia			of national banks in this court's jurisdiction	national banks within jurisdiction	
Coniglio v. Highwood Servs., Inc.	60 F.R.D. 359 (W.D.N.Y.)	1972	All persons who held season tickets for regularly scheduled football games presented by defendant football team	All NFL teams which require their season ticket holders to purchase tickets to exhibition games as well as to regular season games	Defendant class not certified
Ross v. Gerung	69 So.2d 650 (Fla.)	1954	Individual who made repairs to church building	All members of unincorporated religious association	Seeking damages for unpaid bill
O'Connell v. David	35 B.R. 141 (Bankr. E.D. Pa.)	1983	Bankruptcy trustee	"[I]ndividuals and variously styled business entities, none of which are licensed or regulated practitioners of law, nor members in good standing of the Bar of this	Unauthorized practice of law

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				Court or any other court in the Commonwealth of Pennsylvania or the District Court for the Eastern District of Pennsylvania, who have engaged, or in the future will engage, in any of the activities set forth in our Opinion."	
Marchwinski v. Oliver Tyrone Corp.	81 F.R.D. 487 (W.D. Pa.)	1979	Female employees	"[T]hirty to fifty employers in the City of Pittsburgh who are similarly situated to Oliver Tyrone."	Certification not allowed on Title VII claims as matter of law
Research Corp. v. Pfister Associated Growers, Inc.	301 F. Supp. 497 (N.D. Ill.)	1969	Holder of patent	Over 400 seed corn producers	Patent infringement
In re Hotel Tel. Charges	500 F.2d 86 (9th Cir.)	1974	Hotel guests (approximately 40	Hotels (hundreds)	Class not certified

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			million)		
Pipkorn v. Village of Brown Deer	101 N.W.2d 623 (Wis.)	1960	Resident	Beneficiaries of a water trust	Concerning illegal transfer of water trust
Mgmt. Television Sys., Inc. v. NFL	52 F.R.D. 162 (E.D. Pa.)	1971	Operator of closed circuit television systems	Football clubs who are members of the National Football League	Antitrust suit. Defendant class certified.
Kline v. Coldwell, Banker & Co.	508 F.2d 226 (9th Cir.)	1974	Residential home sellers	All real estate brokers who were members of the Los Angeles Realty Board during the 4-year period prior to the filing of the action	Class not certified
Danforth v. Christian	351 F. Supp. 287 (W.D. Mo.)	1972	Attorney General of Missouri	"[A]ll officers and other officials of the state and its political subdivisions charged with enforcement and application of the challenged state laws."	Seeking declaratory judgment related to voting eligibility
Mudd v.	68 F.R.D.	1975	Criminal	All judi-	Defense

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
Busse	522 (N.D. Ind.)		pre-trial detainees	cial officers in Indiana	class not certified
Washington v. Lee	263 F. Supp. 327 (M.D. Ala.)	1966	Current or former prisoners of Alabama state, county, or municipal penal institutions	Wardens, jailers and sheriffs in the State of Alabama	Challenging racial segregation of prisoners
Turpeau v. Fid. Fin. Servs., Inc.	936 F. Supp. 975 (N.D. Ga.)	1996	Borrowers and life insurance insureds	Lenders and life insurers	Defendant class not certified
City of Lebanon v. Holman	402 S.W.2d 832 (Mo. Ct. App.)	1966	City of Lebanon	Property owners of land in dispute	Class not certified, seeking to annex land
City of St. Ann v. Buschard	299 S.W.2d 546 (Mo. Ct. App.)	1957	City of St. Ann	Property owners of land in dispute	Annexation case
Kane v. Fortson	369 F. Supp. 1342 (N.D. Ga.)	1973	Married women who are affected by law in question	All members of Boards of Registrars throughout Georgia.	Challenging law that denies married women in Georgia the right to establish a domicile and residence for voting purposes independent of that of her husband
Adashu-	626 F.2d	1980	School	State and	Classes

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
nas v. Negley	600 (7th Cir.)		children with specific learning disabilities who were not receiving adequate special education	local educational agency members	not certified
Osborn v. Pa.-Del. Serv. Station Dealers Ass'n	94 F.R.D. 23 (D. Del.)	1981	All persons residing in Delaware and Pennsylvania who attempted unsuccessfully to purchase gasoline from a member of the defendant class during the boycott	All service station members of the Pennsylvania-Delaware Service Stations Dealers Association (3,700 members)	Anti-trust
United States v. Truckee-Carson Irrigation Dist.	71 F.R.D. 10 (D. Nev.)	1975	United States	Truckee River Permittees and Newlands Project certificate holders who are members of the TCID	Water rights case
Contract Buyers League v. F & F Inv.	48 F.R.D. 7 (N.D. Ill.)	1969	Minority buyers of houses under land	Home sellers and mortgage lenders	Defendant class not certified, alleged

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			contracts		fraud in the sale of houses under land contracts.
Paxman v. Campbell	612 F.2d 848 (4th Cir.)	1980	Pregnant public school teachers in VA	"[A]ll persons who were or are, during the period December 6, 1969 to June 25, 1975, members of a public county or city school board of the Commonwealth of Virginia which required that a pregnant school teacher cease her teaching at some time during the period of pregnancy other than a time of her own choosing."	Class action reversed on appeal
Benzoni v. Greve	54 F.R.D. 450 (S.D.N.Y.)	1972	Buyers of shares in Sequoyah Industries	There are three actions in this case: 1)Ben-	Securities case

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				zoni: Sequoyah, Merrill Lynch and 15 members of the syndicate of Underwriters, and ten of the selling stockholders who signed the registration statement. 2) Goldman No. 1: Sequoyah, Merrill Lynch, and ten of	

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				the selling stockholders who signed the registration statement. 3)Goldman No. 2: Sequoyah, Greve personally and as a representative of the selling stockholders, and Merrill Lynch individually and as a representative of the under-	

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				writers.	
Tucker v. City of Montgomery Bd. of Comm'rs	410 F. Supp. 494 (M.D. Ala.)	1976	Indigent prisoners	Mayors and recorders throughout the state of AL	Challenging constitutionality of the practice authorized by state statute of confining indigent prisoners to jail in order to allow them to work out fines which they were unable to pay
Guarantee Ins. Agency Co. v. Mid-Cont'l Realty Corp.	57 F.R.D. 555 (N.D. Ill.)	1972	Buyers of stock	Underwriters	Securities case
Hopson v. Schilling	418 F. Supp. 1223 (N.D. Ind.)	1976	All indigent persons in the state	All township trustees in the state that were responsible for administering the state's welfare laws	Certification of defendant class of township trustees is granted based on common juridical link.
Ragsdale v. Turnock	625 F. Supp. 1212 (N.D. Ill.)	1985	Physicians performing or	State's attorneys for all of the coun-	Challenging state law

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
			desiring to perform abortions in the state of Illinois	ties in the state.	
Endo v. Albertine	147 F.R.D. 164 (N.D. Ill.)	1993	Stock purchasers	Underwriters	Defendant class certified only as to the issue of materiality of the alleged misstatements and omissions
Ruocco v. Brinker	380 F. Supp. 432 (S.D. Fla.)	1974	All real property owners in the State of Florida whose real property has or may be encumbered by a Claim of Lien under the Mechanics' Lien Law of Florida	All clerks of the judicial circuits in the State of Florida	Bilateral class action is granted. Plaintiff and defendant classes are certified.
Dudley v. Se. Factor & Fin. Corp.	57 F.R.D. 177 (N.D. Ga.)	1972	Receiver for the Insurance Investors Trust Company	All present and former shareholders of SEFAF who re-	Case may proceed as a class action; defendant class certified.

Appendix Table 1. Defendant Class Actions					
Case	**Cite**	**Year**	**Plaintiff / Plaintiff Class**	**Defendant Class**	**Notes**
				ceived preferred shares of stock in Atlantic Services	

REIMAGINING HUMAN RIGHTS LAW: TOWARD GLOBAL REGULATION OF TRANSNATIONAL CORPORATIONS

RACHEL J. ANDERSON†

INTRODUCTION

Existing human rights law, the body of law that delineates the contours of legal protections for human rights, does not do enough to prevent or provide remedies for corporate-related human rights abuses.[1] Transnational corporations are generally excluded from direct responsibility under international human rights law.[2] The state-centered nature of modern human rights law is inconsistent with the actual power and influence of many transnational corporations.[3] Current human rights law has been conflated with international human rights law and so looks almost exclusively to states to create laws to protect human rights and mechanisms to enforce those laws.[4] However, many modern transnational cor-

---

† Associate Professor of Law, University of Nevada, Las Vegas, William S. Boyd School of Law; J.D. 2005, University of California, Berkeley School of Law; M.A. 2002, Stanford University International Policy Studies; *Zwischenprüfung* 1998, Humboldt-Universität zu Berlin. Research for this article was supported by Dean John V. White. I would like to thank Larry Catá Backer, Monica Bell, Richard Buxbaum, David Caron, Linda Edwards, Mary LaFrance, Andrew Guzman, Fatma Marouf, Saru Matambanadzo, Audrey McFarlane, Ann McGinley, Tom McAffee, David Millon, Jay Mootz, Angela Onwuachi-Willig, Janewa Osei-Tutu, Nancy Rapoport, Paulette Reed-Anderson. Jean Sternlight, Asmara Tekele, Marketa Trimble, the participants at the Lytle Workshop at the University of Kentucky School of Law in June 2010, the participants at Junior Scholars Workshop at the Southeastern Association of Law Schools annual meeting in August 2010, and the participants at the Vulnerability and the 'Corporation' Workshop as part of the Feminism and Legal Theory Project at Emory University School of Law in October 2010 for their comments and suggestions. I am indebted to Jeanne Price and David McClure at the Wiener-Rogers Law Library for their outstanding assistance and support. I also would like to express my deep appreciation to the editors at the *Denver University Law Review* for their hard work, excellent editing, and consistent professionalism.

1. *See* SARAH JOSEPH, CORPORATIONS AND TRANSNATIONAL HUMAN RIGHTS LITIGATION 9 (2004); *see also* Edward L. Rubin, *Response to Comments*, 9 INT'L LEGAL THEORY 167, 173 (2003) ("[H]uman rights discourse fails to reflect the conceptual structure and practical realities of the modern administrative state."). *See generally* JOSEPH, *supra* note 1 (analyzing human rights litigation against transnational corporations).

2. *See* Surya Deva, *Human Rights Violations by Multinational Corporations and International Law: Where From Here?*, 19 CONN. J. INT'L L. 1, 1 (2003) ("[S]ince the existing international mechanism was not designed to apply to [transnational corporations], its inadequacy is exposed.").

3. *See id.* ("States no longer enjoy the monopoly as violators of human rights and no longer solely bear the duty to protect human rights."). States are at the center of international human rights law because they are at the center of international law. *See* IAN BROWNLIE, PRINCIPLES OF PUBLIC INTERNATIONAL LAW 57–58 (5th ed. 1998).

4. *See, e.g.*, International Covenant on Civil and Political Rights art. 2, Dec. 16, 1966, 999 U.N.T.S. 171 [hereinafter ICCPR], *available at* http://www2.ohchr.org/english/law/ccpr.htm ("Where not already provided for by existing legislative or other measures, each State Party to the present Covenant undertakes to take the necessary steps, in accordance with its constitutional processes and with the provisions of the present Covenant, to adopt such laws or other measures as may be necessary to give effect to the rights recognized in the present Covenant."); *see also, e.g.*, International Covenant on Economic, Social, and Cultural Rights art. 2, Dec. 16, 1966, 993 U.N.T.S. 3 [hereinafter ICESCR], *available at* http://www2.ohchr.org/english/law/cescr.htm ("Each State Party

porations have achieved a level of power, wealth, and influence that rivals that of states.[5] Failure to regulate the power, wealth, and influence of transnational corporations is a weakness in human rights law that should be remedied.[6]

This Article argues that transnational corporations require specialized and targeted regulations and laws, and that the conflation of human rights law and international human rights law should be reversed to allow the advancement of other forms of human rights law. It makes two proposals. First, reimagine human rights law and international human rights law as separate categories. Specifically, classify international human rights law as a sub-category of human rights law. This distinction highlights the need to encourage the development of other forms of human rights law, for example, global human rights law and national human rights law. Second, establish global human rights law as a sub-category of human rights law. Specifically, create a new global human rights regime with three main elements: a Global Law Commission, global laws and regulations, and universal civil jurisdiction.

In the summer of 2009, the U.S. news media was dominated by reports about the BP oil spill.[7] The United States Environmental Protection Agency Administrator described it as "the largest environmental disaster in American history."[8] But although similar events have occurred many times over in developing countries, they have not captured the attention of the U.S. media in the same way.[9] For example, a major oil spill the size of the Exxon Valdez disaster has occurred every year for half a cen-

---

to the present Covenant undertakes to take steps, individually and through international assistance and co-operation, especially economic and technical, to the maximum of its available resources, with a view to achieving progressively the full realization of the rights recognized in the present Covenant by all appropriate means, including particularly the adoption of legislative measures."). However, "international legal institutions typically only have advisory powers and are unable to 'make' states take particular action." Angela M. Banks, *CEDAW, Compliance, and Custom: Human Rights Enforcement in Sub-Saharan Africa*, 32 FORDHAM INT'L L.J. 781, 782 (2009).

5. JOSEPH, *supra* note 1, at 1; Hope Lewis, *Embracing Complexity: Human Rights in Critical Race Feminist Perspective*, 12 COLUM. J. GENDER & L. 510, 519 (2003) ("Many [transnational corporations] have more economic resources at their disposal than the entire budget available to some developing countries; they can exert influence that approximates that of a state in some circumstances.").

6. *See* Deva, *supra* note 2, at 3 ("[T]he approach of indirect regulation has failed to deliver the desired results.").

7. For example, a search on July 3, 2010 for "BP oil spill 2010" in Google news resulted in 22,100 hits.

8. Lisa P. Jackson, *Administrator Jackson: Keep Moving America Forward into Energy Independence*, ENVIRONMENTAL PROTECTION AGENCY , http://blog.epa.gov/administrator/2010/06/08/administrator-jackson-%e2%80%9ckeep-moving-america-forward-into-energy-independence%e2%80%9d/ (last updated Oct. 27, 2010). However, others contend that the "Dust Bowl" of the 1930s may be the largest environmental disaster on U.S. territory. *See, e.g.*, Ed Stoddard, *Is Gulf Spill the Worst Ecological Disaster in U.S. History?*, REUTERS ENVIRONMENT FORUM (June 22, 2010, 10:09 AM), http://blogs.reuters.com/environment/2010/06/is-gulf-spill-the-worst-ecological-disaster-in-u-s-history/.

9. A search for "Niger Delta oil spill" in Google news on July 3, 2010 resulted in 154 hits.

tury in the Niger Delta.[10] Environmental disasters overseas often involve a subsidiary of a U.S. corporation, as was the case in the Bhopal disaster in which toxic gas leaked from a pesticide plant, killing an estimated 2,100 people and injuring over 200,000 others.[11] Nonetheless, oil spills in the Niger Delta and other developing countries—and the harms they cause— do not receive the same level of attention in the U.S. media.[12]

Environmental catastrophes like the 2010 BP oil spill and the decades of oil spills in other places like the Niger Delta impinge upon human rights—such as the rights to life, health, adequate food and housing, and clean water.[13] Life, health, and other human rights are enumerated, for example, in the 1948 Universal Declaration of Human Rights, the 1966 International Covenant on Civil and Political Rights, the 1966 International Covenant on Economic, Social, and Cultural Rights, and other human rights documents.[14] In the BP oil spill, eleven people were killed in the explosion.[15] In some cases in the Niger Delta, oil drilling and associated gas flares have made entire villages uninhabitable.[16] Oil drilling and gas flares have had devastating effects on both the environment and human health, and have led to convulsions, chromosomal dam-

10. Adam Nossiter, *Far From Gulf, a Spill Scourge 5 Decades Old*, N.Y. TIMES, June 17, 2010, at A1, *available at* http://www.nytimes.com/2010/06/17/world/africa/17nigeria.html ("The Niger Delta, where the wealth underground is out of all proportion with the poverty on the surface, has endured the equivalent of the Exxon Valdez spill every year for 50 years by some estimates."). For details of the amount of oil spilled in the Niger Delta between 1976 and 1996, see Odjuvwuederhie Emmanuel Inoni, Douglason Gordon Omotor & Felicia Nkem Adun, *The Effect of Oil Spillage on Crop Yield and Farm Income in Delta State, Nigeria*, 7 J. CENT. EUROPEAN AGRIC. 41, 43 (2006), *available at* http://www.agr.hr/jcea/issues/jcea7-1/pdf/jcea71-6.pdf.

11. *In re* Union Carbide Corp. Gas Plant Disaster at Bhopal, India in Dec., 1984, 634 F. Supp. 842, 844 (S.D.N.Y. 1986), *aff'd and modified*, 809 F.2d 195 (2d Cir. 1987).

12. A search for "Niger Delta oil spill" on July 3, 2010 resulted in less than one percent of the hits for "BP oil spill 2010" on the same date although the search is narrower and there has been the equivalent of a major spill in the Niger Delta every year for at least 50 years.

13. *See* U.N. Special Representative of the Secretary General, *Protect, Respect & Remedy: A Framework for Business & Human Rights: Second Addendum*, 3, 29, U.N. Doc. A/HRC/8/5/Add.2 (May 23, 2008) [hereinafter *Protect, Respect & Remedy*], *available at* http://daccess-dds-ny.un.org/doc/UNDOC/GEN/G08/136/61/PDF/G0813661.pdf?OpenElement.

14. ICCPR, *supra* note 4, at art. 6; ICESCR, *supra* note 4, at art. 7, 12; Universal Declaration of Human Rights, G.A. Res. 217 (III) A, U.N. Doc. A/Res/217(III) (Dec. 10, 1948) [hereinafter UDHR]. However, there continues to be disagreement about the legitimacy and enforceability of economic, social, and cultural rights. *See, e.g.*, Edward L. Rubin, *Rethinking Human Rights*, 9 INT'L LEGAL THEORY 5, 5–6 (2003) (discussing the controversy associated with the extent to which social, cultural, and economic rights are accepted).

15. Bradley Blackburn, *BP Oil Spill: Families Gather to Honor 11 Who Died, Express Frustration With BP, Transocean*, ABC NEWS (May 25, 2010), http://abcnews.go.com/WN/bp-oil-spill-transocean-holds-memorial-11-lost/story?id=10739080.

16. *Factsheet: Shell's Environmental Devastation in Nigeria*, CENTER FOR CONST. RTS., http://ccrjustice.org/learn-more/faqs/shell%2526%2523039%3Bs-environmental-devastation-nigeria [hereinafter CCR] (last visited Oct. 27, 2010). The devastation was so extensive that Ken Saro-Wiwa, the leader of the Movement for the Survival of the Ogoni People, described the oil drilling operations as "an ecological war against the Ogoni." Fifth Amended Complaint and Demand for Jury Trial ¶¶ 1, 53, Wiwa v. Shell Petroleum, No. 96 Civ. 8386 (S.D.N.Y. Mar. 16, 2009), http://ccrjustice.org/files/3.16.09%205th%20Amended%20Complaint.pdf [hereinafter Wiwa Fifth Amended Complaint].

age, birth defects, and other serious illnesses.[17] The U.S. news was filled with pictures of the extreme damage of BP's spills and yet most people in the United States are completely unaware of the Niger Delta disasters.

However, environmental catastrophes are not the only way that the operations of transnational corporations like BP negatively affect human rights.[18] Transnational corporations, as "economic entit[ies] operating in more than one country or [as] a cluster of economic entities operating in two or more countries," help shape the economic, political, social, and legal environments in which they operate.[19] Transnational corporations are a subset of corporations and, therefore, much of the discussion of transnational corporations in this Article also applies to domestic corporations, although there are, of course, important differences.[20]

Transnational corporations also impinge on human rights in the labor context.[21] One well-known example is the Nike scandal in the 1990s, when Life magazine exposed Nike's involvement with the use of child labor in the production of soccer balls by publishing a picture of a child assembling Nike soccer balls in Pakistan.[22] More recently, underpaid workers, including child workers, produced the soccer balls used in the 2010 World Cup.[23] This case of corporate-related human rights abuse—thirteen years after the initial Nike scandal—received only minimal media attention.

A 2007 U.N. study reviewed 320 alleged cases of corporate-related human rights abuses and concluded that corporations affect "the full range of human rights" through their acts or omissions.[24] The study fo-

---

17. CCR, *supra* note 16.

18. JOSEPH, *supra* note 1, at 2; *Protect, Respect & Remedy*, *supra* note 13, at 2.

19. Comm'n on Human Rights, Subcomm'n on Promotion and Protection of Human Rights, Rep. on its 55th Sess., July 28–Aug. 15, 2003, U.N. Doc. E/CN.4/Sub.2/2003/12/Rev.2 (Aug. 26, 2003), *available at* http://www.unhchr.ch/huridocda/huridoca.nsf/(Symbol)/E.CN.4.Sub.2.2003.12.Rev.2.En. The influence of transnational corporations is discussed *infra* in Part II.

20. *See, e.g.*, Mark Gibney & R. David Emerick, *The Extraterritorial Application of United States Law and The Protection of Human Rights: Holding Multinational Corporations to Domestic and International Standards*, 10 TEMP. INT'L & COMP. L.J. 123, 126–27 (1996) (discussing "the vastly different standards that exist under U.S. law that give [transnational corporations from the U.S.] license to treat U.S. citizens in one manner, but allow foreigners to be treated in a completely different (and inferior) way").

21. JOSEPH, *supra* note 1, at 2; *Protect, Respect & Remedy*, *supra* note 13, at 2.

22. Richard M. Locke, *The Promise and Perils of Globalization: The Case of Nike* 11–13 (Mass. Inst. of Tech. Indus. Performance Ctr. Working Paper Series, Paper No. 02-007, 2002), *available at* http://web.mit.edu/ipc/publications/pdf/02-007.pdf.

23. *See* Michelle Chen, *FIFA's World Cup Having a Ball with Child Labor*, COLOR LINES (June 11, 2010, 10:42 AM), http://colorlines.com/archives/2010/06/south_africas_world_cup_brims_with_broken_promises.html; James Rupert, *World Cup Profits Bypass Asian Soccer-Ball Stitchers (Update 1)*, BLOOMBERG BUSINESSWEEK (June 9, 2010, 4:39 AM), http://www.businessweek.com/news/2010-06-09/world-cup-profits-bypass-asian-soccer-ball-stitchers-update1-.html; Trina Tocco, *World Cup Soccer Balls Missed the Goal Set 13 Years Ago: Child Labor, Poverty Wages, Temporary Workers*, COMMONDREAMS.ORG (June 7, 2010, 10:46 AM), http://www.commondreams.org/newswire/2010/06/07-6.

24. *Protect, Respect & Remedy*, *supra* note 13, at 1–3, 29. The cases were posted on the Business and Human Rights Resource center webpage (http://www.business-humanrights.org),

cused on the rights enumerated in the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, the International Covenant on Economic, Social and Cultural Rights, and International Labour Organization Core Conventions.[25] These include "civil and political rights; economic, social and cultural rights; and labor rights."[26] The study found that corporate-related human rights abuses take place in all industrial sectors and in all regions of the world.[27] In some cases, the harms include loss of life.[28]

Although corporate-related human rights abuses are not proportionately represented in the U.S. media, scholars and decision-makers are aware of the issue.[29] So, if scholars and decision-makers have been aware of corporate-related human rights abuses for decades at the very least, why are there still such wide-scale problems with corporations negatively affecting human rights? This Article argues that the answer is not simply a lack of regulations or a lack of enforcement. Instead, it argues that part of the answer lies in the conflation of human rights law with international human rights law, and the ways that human rights law intersects with corporate law and foreign direct investment law.[30]

Addressing the complex problem of human rights abuses by transnational corporations requires a comprehensive framework.[31] One avenue toward a comprehensive framework is international law; another avenue is global law. International law is the law between nation-states

---

which "is the most comprehensive" and "objective" resource on the intersection of business and human rights. *Id.* at 6.

25. *Protect, Respect & Remedy*, *supra* note 13, at 2. *See generally* ICCPR, *supra* note 4; ICESCR, *supra* note 4; UDHR, *supra* note 14; INT'L LABOUR OFFICE, THE INTERNATIONAL LABOUR ORGANIZATION'S FUNDAMENTAL CONVENTIONS 8 (2002), *available at* http://www.ilo.org/wcmsp5/groups/public/---ed_norm/---declaration/documents/publication/wcms_095895.pdf. There are eight core International Labour Organization conventions: Freedom of Association and Protection of the Right to Organise Convention, 1948 (No. 87); Right to Organise and Collective Bargaining Convention, 1949 (No. 98); Forced Labour Convention, 1930 (No. 29); Abolition of Forced Labour Convention, 1957 (No. 105); Minimum Age Convention, 1973 (No. 138); Worst Forms of Child Labour Convention, 1999 (No. 182); Equal Remuneration Convention, 1951 (No. 100); Discrimination (Employment and Occupation) Convention, 1958 (No. 111). *Id.*

26. *Protect, Respect & Remedy*, *supra* note 13, at 2.

27. *Id.* Based on the cases reviewed in the U.N. study, the regional breakdown of alleged human rights incidents is as follows: Asia & the Pacific – 28%; Africa – 22%; Latin America – 18%; Global – 15%; North America – 7%; Europe – 3%; and Middle East – 2%. *Id.* at 8.

28. JOSEPH, *supra* note 1, at 2–4.

29. *See generally, e.g.*, JOSEPH, *supra* note 1; Locke, *supra* note 22; *Protect, Respect & Remedy*, *supra* note 13.

30. For a parallel argument addressing the question of why transnational corporations do not have obligations under international human rights law, see Iris Halpern, *Tracing the Contours of Transnational Corporations' Human Rights Obligations in the Twenty-First Century*, 14 BUFF. HUM. RTS. L. REV. 129, 131 (2008) (positing that the omission of transnational corporations is "a product of the systemic separation between international economic development, human rights enforcement, and the regulation of private players").

31. *See id.* at 134 ("The international legal system must be refashioned so as to be capable of simultaneously regulating all the numerous important actors vis-à-vis their human rights behavior.").

and, as such, it does not directly regulate transnational corporations.[32] Instead, it regulates the regulators, the nation-states, thereby concentrating on nation-states and marginalizing issues that do not directly implicate nation-states.[33] Since transnational corporations are (by definition) not nation-states, they do not fall within the natural scope of international law. Further, since transnational corporations operate in multiple jurisdictions, the laws of any one jurisdiction are not sufficient to govern their activities.[34]

Global law is an emerging legal order. It is a next iteration of law, following the law of nations and international law.[35] Global law is neither superior nor inferior to other legal orders.[36] Instead, it presupposes the interconnection and interdependency of all legal orders of the world, including international law and national law.[37] Human rights are a core value of global law.[38] There are multiple sources of global law, including specific economic or other subsectors, and organizational and functional networks.[39] *Lex mercatoria*, also known as commercial law, transnational law, or the New Law Merchant, is an example of global law.[40] Through various means, global law provides an opportunity to address corporate-related human rights abuses, in part because it is not state-centered.

This Article is part of a larger project. My previous article, *Toward Global Corporate Citizenship: Reframing Foreign Direct Investment Law*, argued that the asymmetry and fragmentation of foreign direct in-

---

32. *See* PHILIP C. JESSUP, TRANSNATIONAL LAW 1–2 (1956). As early as 1956, Philip Jessup noted the inadequacy of using "international law" to address issues that arise within the "complex interrelated world community." *Id.* at 1. Instead, Jessup used the term "'transnational law' to include all law which regulates actions or events that transcend national frontiers." *Id.* at 2 ("Both public and private international law are included, as are other rules which do not wholly fit into such standard categories."). For a more detailed discussion of the term "transnational," see generally *id.* Transnational law, thus, may regulate "individuals, corporations, states, organizations of states, or other groups." *Id.* at 2–3.

33. *See id.* at 11.

34. *See id.* at 4–5; *see also* Rachel J. Anderson, *Toward Global Corporate Citizenship: Reframing Foreign Direct Investment Law*, 18 MICH. ST. J. INT'L L. 1, 2 (2009).

35. RAFAEL DOMINGO, THE NEW GLOBAL LAW, at xiv (2010); *see also* Harold J. Berman, *World Law*, 18 FORDHAM INT'L L.J. 1617, 1617 (1995). For more on national human rights law, see generally UNITED NATIONS, BUSINESS AND HUMAN RIGHTS: A SURVEY OF NHRI PRACTICES (2008), *available at* http://www.reports-and-materials.org/OHCHR-National-Human-Rights-Institutions-practices-Apr-2008.doc.

36. *See* DOMINGO, *supra* note 35, at 147.

37. *Id.*; *see also* Ralf Michaels, *The True Lex Mercatoria: Law Beyond the State*, 14 IND. J. GLOBAL LEGAL STUD. 447, 447 (2007) (arguing that "the dichotomy of anational law and state law is false").

38. DOMINGO, *supra* note 35 , at 142–44.

39. Gunther Teubner, *'Global Bukowina': Legal Pluralism in the World Society*, *in* GLOBAL LAW WITHOUT A STATE 3, 3–4 (Gunther Teubner ed., 1997). See generally *id.* for a discussion of global law.

40. *Id.* at 3. *Lex mercatoria* is also defined as "the transnational law of economic transactions." *Id.*; *see also* ANA M. LÓPEZ RODRÍGUEZ, LEX MERCATORIA AND HARMONIZATION OF CONTRACT LAW IN THE EU 90 (2003) (defining global law as "a body of rules, different in origin and content, created by the community of merchants to serve the needs of international trade").

vestment law encourages excesses by transnational corporations.[41] That article proposed transforming the theories and practices of voluntary Global Corporate Citizenship into a mandatory legal framework and developing a legal theory of Global Corporate Citizenship that reconceptualizes the role of transnational corporations in the global economy. This Article builds on *Toward Global Corporate Citizenship* and attempts to reimagine human rights law.[42] It argues for the development of a global law regime as a sub-category of human rights law, distinct from international human rights law, and proposes a global institution, global laws, and global enforcement to regulate transnational corporations and help fill gaps in human rights law.[43] This Article makes that argument in three stages.

Part I, Limits of a State-Centered Human Rights Regime, argues for the re-remembering of human rights law as the super-category and international human rights law as one of several possible sub-categories of human rights law. In the wake of the dramatic expansion of international human rights law in the post-World War II era, it came to be thought of as synonymous with human rights law. This conflation of human rights law and international human rights law has inhibited the development of other sub-categories of human rights law such as global human rights law and national human rights law. Although international human rights law has achieved significant progress, the state-centered human rights regime is limited, and, in the absence of other forms of human rights law, leads to the under-regulation of transnational corporations as a result of state resistance, impotency, and complicity. Part I proposes decoupling and distinguishing between human rights law and international human rights law to allow for the development of other forms of human rights law, including global human rights law.

Part II, Transnational Corporations Need Dedicated Regulation, posits that transnational corporations require dedicated regulation under human rights law that goes beyond the ambit of international human rights law. International human rights law requires states to enact and enforce laws to protect human rights within their jurisdiction. However, international human rights law alone is an insufficient tool with which to regulate transnational corporations when their economic, political, and

41. *See* Anderson, *supra* note 34, at 6.

42. This article attempts to respond to the challenge identified by Philip Alston in *The 'Not-a-Cat' Syndrome: Can the International Human Rights Regime Accommodate Non-State Actors?*, *in* NON-STATE ACTORS AND HUMAN RIGHTS 3, 4 (Philip Alston ed., 2005) ("The challenge that [the state-centered focus of the international human rights regime] lays down is one of re-imagining, as the social scientists would put it, the nature of the human rights regime and the relationships among the different actors within it. Lawyers, not being noted for their willingness to depart from precedents, might prefer to see the task in terms for re-interpreting existing concepts and procedures rather than re-imagining.").

43. Sarah Joseph has argued, "Ultimately, a preferable approach might be for all nations to agree on international minimum human rights standards for TNCs, which could be incorporated into national legislation and enforced by domestic courts." JOSEPH, *supra* note 1, at 153.

legal influence exceeds that of the countries in which they operate. The interests of states in preserving sovereignty inhibit the development of comprehensive multijurisdictional international human rights enforcement mechanisms. Finally, the importance of human rights as core values contrasted with the inadequacy of international human rights law as the sole tool to protect them at the global level, demonstrate the need for protection and enforcement of these rights in multiple forms of human rights law.

Part III, Global Regulation of Transnational Corporations, sets out a proposal for the development of global human rights law as a subcategory of human rights law that could address the problem of corporate-related human rights abuses. This proposal has three main components: creation of a Global Law Commission, development of global laws, and implementation of universal civil jurisdiction. A primary purpose of the Global Law Commission would be to develop global human rights law regulations and legislation to prevent and address corporate-related human rights abuses. The Global Law Commission would develop model global regulations and laws informed by theories of Global Corporate Citizenship and promote their enactment. Global regulations and laws could be enforceable via multiple avenues, including national courts, alternative dispute resolution, and universal civil jurisdiction for corporate-related human rights abuses. However, this proposal does not preclude the creation of a global court or other global mechanism for adjudication or alternative dispute resolution.

I. LIMITS OF A STATE-CENTERED HUMAN RIGHTS REGIME

International human rights law categorizes all actors on the global stage as either state actors or non-state actors.[44] For the purposes of international human rights law, non-state actors can be defined only by their relationship to the state.[45] This state-centered focus inhibits an accurate analysis of corporate-related human rights abuses and limits the development of measures with which these issues can be addressed.[46] In light of these weaknesses, this Article argues for the development of other forms of human rights law.

Although states are no longer assumed to be the only actors in the international arena, modern human rights law remains state-centered.[47]

---

44. Alston, *supra* note 42, at 3.

45. *Id.* at 3–4.

46. *See id.* at 4 ("[S]uch a uni-dimensional or monochromatic way of viewing the world is not only misleading, but also makes it much more difficult to adapt the human rights regime in order to take adequate account of the fundamental changes that have occurred in recent years.").

47. *See* Deva, *supra* note 2, at 1 ("The conventional international framework for protection of human rights is state-centric; it obligates primarily states to promote, and not violate, human rights."); PETER MALANCZUK, AKEHURST'S MODERN INTRODUCTION TO INTERNATIONAL LAW 209 (7th ed. 1997) ("[E]very individual has certain inalienable and legally enforceable rights protecting him or her against state interference and the abuse of power by governments."); *see also* Rafael

This means that when we think of human rights law, we think primarily of *international* human rights law. Modern international human rights law is anchored in the series of international declarations and conventions that were generated in the decades following the fall of the Third Reich and the end of World War II.[48] The atrocities perpetuated by the Nazi state against its own citizens and those of other countries were still fresh in the minds of the policy makers who organized the Nuremburg trials and drafted and signed the Universal Declaration of Human Rights.[49] This explains, in part, the post-World War II focus on state action.

However, although corporations were often complicit and, in many cases, actively involved in human rights abuses under the Nazi regime and in many other contexts, the emerging international human rights regime did not incorporate direct rules governing transnational corporations.[50] Although some argue that the primary focus of international decision-makers in the post-World War II period was responding to the atrocities of the Nazi regime, the reasons more likely lie in the complexities of political machinations and conflicts of interest set in a particular historical context.[51] The exclusion of transnational corporations from international human rights law is not inevitable, but rather results from historical events, flawed assumptions, and lack of political will among some influential policy makers.[52]

Domingo, *The Crisis of International Law*, 42 VAND. J. TRANSNAT'L L. 1543, 1549–50 (2009). For a discussion of the nation-state paradigm and its history, see *id.* at 1556–66.

48. *See* Austen L. Parrish, *Changing Territoriality, Fading Sovereignty, and Development of Indigenous Rights*, 31 AM. INDIAN L. REV. 291, 296–97 (2007). The post-World War II period marks the beginning of broad international action in the area of human rights. *See* MALANCZUK, *supra* note 47, at 209. Prior to World War II, actions addressing human rights concerns tended to be focused on specific regions, groups, or abuses. *Id.*

49. *See* LOUIS HENKIN, THE AGE OF RIGHTS 1 (1990).

50. *See* Halpern, *supra* note 30, at 130–31.

51. *See* Mark Mazower, *The Strange Triumph of Human Rights, 1933–1950*, 47 HISTORICAL J. 379, 380, 397 (2004) ("It does no service to the cause of human rights to disguise the political struggles and conflicts of interest that accompanied their emergence into the international arena. On the contrary, a better understanding of that story, their relationship to prior rights regimes, and their dependence on the international balance of power may help us recognize their true weight and worth."); *see also* Kenneth Cmiel, *The Recent History of Human Rights*, 109 AM. HIST. REV. 117, 119 (2004) ("While university-based historians such as Paul Lauren, Lynn Hunt, and Jeffrey Wasserstrom have addressed the subject, journalists, legal scholars, political activists, and political scientists have still done far more of this history writing. The field remains refreshingly inchoate.").

52. *See* ANTONY ANGHIE, IMPERIALISM, SOVEREIGNTY AND THE MAKING OF INTERNATIONAL LAW 111–14 (2004) (arguing that modern international law is deeply and fundamentally shaped by imperialism and colonialism). Colonial trading companies, the predecessors of modern transnational corporations, were arms of the state and, therefore, did not require a separate body of governing law. Their purpose was to further political, legal, economic, and social goals. For example, the Dutch West India Company was "started as a move in the war game" and "[t]he aim and objective of the Company had from the first been to carry on active war with Spain." Hugh E. Egerton, *The Transference of Colonial Power to the United Provinces and England* 728, 749, *in* 4 THE CAMBRIDGE MODERN HISTORY: THE THIRTY YEARS' WAR, 749 (A.W. Ward, G.W. Prothero & Stanley Leathes eds., 1906). The shareholders of the Dutch West India Company were five Dutch government institutions, the Dutch legislature made annual payment to the Dutch West India Com-

Protected from the reach of international human rights law, transnational corporations continue to impinge upon human rights. The most common labor-related rights that the operations of transnational corporations affect include "the right to work (34%), [the] right to just and favorable remuneration (30%), the right to a safe work environment (31%), and the right to rest and leisure (25%)."[53] For example, in the late 1990s, an Ernst and Young[54] report on a Nike subcontractor in Vietnam claimed that concentrations of a chemical solvent, chemical releases, and excess dust in the shoe plant (in each case, many multiples above the allowed levels) had caused extensive harm to the workers' human rights.[55] Specifically, workers suffered from respiratory ailments and skin and heart disease that were allegedly caused by the health and safety violations at the factory.[56] The Ernst and Young report also stated that the employees were forced to work more hours than allowed by Vietnamese law.[57]

As mentioned above, not only do corporate-related human rights abuses take place in the labor context, they are also linked to environmental harms. Those human rights affected include the right to physical health and to an adequate standard of living, and the right to life, liberty and personal security.[58] Corporate-related human rights harms are often caused by pollution, contamination, and environmental degradation.[59] For example, the plaintiffs in *Sarei v. Rio Tinto PLC*[60] alleged that Rio Tinto's Panguna Mine in Papua New Guinea polluted the Kawerong-Jaba River with waste and poisoned the air with dust and emissions from a copper concentrator.[61] The plaintiffs asserted that this air and water pollution led to an increase in respiratory infections and asthma, and that the decreased food supply due to crop damage and the deaths of traditional sources of food like fish led to health problems in the local popula-

---

pany, and the Dutch legislature promised to provide the Dutch West India Company with a fleet of ships in the event of a serious war. *Id.* at 749–50.

53. *Protect, Respect & Remedy*, *supra* note 13, at 11.

54. Ernst and Young is a global accounting firm that provides assurance, tax, transaction, and advisory services. *See About Us*, ERNST AND YOUNG, http://www.ey.com/US/en/About-us (last visited Oct. 28, 2010).

55. Locke, *supra* note 22, at 13; *see also* Steven Greenhouse, *Nike Shoe Plant in Vietnam is Called Unsafe for Workers*, N.Y. TIMES, Nov. 8, 1997, at A1, *available at* http://www.nytimes.com/1997/11/08/business/nike-shoe-plant-in-vietnam-is-called-unsafe-forworkers.html?scp=1&sq=nike%20shoe%20plant%20in%20vietnam%20is%20called%20unsafe&st=cse.

56. Locke, *supra* note 22, at 13.

57. Greenhouse, *supra* note 55 (stating that employees were forced to work 65 hours a week for $10). Over 9,000 workers were employed at the factory. Locke, *supra* note 22, at 13.

58. *Protect, Respect & Remedy*, *supra* note 13, at 2.

59. *See id.*

60. 221 F. Supp. 2d 1116 (C.D. Cal. 2002), *aff'd in part, vacated in part, rev'd in part*, 456 F.3d 1069 (9th Cir. 2006), *withdrawn and superseded on reh'g in part*, 487 F.3d 1193 (9th Cir. 2007), *reh'g en banc granted*, 499 F.3d 923 (9th Cir. 2007). On October 26, 2010, an en banc panel of the Ninth Circuit referred the case to a mediator to "explore the possibility of mediation." Sarei v. Rio Tinto, 02-cv-56256 (9th Cir. Oct. 26, 2010). The mediator is scheduled "to report to the en banc court within twenty-eight (28) days as to whether mediation should proceed or whether this case should be returned to the en banc court." *Id.*

61. *Id.* at 1124 & n.31.

tion.[62] Corporate-related human rights abuses harm not only individuals but also entire communities.[63]

In addition to direct abuses such as under-compensation or chemical spills, abuses by transnational corporations create ripple effects that lead to additional harms.[64] For example, exploitative compensation practices can contribute to harassment and sexual abuse of women.[65] The use of child labor affects enjoyment of the right to education and, in some cases, the rights to health and even life.[66] By way of illustration, children between the ages of five and seventeen are employed at vanilla orchards in Madagascar.[67] They are among the twenty-eight percent of children in Madagascar who are employed in the agriculture and fishing industries.[68] These children are unable to enjoy the right to education because they work six to seven hours a day for approximately twelve cents a day.[69]

Corporations also contribute to or benefit from indirect involvement in human rights abuses by third parties—including governments, other businesses, and individuals.[70] For example, the Swedish Company Lundin Oil AB, together with Sudapet Ltd.—which is wholly owned by the Sudanese government—has been accused of complicity in war crimes and crimes against humanity perpetrated by government security forces.[71] Most of the indirect cases are alleged in Africa, Asia and the Pacific, Latin America, and the Middle East.[72]

Even though we tend to think of human rights in terms of international human rights law—that is, in a post-1948, post-Universal Declaration of Human Rights context—human dignity and other concepts underlying human rights predate the Universal Declaration of Human Rights and even predate the nation state. Louis Henken,[73] a leading international

---

62. *Id.*

63. *Protect, Respect & Remedy*, *supra* note 13, at 4, 29.

64. *Id.* at 3, 29.

65. Lewis, *supra* note 5, at 519 ("Transnational corporations exploit the low-wage status of women (often women of color) and, in doing so, create new avenues for sexual abuse and harassment."); *see also Nike Admits Abuse at Indonesian Plants*, BBC NEWS (Feb. 22, 2001, 13:55 GMT), http://news.bbc.co.uk/2/hi/asia-pacific/1184103.stm (reporting that underpaid workers in a Nike plant in Indonesia, who are 85% women, have reported being coerced into sex and being fondled by managers).

66. *Protect, Respect & Remedy*, *supra* note 13, at 3.

67. Dan McDougall, *Bitter Plight of the Vanilla Trade Children*, THE SUNDAY TIMES (Mar. 14, 2010), http://www.timesonline.co.uk/tol/news/world/africa/article7060962.ece.

68. *Id.*

69. *See id.*

70. *Protect, Respect & Remedy*, *supra* note 13, at 4.

71. EUROPEAN COAL. ON OIL IN SUDAN, UNPAID DEBT: THE LEGACY OF LUNDIN, PETRONAS AND OMV IN BLOCK 5A, SUDAN 1997-2003, at 10 (2010), *available at* http://www.ecosonline.org/reports/2010/UNPAID_DEBT_fullreportweb.pdf. ("[T]here are grounds to investigate whether the Consortium provided financial and material support to the security agencies that were responsible for the commission of international crimes and gross violations of human rights.").

72. *Protect, Respect & Remedy*, *supra* note 13, at 15.

73. Louis Henkin (1917-2010) was a professor of law at Columbia University, a president of the American Society of International Law, and held many important positions over the course of his

law and human rights scholar, defined human rights as "legitimate, valid, [and] justified claims" of individuals and communities upon society.[74] Although human rights are claims or entitlements that can be asserted, they are not always coexistent with legal rights.[75] Human rights arise on the basis of a person's humanity and do not require anything more than the nature of a human being to exist; legal rights are government created and require a government or other institutional act to exist.[76] Thus, human rights may exist, even when law does not yet protect them.

A. Underregulation of Transnational Corporations

Modern human rights law does not generally directly regulate transnational corporations. International human rights agreements are—by definition—agreements between nation-states.[77] International human rights documents range from recording understandings of human rights and human dignity, to setting out aspirational goals, to establishing requirements and guidelines for state actors.[78] The idea is that state signatories of international agreements will enact laws that protect human rights.[79] Such laws would also apply to domestic and transnational corporations where appropriate.[80]

The exclusion of transnational corporations from the international human rights regime puts the burden on states to enact and enforce laws that protect human rights.[81] However, this expectation, although arguably correct in principle, does not account for realities that make state en-

---

life, including at the State Department, the United Nations, and the Permanent Court of Arbitration. *Louis Henkin*, COLUM. L. SCH., http://www.law.columbia.edu/fac/Louis_Henkin (last visited Dec. 20, 2010). Henkin was a prolific and influential scholar and was "credited with founding the study of human rights law and inspiring generations of legal scholars." *Louis Henkin: Preeminent Scholar in Constitutional and International Law*, COLUM. L. SCH., http://www.law.columbia.edu/louis_henkin/55703 (last visited Dec. 20, 2010).

74. HENKIN, *supra* note 49, at 2; *see also id.* at 2–5 (providing a more detailed discussion of this conception of human rights); Rubin, *supra* note 14, at 9 ("[H]uman rights must be rights, that is, they must represent some claim or entitlement that can be asserted by the human beings in question.").

75. See Rubin, *supra* note 14, at 9 for a discussion of the difference between human rights and legal rights.

76. *See id.* at 8–9.

77. *See* Domingo, *supra* note 47, at 1549 (quoting HANS KELSEN, PURE THEORY OF LAW 320 (Max Knight trans., 1967)).

78. For example, at the time it was passed by the United Nations General Assembly, most states did not regard the Universal Declaration of Human Rights as legally binding. MALANCZUK, *supra* note 47, at 213. However, in the time since the Universal Declaration of Human Rights was passed some of its provisions may have become binding customary international law. *Id.* (discussing the resolution passed at the United Nations Conference on Human Rights at Teheran in 1968 "proclaiming, inter alia, that 'the Universal Declaration of Human Rights . . . constitutes an obligation for the members of the international community.'" (alteration in original)).

79. *See, e.g.*, ICCPR, *supra* note 4, art. 2; ICESCR, *supra* note 4, art. 2.

80. *See* Dinah Shelton, *Protecting Human Rights in a Globalized World*, 25 B.C. INT'L & COMP. L. REV. 273, 305 (2002) (suggesting that, under the ICESCR, states will be held responsible for controlling non-state actors "over which they exercise jurisdiction").

81. *See, e.g.*, ICCPR, *supra* note 4, art. 2; ICESCR, *supra* note 4, art. 2.

forcement improbable at best and completely unrealistic at worst.[82] For example, developing countries have a disincentive to enact and enforce laws to protect human rights and remedy abuses because "developing countries compete among themselves for a limited pool of investment."[83] This environment of competition leads governments to make different choices about legislation and enforcement than might be the case if they were working in concert with other governments.[84] This situation leaves transnational corporations under-regulated.

Under-regulation of transnational corporations is problematic because it encourages decision-making that results in corporate-related human rights abuses.[85] Although transnational corporations are subject to the laws of multiple jurisdictions, the decisions that guide the acts of transnational corporations are, in practice, rarely exposed to judicial scrutiny.[86] Only a small percentage of alleged corporate-related abuses make it into court, and even fewer result in an outcome that would create incentives for corporations to change their business practices to reduce negative effects on human rights.[87]

There are at least five ways to address the problem of corporate-related human rights abuses: laws, pressure by consumers and non-governmental organizations (NGOs), self-regulation, socially responsible investment, and enforcement.[88] These can be categorized into legal and non-legal remedies. Nonetheless, in their current forms, each of these avenues is insufficient.

Legal measures include legislation and litigation. While domestic and international laws do exist, they are fragmented and are often not enforced.[89] In addition, many plaintiffs do not have the financial resources to pursue legal remedies. As a result, most cases of alleged violations never make it to court. Reaching a settlement often first requires the incentive of a pending court case. Even then, there is a possibility that

82. *See* Deva, *supra* note 2, at 3 ("[T]he approach of indirect regulation has failed to deliver the desired results."); Steven R. Ratner, *Corporations and Human Rights: A Theory of Legal Responsibility*, 111 YALE L.J. 443, 448 (2001) ("Without some international legal standards, we will likely continue to witness both excessive claims made against such actors for their responsibility and counterclaims by corporate actors against such accountability.").

83. Andrew T. Guzman, *Why LDCs Sign Treaties that Hurt Them: Explaining the Popularity of Bilateral Investment Treaties*, 38 VA. J. INT'L L. 639, 674 (1998).

84. *See id.* (suggesting that the environment of competition influences the policies adopted by developing countries).

85. *See* JOSEPH, *supra* note 1, at 153 (arguing that measures are needed to incentivize transnational corporations to adopt better practices with regard to human rights).

86. *See id.* (explaining that the human rights abuses by transnational corporations actually subjected to transnational litigation represent "only the tip of the iceberg").

87. *See id.*

88. *See id.*

89. *See id.* at 8–12.

the plaintiffs may not meet all jurisdictional, substantive, and evidentiary hurdles—which are often substantial in these types of cases.[90]

Non-legal measures can contribute to filling the gaps in law and enforcement. These measures include external incentives, such as public pressure and socially responsible investment, and internal incentives, such as self-regulation.[91] Pressure by consumers and NGOs has achieved some high profile successes.[92] However, this option is limited by the need for abuses to be sufficiently widespread or public, or both, to come to the attention of consumers and NGOs, at which point substantial harms have often already occurred.[93] Socially responsible investment offers shareholders an opportunity to avoid financing transnational corporations with questionable practices.[94] However, socially responsible investment itself does not prevent other people from investing in companies that are less socially responsible. In addition, the effectiveness of socially responsible investment remains disputed.[95]

The good works for goodwill model underlies more recent developments in scholarship such as corporate citizenship and enlightened shareholder value.[96] The essence of the good works for goodwill model is that, although good works may be more costly from a short-term perspective, they have a beneficial effect on profits from a long-term perspective.[97] The good works for goodwill model harkens back to issues raised in cases like *Dodge v. Ford Motor Co.*,[98] *A.P. Smith Manufactur-*

90. For a discussion of some of these hurdles, see Cynthia A. Williams, *Corporate Social Responsibility in an Era of Economic Globalization*, 35 U.C. DAVIS L. REV. 705, 764–71 (2002).

91. JOSEPH, *supra* note 1, at 153.

92. *See* Halpern, *supra* note 30, at 135 ("Currently, the most acute pressure felt by TNCs to modify their behavior results from concerted NGO and consumer action campaign activity.").

93. For example, in the Nike case, the underpayment of the workers in Indonesia "became publicized through the skillful use of media by several NGOs" and reported in "The New Republic, Rolling Stone, The New York Times, Foreign Affairs, and The Economist." Locke, *supra* note 22, at 10–11. However, this practice had gone on for several years before it was brought to the attention of the public. *See id.* (noting that the practice was taking place in the early 1990s and was not discontinued until the mid-1990s). For a discussion of other examples of successful applications of pressure by consumers and NGOs, see *id.* at 18–19.

94. *See* Michael S. Knoll, *Ethical Screening in Modern Financial Markets: The Conflicting Claims Underlying Socially Responsible Investment*, 57 BUS. LAW. 681, 690 (2002).

95. *See* PAUL HAWKEN & THE NATURAL CAPITAL INSTITUTE, SOCIALLY RESPONSIBLE INVESTING: HOW THE SRI INDUSTRY HAS FAILED TO RESPOND TO PEOPLE WHO WANT TO INVEST WITH CONSCIENCE AND WHAT CAN BE DONE TO CHANGE IT 16 (2004) [hereinafter HAWKEN & NCI], *available at* http://www.community-wealth.org/_pdfs/articles-publications/sri/report-harkin.pdf.

96. *See generally* DAVID LOGAN, DELWIN ROY & LAURIE REGELBRUGGE, GLOBAL CORPORATE CITIZENSHIP: RATIONALE AND STRATEGIES (1997) (discussing Global Corporate Citizenship); David Millon, *Enlightened Shareholder Value, Social Responsibility, and the Redefinition of Corporate Purpose Without Law* (Wash. & Lee Pub. Legal Studies Research Paper Series, Paper No. 2010-11, 2010), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1625750 (discussing enlightened shareholder value).

97. *See* Millon, *supra* note 96, at 1–2.

98. 170 N.W. 668, 683–84 (Mich. 1919) (examining when shareholders challenged the authority of the board of directors to prioritize philanthropic contributions over profit maximization).

ing Co. v. Barlow,[99] and *Shlensky v. Wrigley*[100] with which U.S. law students and corporate law scholars alike are familiar.[101] The good works for goodwill model essentially shifts the framework from a short-term to a long-term emphasis but retains a focus on profit-maximization as a primary goal.[102]

Finally, the level of judicial review of business decisions in the United States has global implications because approximately one-fifth of the top 100 non-financial transnational corporations are located in the United States. Currently, the main forum for corporate-related human rights abuse cases is also the United States.[103] U.S. courts subscribe to a doctrine of minimal judicial review of the substance of business decisions called the business judgment rule.[104] This rule presents a particularly problematic hurdle for enforcement in cases of corporate-related human rights abuses because it means that they, too, are subjected to only minimal judicial review.

B. State Resistance, Impotency, and Complicity

To date, international human rights law has had a limited trickle-down effect on corporate-related human rights abuses. Not all countries have signed on to all international human rights agreements. For example, the United States has ratified only seventeen of fifty-one international human rights agreements, and five of the agreements ratified by the United States relate to terrorism.[105] In addition, ratification of an in-

99. 98 A.2d 581, 585–87 (N.J. 1953) (discussing shareholders who challenged the authority of the board of directors to make a contribution to a university).

100. 237 N.E.2d 776, 778 (Ill. App. Ct. 1968) (discussing shareholders that challenged the authority of the board of directors to act "for a reason or reasons contrary and wholly unrelated to the business interests of the corporation").

101. It is worth noting here that Geoffrey P. Miller argues that although the narratives of these situations are often told as charitable corporate giving cases, these narratives did not accurately reflect the underlying motives and issues in these cases. *See generally* Geoffrey P. Miller, *Narrative and Truth in Judicial Opinions: Corporate Charitable Giving Cases* (N.Y. Univ. Sch. of Law Pub. Law & Legal Theory Research Paper Series, Working Paper No. 09-56, 2009), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1495069.

102. *See generally* Millon, *supra* note 96, at 4 (discussing corporate purpose and the shift from short-term shareholder value to be broadened to encompass nonshareholder interest as well).

103. JOSEPH, *supra* note 1, at 21 ("Most of the transnational human rights cases against corporations have arisen under the Alien Tort Claims Act [ATCA] in the United States."). However, the ability to bring claims against corporations under the Alien Tort Statute was recently called into question in Kiobel v. Royal Dutch Petroleum Co., 621 F.3d 111, 149 (2d Cir. 2010) (holding that (1) customary international law governs the scope of Alien Tort Statute liability, (2) in matters of first impression, the Alien Tort Statute does not confer jurisdiction over claims against corporations, and (3) corporate defendants were not subject to Alien Tort Statue liability because corporations were not subject to liability under customary international law).

104. The business judgment rule is based on the presumption that directors possess more expertise than judges when it comes to making business decisions and so should not be second-guessed by judges as long as appropriate procedures have been followed in the decision-making process. *See, e.g.*, *In re* Walt Disney Co. Derivative Litig., 731 A.2d 342, 361–62 (Del.Ch. 2000); Aronson v. Lewis, 473 A.2d 805, 812–13 (Del. 1984); Zapata Corp. v. Maldonado, 430 A.2d 779, 782 (Del. 1981).

105. *See Ratification of International Human Rights Treaties – USA*, U. MINN. HUM. RTS. LIBR., http://www1.umn.edu/humanrts/research/ratification-USA.html (last visited Oct. 29, 2010).

ternational human rights agreement does not automatically ensure that a country will enact comprehensive laws to implement the agreements to which it is a party.[106]

Further, even if enacted, such laws may not achieve the goals of the agreement under which they were mandated. For example, the United States ratified the International Convention on the Elimination of All Forms of Racial Discrimination (CERD) in 1994.[107] However, in its October 2000 report on efforts made to implement the CERD, the U.S. government admitted that enactment of laws has been insufficient to fulfill the goals of the CERD.[108]

There are numerous factors that affect states' ability to protect human rights.[109] These include historical, economic, and institutional constraints as well as weak legal systems, regime change, wars, politics, and geopolitical power plays.[110] Some states—for example, failed states (states that have collapsed and can no longer perform basic functions)—are unable to protect human rights.[111] In some cases, developing countries do not have the political will or legal processes to protect human rights.[112]

In other cases, states actively resist the development, implementation, and enforcement of the international human rights regime.[113] Even at the inception of the modern international human rights regime, politi-

106. *See* Mazower, *supra* note 51, at 379 ("To many people, [human rights] are honoured in the breach, an ideal which statesmen pay lip-service to but flout in practice.").

107. Comm. on the Elimination of Racial Discrimination (CERD), *Reports Submitted by States Paties Under Article 9 of the Convention: Third Periodic Reports of States Due in 1999*, ¶ 3, U.N. Doc. CERD/C/351/Add.1 (Oct. 10, 2000) [hereinafter United States 2000 CERD Report], *available at* http://www.state.gov/documents/organization/100294.pdf (combining the first three periodic reports of the United States). It may seem surprising that the United States has ratified CERD since the United States has not ratified the Convention on the Elimination of All Forms of Discrimination Against Women and the Convention on the Rights of the Child. *See Ratification of International Human Rights Treaties – USA*, U. MINN. HUM. RTS. LIBR., http://www1.umn.edu/humanrts/research/ratification-USA.html (last visited Oct. 29, 2010). It is worth noting that this exception—U.S. ratification of an international human rights agreement—was prompted by the very public presentation of petition to the United Nations in 1947 by W.E.B. Du Bois on behalf of the National Association for the Advancement of Colored People. Mazower, *supra* note 51, at 395.

108. United States 2000 CERD Report, *supra* note 108, at 18–21. Reasons given in the report for the inadequate implementation of the CERD include inadequate enforcement of existing laws, inefficient use of data, economic disparities, and lack of access to educational opportunities, technology, and high technology skills. *Id.* at 19–20. In addition, the report notes that persistent discrimination is a primary factor affecting the implementation of laws designed to eliminate discrimination. *Id.* at 19–21. Ironic. However, it is worth noting that the section on factors affecting implementation has been reduced to less than a page in the 2007 report and no longer gives discrimination as a reason for persistent discrimination despite the enactment of laws. *See* Periodic Report of the U.S., U.N. Comm. on the Elimination of Racial Discrimination, *Periodic Report of the U.S.*, ¶¶ 52–54 (April 2007), *available at* http://www.state.gov/documents/organization/83517.pdf.

109. *See* Randall Peerenboom, *Human Rights and Rule of Law: What's the Relationship?*, 36 GEO. J. INT'L L. 809, 943–44 (2005).

110. *See id.*

111. *See id.* at 858.

112. JOSEPH, *supra* note 1, at 11.

113. *See* Mazower, *supra* note 51, at 393.

cal decision-makers attempted to limit the scope, implementation, and enforcement of international human rights law.[114] For example, during the establishment of the United Nations in the 1940s, the U.S. Congress actively took steps to protect U.S. domestic jurisdiction and prevent direct applicability of the U.N. Charter's human rights provisions.[115]

As a result of U.S. and British concerns about national jurisdiction, a domestic jurisdiction clause was incorporated into the U.N. Charter.[116] Article 2 of the U.N. Charter states that "[n]othing contained in the present Charter shall authorize the United Nations to intervene in matters which are essentially within the domestic jurisdiction of any state or shall require the Members to submit such matters to settlement under the present Charter."[117] The inclusion of this domestic jurisdiction clause effectively circumscribed the ability of the U.N. Charter to serve as a direct basis for effective implementation and enforcement of the emerging post-World War II international human rights regime.[118]

The state-centered focus of international human rights law relegates to the periphery complex relationships between public and private actors—as well as relationships between private actors, and individuals and communities.[119] For example, the real and perceived need for foreign capital creates incentives for states to establish an investment climate that is attractive to transnational corporations.[120] In response, states that are looking to foreign direct investment to provide capital and know-how for infrastructure development may choose not to monitor or enforce laws and regulations that would increase the cost of operating within their borders.[121]

In some cases, states are also complicit in corporate-related human rights abuses[122] and, therefore, have a disincentive to implement and enforce protection from and remedies for human rights abuses. For example, in response to the Niger Delta disasters discussed above, Ken Saro-Wiwa, a businessman, novelist, television producer, and former government official, spoke out against land appropriation, pollution, and other alleged negative effects of the acts of multinational oil companies, including Royal Dutch/Shell.[123] In November 1995, Ken Saro-Wiwa and

114. *See id.* ("The higher human rights moved up the agenda, the greater the pressure for a *further* limitation on the new [United Nation]'s ability to intervene in the domestic affairs of member states." (emphasis added)).

115. *See id.*

116. *See id.* at 393.

117. U.N. Charter art. 2, para. 7.

118. *See* Mazower, *supra* note 51, at 393 (discussing how Article 2 undermined the scope and effectiveness of the U.N. Charter's ability to serve as the basis for an effective human rights regime).

119. *See* Hope Lewis, *Transnational Dimensions of Racial Identity: Reflecting on Race, the Global Economy, and the Human Rights Movement at 60*, 24 MD. J. INT'L L. 296, 306 (2009).

120. *See id.* at 306–07.

121. *See id.*

122. *See generally* Wiwa Fifth Amended Complaint, *supra* note 16.

123. *See id.* at 2.

five other men were beaten, and denied food, water and bedding for several days.[124] On November 10, 1995, the six men were hanged after a "special military trial" based on fabricated charges.[125] The Royal Dutch Shell consortium had knowledge of these acts and gave its consent and support.[126] Royal Dutch/Shell made payments to the military police, contracted for the purchase of weapons supplied to the Nigerian police and military, exchanged information with and provided logistical support to the Nigerian police and military, and participated in the planning and coordination of raids and terror campaigns in Ogoni and the Niger Delta.[127] At one point, the Nigerian subsidiary, a member of the Royal Dutch Shell consortium, publicly praised the working relationship between the subsidiary and the Nigerian Police Force.[128]

Even in cases in which states are not acting in collusion with transnational corporations, the state-centered human rights regime does not provide sufficient incentives for transnational corporations to respect human rights. In the 2007 U.N. study of cases of alleged corporate-related human rights abuses discussed above, two of the main contexts in which transnational corporations impinged upon human rights were labor and the environment.[129] In most of the cases reviewed in the U.N. study, each alleged instance of corporate-related human rights abuse affected at least one hundred people, and the numbers were often much higher.[130] For example, in one case, one alleged instance of corporate-related human rights abuse affected approximately 60,000 people.[131]

Although the international human rights regime has made progress since the United Nations General Assembly adopted the 1948 Universal Declaration of Human Rights, state resistance continues to inhibit development of an implementable and enforceable human rights regime domestically and internationally.[132] States, which bear the responsibility for implementing and enforcing international human rights law, fail to ade-

---

124. *See id.* at 3, 17.

125. *Id.* ("Defendants Royal Dutch/Shell, together with the military regime governing Nigeria, acting through the Shell Petroleum Development Company of Nigeria Limited . . . , and acting with other agents and co-conspirators have, in the past and continuing through the present, used force and intimidation to silence any opposition to their activities in Nigeria which includes the exploitation of the petroleum resources of the Delta and spoliation of the environment there.").

126. *Id.* at 2 ("The executions of Ken Saro-Wiwa, John Kpuinen, Saturday Doobee, Felix Nuate, Daniel Gbokoo, and Dr. Barinem Kiobel and the imprisonment and torture of Michael Tema Vizor by the Nigerian military junta and the campaign to falsely accuse them were carried out with the knowledge, consent, and/or support of Defendants Royal Dutch Petroleum Company and Shell Transport and Trading Company, p.l.c., . . . and their agents and officers, as part of a pattern of collaboration and/or conspiracy between [Royal Dutch/Shell] and the military junta of Nigeria to violently and ruthlessly suppress any opposition to Royal Dutch/Shell's conduct in its exploitation of oil and natural gas resources in Ogoni and in the Niger Delta.").

127. *See id.* at 8–9.

128. *Id.* at 13.

129. *Protect, Respect & Remedy*, *supra* note 13, at 2.

130. *Id.* at 13–14.

131. *Id.* at 13.

132. *See* Peerenboom, *supra* note 109, at 824.

quately do so. At the same time, the focus on international human rights law diverts attention and energy that could be used to develop alternative areas of human rights law. State resistance takes many forms including non-ratification, reservations to international agreements, non-compliance, and even withdrawal from international agreements.[133] The extent to which states make reservations to international human rights treaties undermines the effectiveness of those treaties.[134]

C. Extraterritorial Application of Domestic Law

The combination of sovereignty and territorial integrity imbues states with jurisdiction within a state's territories.[135] Generally, under modern international law, sovereign states have the jurisdiction to make and enforce laws within their own territories.[136] This also means that, generally, states may not interfere in affairs that are within the jurisdiction of another state or in another state's exercise of its authority within its jurisdiction.[137]

States also exercise extraterritorial jurisdiction.[138] In fact, there is an increasing trend toward extraterritoriality.[139] Familiar examples include exterritorial application of antitrust, securities, and merger and takeover laws.[140] However, the extraterritorial application of domestic laws is problematic.[141] For example, the checks and balances that develop out of mutual obligations, such as those embodied in international treaties, do not constrain extraterritorial application of domestic law.[142] Individual nations are able to use their domestic laws to influence international policy in a non-transparent manner.[143] Such laws are not first subjected to

---

133. *Id.*

134. *See id.* However, ratification alone is not inherently correlated with increased protections for human rights.

135. For a discussion of sovereignty, territoriality, and jurisdiction, see Domingo, *supra* note 47, at 1556–76.

136. *See* PHILLIP I. BLUMBERG, THE MULTINATIONAL CHALLENGE TO CORPORATION LAW: THE SEARCH FOR A NEW CORPORATE PERSONALITY 175 (1993).

137. U.N. Charter art. 2, para. 7.

138. *See* BLUMBERG, *supra* note 136, at 175–77 (discussing the bases for extraterritoriality). Blumberg differentiates between home and host country territoriality. *Id.* at 177–91. For examples of home and host country extraterritoriality, see generally *id.* For a brief history of extraterritorial jurisdiction in the international criminal context, see Christopher L. Blakesley & Dan E. Stigall, *The Myopia of* U.S. v. Martinelli*: Extraterritorial Jurisdiction in the 21st Century*, 39 GEO. WASH. INT'L L. REV. 1, 3–11 (2007).

139. Austen L. Parrish, *Reclaiming International Law from Extraterritoriality*, 93 MINN. L. REV. 815, 818 (2009). For a discussion of U.S. attempts to regulate transnational corporations, see Mark B. Baker, *Tightening the Toothless Vise: Codes of Conduct and the American Multinational Enterprise*, 20 WIS. INT'L L.J. 89, 107–18 (2001).

140. BLUMBERG, *supra* note 136, at 192–93.

141. *See* Parrish, *supra* note 139, at 820.

142. *Id.* at 846. For a response to arguments that extraterritorial laws encourage international lawmaking, see *id.* at 871–72.

143. *See id.* at 846, 857, 860, 862.

public debate and discourse at the global level, which poses a threat to democratic sovereignty and to international cooperation and stability.[144]

The extraterritorial application of domestic law also tends to make enforcement of human rights dependent on private litigation.[145] On its face, private litigation has certain benefits because it affords potential plaintiffs an opportunity to pursue a remedy for their harms. However, as mentioned above, many plaintiffs do not have the significant financial resources necessary to pursue remedies through private litigation. Further, even if private litigants have the resources to successfully pursue litigation against a corporation, there is no guarantee that those plaintiffs would have the financial resources to enforce the judgment. After all, large transnational corporations have comparatively infinite time and resources with which to oppose litigation and the enforcement of judgments.

In addition, "[e]xtraterritorial laws undermine international law . . . ."[146] Domestic litigation can divert energy toward short-fixes with limited scope that could be better concentrated on working toward comprehensive laws and enforcement with long-term effects.[147] For example, when decisions are not recognized and enforced abroad, battles won in one domestic court must then be re-litigated over and over again in all of the jurisdictions around the world to achieve truly international protections and remedies.

The state-centered human rights regime is characterized by an emphasis on domestic jurisdiction.[148] Thus, laws governing the protection and enforcement of human rights are at the whim of the prevailing political will or lack thereof in any given country at any given time. Arguably, it might be possible to fill gaps in the human rights protections and enforcement of one country with the extraterritorial application of the domestic laws of another country. However, as this Part I demonstrated, such extraterritorial application of domestic law is problematic and can undermine the international human rights law regime.[149]

This Part I highlighted the weaknesses of the state-centered international human rights regime and its inability to respond to the realities of

144. *See id.*

145. *See id.* at 862–63.

146. *Id.* at 865. However, this is not to suggest that the territorial jurisdiction of a state is holy and should be protected from any and all encroachments. *See* BLUMBERG, *supra* note 136, at 201 ("[T]he extraterritorial assertion of national law inherent in the application of enterprise principles to components of multinational groups inevitably will engender international confrontation and disrupt international trade and relations."); Domingo, *supra* note 47, at 1575–76 (arguing that universal harms should be resolved in a universal way).

147. *See* Parrish, *supra* note 139, at 865–66.

148. U.N. Charter art. 2, para. 7.

149. *See* Parrish, *supra* note 139, at 865–66.

modern transnational corporations.[150] State resistance, impotency, and complicity lead to under-protection of human rights.[151] Extraterritorial application of domestic laws is insufficient to fill these gaps and can have far-reaching detrimental effects.[152] Although the weaknesses of the existing regime are well-known and well-documented, the question of how to comprehensively address these issues remains unanswered.

This Article argues that this question is unanswered largely because the conflation of human rights law and international human rights law has inhibited the evolution of human rights law in other directions. The development of new forms of human rights law could address the under-protection of human rights that results from an over-reliance on international human rights law. One issue that new forms of human rights law could address is corporate-related human rights abuse. Part II argues that human rights law should be extended directly to transnational corporations because the intersection of human rights and transnational corporate activity is special and, therefore, requires a specialized regulatory regime.

II. TRANSNATIONAL CORPORATIONS NEED DEDICATED REGULATION

Alternative forms of human rights law are needed because international human rights law is inadequate to enact and enforce comprehensive protections from corporate-related human rights abuses. The intersection of transnational corporations and human rights presents particular challenges which existing legal regimes are ill-equipped to regulate. One reason for this is the combination of the nature of transnational operations and the importance of human rights. Many transnational corporations have achieved a level of wealth and influence that strain the regulatory competence of nation-states.[153] Transnational corporations operate and cause or contribute to harms in multiple jurisdictions, which challenges the effectiveness of traditional territorial jurisdiction. Finally, human rights represent a special set of societal values that require a level of regulation that may exceed those deemed sufficient in other areas of the law. This Part II analyzes each of these characteristics of corporate-related effects on human rights in turn.

150. *See* Erika R. George, *The Place of the Private Transnational Actor in International Law: Human Rights Norms, Development Aims, and Understanding Corporate Self-Regulation as Soft Law*, 101 AM. SOC'Y INT'L L. PROC. 473, 474 (2007).

151. *See* Peerenboom, *supra* note 109, at 943–44.

152. *See* Parrish, *supra* note 139, at 866.

153. *See generally* JOEL BALKAN, THE CORPORATION: THE PATHOLOGICAL PURSUIT OF PROFIT AND POWER 3 (2004) (examining the strength and wealth of transnational corporations).

A. Wealth and Influence of Transnational Corporations

Transnational corporations have achieved vast wealth and immense influence.[154] They exercise economic, political, and legal influence in home countries, the country of incorporation, and host countries—the country in which assets or operations are located.[155] This exercise of influence inhibits the ability of states to prevent and redress corporate-related human rights abuses.[156]

1. Economic Might of Transnational Corporations

Foreign direct investment, which takes the form of owning, operating, or managing a business in host countries, is an important component of transnational corporations' operations.[157] States allow foreign direct investment by transnational corporations within their borders for a variety of reasons. Two of the main reasons that host countries allow—and more importantly *encourage*—foreign direct investment are to gain infusions of capital and know-how in the short term and to achieve economic prosperity in the long term.[158]

In 2007, there were approximately 77,000 transnational corporations.[159] The foreign assets of the top twenty transnational corporations amounted to over $2,904,995 million.[160] The majority of the largest transnational corporations are incorporated in more developed countries.[161] However, at least since the European occupation of Africa and other former colonies, transnational corporations have chosen to own, operate, and manage business operations in countries in which they are not incorporated, and they continue to do so today in increasing numbers.[162] In some cases, the wealth of a transnational corporation may ex-

---

154. *See generally* United Nations Conference on Trade and Development, *The Universe of the Largest Transnational Corporations*, U.N. Doc. UNCTAD/ITE/IIA/2007/2 (2007) [hereinafter UNCTAD], *available at* http://www.unctad.org/en/docs/iteiia20072_en.pdf (compiling a list of the 100 largest non-financial transnational corporations and the top fifty transnational corporations from developing countries).

155. Jernej Letnar Cernic, *Corporate Human Rights Obligations under Stabilization Clauses*, 11 GER. L.J. 210, 218 (2010); *see also* Citizens United v. FEC, 130 S. Ct. 876, 962–63 (2010) (Stevens, J., concurring in part and dissenting in part) (discussing the ability of corporations to exert political influence through political speech and election contributions).

156. Cernic, supra note 155, at 211.

157. UNCTAD, supra note 154, at 3 (stating that 770,000 foreign affiliates of transnational corporations have globally generated an estimated $4.5 trillion in value since the early 1990's).

158. Padma Mallampally & Karl P. Sauvant, *Foreign Direct Investment in Developing Countries*, FINANCE & DEVELOPMENT, Mar. 1999, at 34–36, *available at* http://www.imf.org/external/pubs/ft/fandd/1999/03/pdf/mallampa.pdf.

159. UNCTAD, *supra* note138, at 3.

160. *Id.* at 40.

161. *Id.* at 4. The European Union, Japan, and the United States are the home countries for eighty-five percent of the top one hundred transnational corporations. *Id.* The number of transnational corporations incorporated in less developed economies is growing but only a handful have entered the ranks of the world's largest. *Id.* at 15–16.

162. *See* Cernic, *supra* note 155, at 211.

ceed the wealth of the host country itself.[163] The wealth of transnational corporations gives them room to influence the abuse or enjoyment of human rights by individuals and communities in both home and host countries.[164]

2. The Siren Song of Foreign Direct Investment

Foreign direct investment by transnational corporations is considered to be a significant source of private external financing for developing countries.[165] In 2008, approximately forty-three percent of all inward foreign direct investment flowed into transitional and developing countries.[166] Foreign direct investment is also important to developing countries as an opportunity for innovation, technology transfer, human capacity development, and access to various forms of corporate governance.[167] The need (real or perceived) for foreign capital makes many host countries dependent on it for development and, therefore, susceptible to the influence of transnational corporations.[168]

In addition to potential economic benefits, foreign direct investment by transnational corporations also exerts socio-cultural, political, and legal influence in host countries.[169] In some countries, transnational corporations may reinforce gender hierarchies by paying, or allowing their subsidiaries to pay, women less than men for the same work.[170] The effects of foreign direct investment may also be indirect and unexpected. Between 1989 and 1992 in Papua New Guinea, for example, increased income resulting from foreign direct investment was linked to increased

---

163. Halpern, *supra* note 30, at 144 ("[T]he process of economic globalization has allowed many TNCs to accumulate vast sums of resources and power, often times in excess of the host state's own."); *see also* Baker, *supra* note 139, at 94.

164. JOSEPH, *supra* note 1, at 1–2.

165. Mallampally & Sauvant, *supra* note 158, at 35; *see also*, United Nations Conference on Trade and Development, *World Investment Directory: Volume X Africa 2008*, at 1, U.N. Doc. UNCTAD/ITE/IIT/2007/5 (2008), *available at* http://www.unctad.org/en/docs/iteiit20075_en.pdf ("The ratio of FDI inflows to the region's gross fixed capital formation . . . [was] 20 per cent in 2006.").

166. *See Inward & Outward Foreign Direct Investment Flows*, UNCTADSTAT, http://unctadstat.unctad.org/TableViewer/tableView.aspx?ReportId=88 (last visited Oct. 29, 2010) (calculated using 2008 data available on UNCTADSTAT).

167. Mallampally & Sauvant, *supra* note 158, at 35–36.

168. *See id.* at 36.

169. Baker, *supra* note 139, at 89 ("[Transnational corporations are] potentially more economically powerful than Stalin's Soviet Union, and with more broad-based political influence than The Third Reich.").

170. *See* U.N. Secretary-General, *The World's Women 2005: Progress in* Statistics, 54, U.N. Doc. ST/ESA/STAT/SER.K/17 (2006), *available at* http://unstats.un.org/unsd/demographic/products/indwm/ww2005_pub/English/WW2005_text_complete_BW.pdf. For a discussion of the gendered effects of foreign direct investment, see generally Rachel J. Anderson, *Foreign Direct Investment and Distributional Equity for Women*, 31 WOMEN'S RTS. L. REP. (forthcoming 2010). See also Darren Rosenblum, *Feminizing Capital: A Corporate Imperative*, 6 BERKELEY BUS. L.J. 55, 74–80 (2009) for a more comprehensive discussion of women's work and DOUGLAS BRANSON, NO SEAT AT THE TABLE: HOW CORPORATE GOVERNANCE AND LAW KEEP WOMEN OUT OF THE BOARDROOM (2006) for a discussion of implicit male bias in corporate governance and the effects of this bias on women's rights.

consumption, which was linked to increased rates of polygamy.[171] Men with more income were able to take more wives.[172]

3. Political Influence of Transnational Corporations

Transnational corporations exert political influence in both home and host countries. This influence may be direct or indirect or both. In home countries, for example, transnational corporations exert political influence through lobbying and campaign contributions.[173] When a transnational corporation exerts political influence, it affects government decision making in a way that encourages government action or inaction.[174]

Transnational corporations may also use their political influence in their home country to influence decisions in host countries. For example, transnational corporations incorporated in countries like the United States pressure their home governments to pressure governments in developing countries to provide protections for the corporations' intellectual property.[175]

4. Legal Influence of Transnational Corporations

Transnational corporations are often able to influence the application and enforcement of laws in the countries in which they operate. For example, Nike's Korean suppliers operating shoe factories in Indonesia in the early 1990s successfully petitioned for and received an exemption from the minimum wage.[176] The companies claimed that paying mini-

171. *See* Glenn Banks, *Globalization, Poverty, and Hyperdevelopment in Papua New Guinea's Mining Sector*, 46 FOCCAL–EUR. J. ANTHROPOLOGY 128, 135 (2005).

172. *See id.*

173. *See* Citizens United v. FEC, 130 S. Ct. 876, 881–86 (2010). This recent U.S. Supreme Court case ensured the ability of corporations, whether domestic and foreign, to exert political influence through political speech and election contributions. *Id.*; *cf.* 2 U.S.C. § 441e(a)(1) (2006) (prohibiting foreign nationals from directly or indirectly making contributions or independent expenditures in connection with a U.S. election). However, the majority opinion does not address whether the "[g]overnment has a compelling interest in preventing foreign individuals or associations from influencing" the domestic political process. *Citizens United*, 130 S. Ct. at 911. Further, the majority does not consider the issues in this case in the context of transnational corporations. *Id.* at 936 (Stevens, J., concurring in part and dissenting in part).

174. *Id.* at 961–63.

175. *See* Peter Straub, *Farmers in the IP Wrench—How Patents on Gene-Modified Crops Violate the Right to Food in Developing Countries*, 29 HASTINGS INT'L & COMP. L. REV. 187, 193 (2006). For example, U.S. companies have tried to "'invent restrictions that do not exist within international law and compel developing countries to accept them through U.S. trade pressures.'" *Special 301 Review Public Hearing: Hearing Before the Special 301 Subcommittee of the Office of the U.S. Trade Representative*, at 104 (2010) (from testimony given by a representative from Doctors Without Borders) *available at* http://www.ustr.gov/webfm_send/1726. . Strong intellectual property rights can have detrimental effects on human health and life. *Id.* at 97–98 ("People in developing countries are dying because medicines do not exist due to inadequate incentives for their development or because they're unavailable due in part to patent barriers and high costs."). For a more comprehensive discussion of the intersection of trade and human rights, see BERTA HERNANDEZ-TRUYOL & STEPHEN J. POWELL, JUST TRADE: A NEW COVENANT LINKING TRADE AND HUMAN RIGHTS (2009).

176. Locke, *supra* note 22, at 10.

mum wage would be a hardship for them.[177] The local minimum wage at the time was approximately $1 per day (2,100 rupiah) and covered only 70% of the basic needs of one individual.[178] This arrangement allowed factories producing Nike products to underpay more than 25,000 workers without violating the letter of the law in Indonesia—the host country.[179] This practice was discontinued on Nike's request in the mid 1990s—an example of transnational corporations' ability to prevent human rights abuses by their subsidiaries and suppliers, if they choose to do so.[180]

One important way that transnational corporations affect the legal regimes in host countries is through investment agreements between the state and one or more investors. Investor-state investment agreements are agreements between host states and investors to set rules, standards, and even determine the laws that will apply to the transaction that is the object of the contract.[181] Investor-state investment agreements are problematic for the protection of human rights because they insulate "projects from standards of protection of basic rights that apply elsewhere in a host country; and [they] shrink[] certain remedies that victims would otherwise have."[182] Investment agreements between states and transnational corporations are common, for example, in the extraction sector.[183]

Investor-state investment agreements highlight a key problem associated with transaction-based law and policymaking that takes place outside of a comprehensive legal and regulatory framework. This is signaled by the fact that investor-state investment agreements are commonly referred to as *international* investment agreements.[184] This nomenclature supports claims that some transnational corporations have amassed previously unknown influence over the law and policymaking power of these agreements. In the international investment agreement, transnational corporations become quasi-states that enter into agreements with nation-states, and they imbue the transnational corporations with rights and powers that allow them to encroach upon prerogatives of the public

177. *Id.*

178. *Id.*

179. *See id.*

180. *Id.* at 11. Nike first denied the ability to affect the practices of companies in its supply chain and then only did so after substantial bad press. *Id.* at 10–11.

181. *See* Sheldon Leader, *Human Rights, Risks, and New Strategies for Global Investment*, 9 J. INT'L ECON. L. 657, 703–04 (2006) (discussing the types of clauses in investment contracts and international investment contracts in conjunction with human rights); *see also* Wendy N. Duong, *Partnerships with Monarchs—Two Case Studies: Case One*, 25 U. PA. J. INT'L ECON. L. 1171, 1265 (2004) ("In various forms, the [Stabilization] Clause restricts the host jurisdiction's exercise of 'permanent sovereignty' by contractually preventing the nation-state from subsequently modifying the governing law of the investment contract."). Concession agreements between early transnational corporations, formerly colonial trading companies, are the predecessors of modern international investment contracts. *See* ANGHIE, *supra* note 52, at 233–34.

182. Leader, *supra* note 181, at 700.

183. *See id.* at 661, 696 (discussing the BTC pipeline connecting the Caspian and Mediterranean Seas and the Chad-Cameroon pipeline agreement).

184. This is a post-World War II development in international law. ANGHIE, *supra* note 52, at 230.

domain.[185] Taking this one step further, international investment agreements even have the potential to trump both domestic and international law.[186]

These investor-state agreements are, thus, not subject to domestic law but rather a legal system of a different kind.[187] Generally, these agreements are considered to be subject to "an international law system."[188] Thus, although the agreement is not between two nations, it is treated as *inter*-national if one of the parties is a developing country and one of the parties is a transnational corporation, but not if the country involved is a "developed" country.[189] As *international* agreements, arbitration of disputes is subject to international rather than national law.[190]

However, the shifting of transnational corporations into a quasi-state status glosses over important distinctions. Transnational corporations are not states and, although the line between public and private is a topic for debate, according them a quasi-state status contributes to the obfuscation of problematic trends related to transaction-based law and policy making. Transaction-based law and policy making is not subjected to the scrutiny accorded a more public deliberative process, and the immediate interests of the parties are narrower than would be the case, for example, if third parties with differing interests were directly involved.

States are at a greater disadvantage when entering into investment agreements than may be apparent at first blush. Initially, it might seem that the state has all the power since it is a sovereign nation and is the maker and enforcer of law and policy within its territories. However, the reason that states enter into investment agreements with investors is that they do not have the domestic capacity, whether public or private, to achieve the purpose of the agreement.[191] Thus, they need to attract one or more foreign investors, although they may not necessarily need to attract a specific foreign investor. Depending on the industry, the number of potential foreign investors may be extremely limited due to high barriers to entry. For example, only a limited number of corporations have equipment readily available to dedicate to drilling oil, and corporations that want to enter this market must have sufficient capital to lay out in

185. *See* ANGHIE, *supra* note 52, at 232.

186. *See* Cernic, *supra* note 155, at 218 (analyzing the language of the 2005 Mineral Development Agreement between the National Transitional Government of Liberia and Mittal Steel Holdings AG).

187. *See* ANGHIE, *supra* note 52, at 231.

188. *See id.*

189. *See* ANGHIE, *supra* note 52, at 231–32 (quoting Derek Bowett, *State Contracts with Aliens: Contemporary Developments on Compensation for Termination or Breach*, 59 BRITISH Y.B. INT'L L. 49, 51 (1988)).

190. Texaco Overseas Petroleum Co. v. Libyan Arab Republic, 53 I.L.R 389, 433 (Arb. Trib. 1977).

191. *See generally* Leader, *supra* note 181, at 657–694 (discussing investment contracts).

advance to purchase the expensive equipment necessary for oil drilling.[192]

Investment agreements between transnational corporations and host countries create special legal regimes of limited duration with potentially substantial effects.[193] These special legal regimes can weaken legal protections for individuals and communities in host countries.[194] For example, existing human rights standards may be affected indirectly or are limited directly by the terms of the international investment contract.[195] When this happens, the host state may effectively lose its ability to enforce otherwise applicable human rights laws and standards.[196]

Most investor-state investment agreements include stabilization clauses.[197] A stabilization clause is "contract language which freezes the provisions of a national system of law chosen as the law of the contract as of the date of the contract, in order to prevent the application to the contract of any future alterations of this system."[198] The purpose of stabilization clauses is to protect transnational corporations from unfavorable and perhaps unfair changes in legislation or regulations after they are locked into the contract and have sunk money into the project.[199] These stabilization clauses can, for example, prevent host states from applying laws that come into existence after the contract is signed to the transaction that is the object of the contract.[200]

On their face, stabilization clauses make sense because they mitigate the risk to a transnational corporation that the other party to the contract, a sovereign state with law-making power, might change the laws for its own benefit at any time. However, a broad stabilization clause may also have the effect of depriving individuals and communities in the host country of human rights protections that may become enforceable by law after the date of an investment contract. In fact, stabilization

192. *See id.* at 682–83.

193. *See id.* at 660. Although the duration is limited, it still extends for the life of the investment project, which is often as long as seventy years. *Id.*

194. *Id.*

195. *See, e.g.*, Cernic, *supra* note 155, at 220–21 (discussing the stabilization clause in the Host Government Agreements between a consortium of oil corporations and the governments of Azerbaijan, Georgia, and Turkey regarding the Baku-Tbilisi-Ceyhan pipeline project).

196. Leader, *supra* note 181, at 671.

197. Cernic, *supra* note 155, at 213. There are various forms of stabilization clauses. *Id.* at 214. *See* J. Nna Emeka, *Anchoring Stabilization Clauses in International Petroleum Contracts*, 42 INT'L LAW. 1317, 1318–20 (2008) (discussing the rationale for and the origin of stabilization clauses); *see also* Christopher T. Curtis, *The Legal Security of Economic Development Agreements*, 29 HARV. INT'L L.J. 317, 321 (1988) (describing both stabilization and governing-law clauses in an investment agreement).

198. Cernic, *supra* note 155, at 213; *see also* Curtis, *supra* note 197, at 346–47 (discussing the nature of stabilization clauses); Sang-Jick Yoon, Comment, *Critical Issues on the Foreign Investment Laws of North Korea for Foreign Investors*, 15 WIS. INT'L L.J. 325, 364–65 (1997) (describing how stabilization clauses are "grandfathered exceptions to the new law").

199. *See* Cernic, *supra* note 155, at 213–14.

200. *See* Leader, *supra* note 181, at 672–681 (discussing stabilization clauses in international investment contracts).

clauses are sometimes structured so that the entry into an international agreement to protect human rights by a host state can make the host state liable to the transnational corporation for profit lost due to the need to comply with laws protecting human rights.[201]

B. Complexities of Multijurisdictional Regulation and Enforcement

The complex structure of transnational operations often triggers questions of jurisdiction and impedes transnational enforcement.[202] It also creates a need for the development of comprehensive global regulation. Although there are jurisdictional challenges, there are also options that offer promising alternatives to address existing jurisdictional hurdles.

1. Territorial Jurisdiction

Territorial jurisdiction, also known as national jurisdiction, state jurisdiction, domestic jurisdiction, and sovereignty, is the power a state has to make laws that govern its territories and the power of its courts to exercise jurisdiction within the bounds of its territories.[203] Territorial jurisdiction is one of the oldest, most fundamental, and accepted forms of jurisdiction.[204] Article 2 of the U.N. Charter, which prohibits states from interfering in the territorial jurisdiction of other states, further protects the territorial jurisdiction of states.[205]

Territorial jurisdiction makes it possible, at least theoretically, to enact and enforce laws governing transnational corporations in the jurisdictions in which they are incorporated, operate, or have assets.[206] In part, this is because respect for state sovereignty remains a central tenet of international human rights law.[207] Thus, domestic tax, labor, business, and antitrust laws, etc., primarily govern transnational corporations. This is not to suggest that these laws do not sometimes have extraterritorial

---

201. *See* Leader, *supra* note 181, at 673–75.

202. Jurisdiction is the power to regulate and to enforce. *See generally* GARY B. BORN, INTERNATIONAL COMMERCIAL ARBITRATION (2009); Gary B. Born, *Reflections on Judicial Jurisdiction in International Cases*, 17 GA. J. INT'L & COMP. L. 1, 7–25 (1987) (discussing judicial jurisdiction in the international context).

203. Anne-Marie Slaughter, *Defining the Limits: Universal Jurisdiction and National Courts*, *in* UNIVERSAL JURISDICTION 168, 171–72 (Stephen Macedo, ed., 2004); *see also* M. Cherif Bassiouni, *The History of Universal Jurisdiction*, *in* UNIVERSAL JURISDICTION 39, 40 (Stephen Macedo ed., 2004); *cf.* Blakesley & Stigall, *supra* note 138, at 12 (discussing two types of jurisdiction: "rule-making and rule-enforcing jurisdiction," which take three forms: "prescriptive or legislative, adjudicative, and enforcement").

204. *See* Bassiouni, *supra* note 203, at 40.

205. U.N. Charter art. 2, para. 4.

206. *See* JESSUP, *supra* note 32, at 74–76 (discussing territorial jurisdiction governing transnational corporations and the past resentments between the judicial authorities of the United States and Canadian companies).

207. *See* Parrish, *supra* note 139, at 824 (discussing one reason for this central tenant-fears of some scholars, theorists, and policy makers that international law has more democratic legitimacy than domestic law). See also MALANCZUK, *supra* note 47, at 17–18 and Domingo, *supra* note 47, at 1556–65 for a discussion of the theory, doctrine, and concepts of sovereignty.

reach, but rather that—in most cases—they are dependent on domestic enforcement.[208] However, in practice, territorial jurisdiction has proven to be a limited tool. Transnational corporations often operate in the context of a multinational enterprise, which includes multiple entities incorporated in multiple jurisdictions working with various partners and suppliers.[209]

In the case of lawmaking to protect human rights, territorial jurisdiction generally determines which state and law-making bodies have prescriptive jurisdiction, also known as legislative jurisdiction.[210] In addition, as discussed above, prescriptive jurisdiction may determine the application of international law and extraterritorial application of domestic law. For example, bilateral investment treaties are forms of international lawmaking that govern foreign direct investment.[211] However, domestic and international laws, when they exist, are often fragmented and not enforced. Recall the minimum wage exceptions granted by the Indonesian government to Nike suppliers.[212]

When corporate-related human rights abuses occur in the context of a multinational operation, multiple states may have the right to exercise jurisdiction. Domestic conflict of laws rules can be used to determine which of the states that have the right to exercise jurisdiction should exercise that right in a particular case.[213] Although this may seem to offer plaintiffs multiple jurisdictions and thus multiple opportunities to pursue redress for their claims, in practice this can work to the disadvantage of potential plaintiffs. For example, the courts in the home country of a defendant transnational corporation may—based on their own national conflict of law rules—decide not to assert jurisdiction under the principle of forum non conveniens. In dismissing a case for forum non conveniens, the court finds that another forum, the host country for example, would be a better suited—more convenient—alternative forum. At the same time, the plaintiffs claims may be blocked for political or other reasons in the host country. Therefore, if such a claim is dismissed, the effect in practice is often that potential plaintiffs are left without a forum in which to bring their claims. This is particularly true for claims dismissed in the United States since it is one of the primary—and only—forums in which

208. See Parrish, *supra* note 139, at 820, 846 for a discussion of the rise of extraterritorial domestic law and associated problems.

209. *See* Anderson, *supra* note 34, at 2.

210. *See* Michael D. Ramsey, *International Law Limits on Investor Liability in Human Rights Litigation*, 50 HARV. INT'L L.J. 271, 271–72 (2009) (discussing the international law implications for foreign investors responsible for human rights abuses in developing nations in which they do business).

211. *See generally* Leader, *supra* note 165, at 660–70.

212. Locke, *supra* note 22, at 10–11.

213. *See* JESSUP, *supra* note 32, at 35–39 (discussing the difference between international law and conflicts of law and which one holds the power in a particular problem).

is possible to bring claims for corporate-related human rights violations.[214]

2. Other Forms of State Jurisdiction

In addition to territorial jurisdiction, there are numerous other types of jurisdiction in the transnational context. The least controversial form of jurisdiction is where there is a tie between the prescribing or enforcing state and the party who is the object of this prescription or enforcement. The traditional bases of prescriptive jurisdiction, the power to make rules and regulations, are territory and citizenship.[215] As discussed above, countries commonly make laws that govern their citizens and people within their territory.[216] These are also generally accepted as a matter of customary international law.[217]

Forms of judicial jurisdiction, the power of a court to exercise jurisdiction, include personal jurisdiction and jurisdiction based on effects, consent, nationality, and property.[218] However, the recognized forms of judicial jurisdiction are not uniform and vary from country to country. For example, the United States recognizes tag or transient jurisdiction, whereby the courts may assert personal jurisdiction even if the defendant is passing through their jurisdiction on a wholly unrelated matter.[219] French courts will generally grant jurisdiction in any case that involves a French national on either side.[220] German courts grant jurisdiction in cases involving property located in Germany whether or not the owners have any other ties or connections to Germany.[221]

3. Universal Jurisdiction

In addition to the various forms of jurisdiction recognized by different countries, there is also jurisdiction that stems from the nature of the conduct. Universal jurisdiction allows courts to exercise jurisdiction based on the type of conduct alone without any other connection to the state that is exercising jurisdiction.[222] Often, people think only about

214. *See* Anderson, *supra* note 34, at 4–5.

215. Paul B. Stephan, *Symmetry and Selectivity: What Happens in International Law When the World Changes*, 10 CHI. J. INT'L L. 91, 105 (2009).

216. *See* Ramsey, *supra* note 210, at 284.

217. *Id.*

218. *See* Born, *supra* note 202, at 1–3, 14–20.

219. *Id.* at 12, 36 n.147.

220. *See* Born, *supra* note 202, at 14 (citing CODE CIVIL [C. CIV.] art 14–15 (Fr.)).

221. *See* Born, *supra* note 202, at 14–15 (citing ZIVILPROZESSORDNUNG [ZPO] [CODE OF CIVIL PROCEDURE] Sept. 12, 1950, § 23 (Ger.); Christof von Dryander, *Jurisdiction in Civil and Commercial Matters Under the German Code of Civil Procedure*, 16 INT'L LAW. 671, 678 (1982)).

222. Peter Weiss, *Universal Jurisdiction: Past, Present and Future*, 102 AM. SOC'Y INT'L L. PROC. 406, 407 (2008) (citing PRINCETON PROJECT ON UNIVERSAL JURISDICTION, THE PRINCETON PRINCIPLES ON UNIVERSAL JURISDICTION 28 (Stephen Macedo ed., 2001) [hereinafter PRINCETON PRINCIPLES] ("[U]niversal jurisdiction is criminal jurisdiction based solely on the nature of the crime, without regard to where the crime was committed, the nationality of the alleged or convicted perpetrator, the nationality of the victim, or any other connection to the state exercising such juris-

criminal universal jurisdiction when they think about universal jurisdiction.

Criminal universal jurisdiction applies to conduct that falls within one of the following categories: "piracy, slavery, genocide, crimes against humanity, war crimes, torture, and 'perhaps certain acts of terrorism.'"[223] Although much of the discussion about universal jurisdiction focuses on criminal cases, universal jurisdiction can also apply in civil contexts.[224] Despite a lack of consensus as to whether human rights are universal, it may be possible to reach a consensus that human rights should be protected from private actors. However, universal civil jurisdiction for corporate-related human rights abuses is not yet an established principle of law.

C. Human Rights and Core Values in Crisis

Alone, the power and influence of transnational corporations may not be sufficient to justify developing alternative regulatory and enforcement options. Even coupled with the challenges of multijurisdictional regulation and enforcement, it may be difficult to convince many scholars and policy makers that additional administrative and regulatory mechanisms are necessary to restrain and harness the power and influence of transnational corporations. However, the nature of human rights as core values and the history of corporate-related human rights abuses lend support to arguments for the need to regulate transnational corporations at a global level.[225]

In western cultures, the idea of human rights as an individual's claim that can be asserted against society (whether or not there is a legal right) has its roots in the medieval era.[226] The western origins of civil and political rights as protections against government have their roots in the Renaissance and the Reformation.[227] However, protection of human

---

diction.")). Universal jurisdiction is one of the most controversial forms of jurisdiction. *See* Henry J. Steiner, *Three Cheers for Universal Jurisdiction—or Is It Only Two?*, 5 THEORETICAL INQUIRIES L. 199, 224–26 (2004) (describing reaction from other states to the potential exercise of extraterritorial jurisdiction by Belgian courts).

223. Anthony J. Colangelo, *The Legal Limits of Universal Jurisdiction*, 47 VA. J. INT'L L. 149, 151 (2006) (citing United States v. Yunis, 924 F.2d 1086, 1091 (D.C. Cir. 1991)); *see also* PRINCETON PRINCIPLES, *supra* note 222, at 29; Anthony J. Colangelo, *The New Universal Jurisdiction: In Absentia Signaling Over Clearly Defined Crimes*, 36 GEO. J. INT'L L. 537, 578–603 (2005).

224. *See* Weiss, *supra* note 222, at 407.

225. *See* Paul R. Dubinsky, *Human Rights Law Meets Private Law Harmonization: The Coming Conflict*, 30 YALE J. INT'L L. 211, 304 (2005) ("From a compliance point of view, what is most distinctive about human rights law is that it requires extraordinarily high compliance levels.").

226. Rubin, *supra* note 14, at 10 (discussing the origins of the doctrine of natural rights). For a justification for a rights approach, see Barbara Stark, *Women's Rights*, *in* ENCYCLOPEDIA OF HUMAN RIGHTS 341–44 (David P. Forsythe, ed., 2009) (arguing that the benefits of a rights approach include the moral weight carried by rights rhetoric, the provision of a theoretical framework, a well-established infrastructure, and the power and flexibility of rights as a tool).

227. *Id.* at 35–37 (discussing the origins of civil and political rights).

rights through law emerged out of religious and public policy considerations.[228]

Historically, rules and laws protecting human dignity and human rights were often articulated in response to extreme power and influence combined with encroachments upon those rights and dignity. Crisis is often the catalyst that prompts change. There are many examples of these times of crisis and change, including in the history of the Catholic Church, the American Revolution, and the post-World War II creation of the international human rights regime.[229]

Religious institutions have espoused human rights—albeit under another name—as core values for centuries. The predecessor of early western human rights theory, natural rights theory, emerged at a time when the Catholic Church exercised enormous power.[230] This time of crisis is exemplified by the tension between the growth of the Franciscan Order after its recognition by Pope Innocent III in 1209 and the Inquisition's increasing violence, including torture.[231] Under the doctrine of natural rights, there are certain "claims that can be asserted by all human beings."[232] These ideas are captured in the writings of William von Ockham, author of the first Western theory of natural rights and a follower of the Franciscan philosophies.[233] They have also been delineated in the Catholic social doctrine.[234]

Human rights have been articulated as core values as part of the creation of nations. In the eighteenth century, the American Declaration of Independence, one of the most famous articulations of individual rights as a political idea, was signed in the context of economic oppression by the British government of its American colonies.[235] This crisis is exemplified by oppression in the form of, for example, the Sugar, Stamp, and Townshend Acts. These acts imposed taxes and import restrictions

228. *Id.* at 168.

229. See generally THOMAS D. WILLIAMS, WHO IS MY NEIGHBOR?: PERSONALISM AND THE FOUNDATIONS OF HUMAN RIGHTS 31–52 (2005), for a historical perspective of the Church's attitudes concerning human rights.

230. *See* Rubin, *supra* note 14, at 10.

231. *See id.* at 12; *Inquisition*, CATHOLIC ENCYCLOPEDIA, http://www.newadvent.org/cathen/08026a.htm#IIA (last visited Oct. 31, 2010).

232. Rubin, *supra* note 14, at 10.

233. *Id.* at 13–14; *see id.* at 14–17 (discussing Ockham's writings).

234. *See generally* JAY P. CORRIN, CATHOLIC INTELLECTUALS AND THE CHALLENGE OF DEMOCRACY 41–81 (2002) (describing the development of Catholic social action and the principles of Rerum Novarum (On Capital and Labor)); RICHARD W. ROUSSEAU, HUMAN DIGNITY AND THE COMMON GOOD: THE GREAT PAPAL SOCIAL ENCYCLICALS FROM LEO XIII TO JOHN PAUL II (2002) (describing the eight major papal encyclicals on the Church's social teaching); JOHAN D. VAN DER VYVER, LEUVEN LECTURES ON RELIGIOUS INSTITUTIONS, RELIGIOUS COMMUNITIES AND RIGHTS 91–116 (2004) (discussing economic and social human rights and Pope Leo XIII and the Rerum Novarum, the "Magna Carta of Catholic social teaching"). See generally MODERN CATHOLIC SOCIAL TEACHING 72–174 (Kenneth R. Himes ed., 2005) for an extensive discussion of the economic, legal, political, cultural, and social context leading to the creation of the Catholic social doctrine.

235. *See* HENKIN, *supra* note 49, at 1.

upon colonists who did not have a voice in the government and were backed by the threat of and acts of violence. The Declaration of Independence posits that all human beings, or at least all men, possess "certain unalienable Rights," including "Life, Liberty and the pursuit of Happiness."[236]

The international community has championed human rights as core values. Modern international human rights documents delineating the rights of individuals and communities against abuses of state power flourished in the wake of the Nazi-driven horrors of the last century.[237] This time of crisis is exemplified by the genocide committed by the German National Socialist government against Jewish, black, Sinti and Roma, and other peoples as well as political dissidents. The 1948 Universal Declaration of Human Rights was signed in the wake of these atrocities.[238]

Private economic actors and international organizations have endorsed human rights as core values. The crisis of corporate-related human rights abuses has been detailed above and there is increasingly broad recognition that transnational corporations should be subject to human rights law.[239] Many corporations are reconsidering their role in respecting human rights.[240] Voluntary codes of conduct have proliferated.[241] At the international level, the U.N. Global Compact promotes core human rights and other principles in the context of global business activities.[242] However, to effectively regulate transnational corporations, regulations and regulatory control must be comprehensive and enforceable.[243] Part III looks towards the development of a comprehensive global legal framework for regulating transnational corporations.

---

236. THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776).

237. HENKIN, *supra* note 49, at 1; *see* MALANCZUK, *supra* note 47, at 209. However, some scholars argue that the Nazi's human rights abuses are less central to the development of the international human rights regime than is often ascribed to them. *See* Mazower, *supra* note 51, at 381 ("As for Nazi evil, we know now that the Holocaust as such was much less central to perceptions of what the war had been about in 1945 than it is today.").

238. HENKIN, *supra* note 49, at 1.

239. *See* Deva, *supra* note 2, at 1 ("The international community is realizing that in order to achieve fuller and wider realization of human rights, the umbrella of human rights obligations and their enforcement should cover MNCs."); *see also* Ratner, *supra* note 82, at 446–47 (providing examples of human rights violations in different regions and the response toward transnational corporations operating there).

240. *See* George, *supra* note 150, at 474–75.

241. George, *supra* note 150, at 475; Ratner, *supra* note 82, at 448. See generally Lisa M. Fairfax, *Easier Said than Done? A Corporate Law Theory for Actualizing Social Responsibility Rhetoric*, 59 FLA. L. REV. 771, 771–826 (2007), for a discussion of the behavioral significance of corporate rhetoric.

242. *See* JOSEPH, *supra* note 1, at 7–8; Anderson, *supra* note 34, at 17; George, *supra* note 150, at 475.

243. BLUMBERG, *supra* note 136, at 200–01 ("The reality of the matter is that effective regulation of corporate groups or their activities inevitably requires control of all the components participating in the enterprise. . . . [I]t is essential that the legal structure match the economic structure of the enterprise subject to the regulatory system."); *see also* Edward L. Rubin, *Passing Through the*

III. GLOBAL REGULATION OF TRANSNATIONAL CORPORATIONS

This Part III sets out a proposal for the global regulation of transnational corporations. This proposal has three parts that address existing institutional, legislative, and enforcement weaknesses in current regulation of transnational corporations. Specifically, it proposes the creation of a global institution, development of uniform model laws, and implementation of universal enforcement mechanisms for corporate-related human rights abuses.

A global regulatory scheme has the potential to be more successful than a political one. In recent years, regulatory cooperation has successfully developed international rules.[244] In part, this is explained by economic assumptions that regulations are needed in the case of market failure; for example, when sovereign national governments cannot regulate global actors or where market failures are combined with government failures.[245] As discussed in Parts I and II, in the context of corporate-related human rights abuses, regulation by national governments is inadequate for a number of reasons—including the power and influence of transnational corporations.

To be successful, a regulatory system for transnational corporations should meet several criteria. Regulation should be truly global so that transnational corporations cannot avoid responsibility and accountability by jurisdiction hopping or forum shopping. Global regulations and laws also have the potential to reduce the incentives for states to allow standards that are so low that they encourage human rights abuses in order to encourage foreign direct investment. If laws and regulations are global and set a globally adhered to minimum standard, this would not eliminate competition between states but it could increase the enactment and enforcement of human rights laws.

Global regulation should include rulemaking, adjudicatory, and enforcement components.[246] Rulemaking would define appropriate conduct for transnational corporations and protections for potential victims of corporate-related human rights abuses. Providing for global adjudication options would increase access for potential plaintiffs to bring a claim for

---

Door: Social Movement Literature and Legal Scholarship, 150 U. PA. L. REV. 1, 83 (2001) ("The recognition of human rights may be won by the activism of social movements, but this victory must be secured by the development of legal concepts that can be understood and used by public decision makers."); Aneel Karani, *The Case Against Corporate Social Responsibility*, WALL ST. J., Aug. 23, 2010 (arguing that voluntary measures fail when they conflict with financial goals).

244. David Zaring, *International Law by Other Means: The Twilight Existence of International Financial Regulatory Organizations*, 33 TEX. INT'L L.J. 281, 282 (1998).

245. John-ren Chen, *Global Market, National Sovereignty and International Institutions*, *in* THE ROLE OF INTERNATIONAL INSTITUTIONS IN GLOBALISATION: THE CHALLENGES OF REFORM 1, 1 (John-ren Chen ed., 2003).

246. See generally David Zaring, *Rulemaking and Adjudication in International Law*, 46 COLUM. J. TRANSNAT'L L. 563, 563–611 (2008), for a discussion of rulemaking, adjudication, and the increasing importance administrative institutions at the global level.

harms before an adjudicatory body. Global enforcement would create an avenue by which successful plaintiffs could obtain remedies anywhere in the world. Ideally, a global regulatory system should not only create disincentives for harmful behavior, but should also encourage beneficial behavior. Further, it should be informed by other prior and current efforts to create global and transnational regulatory bodies and rules.[247] It should also be compatible with existing structures like the U.N. Framework wherever possible.

Creating a global regulatory framework and component mechanisms will not happen overnight. This is a multi-stage, multi-component proposal comprised of short-term, medium-term, and long-term elements. In the short term, developing a global institution is a matter of the first order. As the proposed site for the drafting of uniform model laws, a global institution is a pre-requisite for the medium-term element of creating a body of uniform model legislation. In the medium term, this proposal foresees the drafting of model uniform laws that can be adopted *in toto* or with some variations by the nations of the world.[248] In the long term, this proposal envisions adjudication of claims in domestic courts or the use of alternative dispute resolution mechanisms or both and enforcement of remedies for successful plaintiffs.

---

247. Earlier attempts include the failed UN Norms on the Responsibilities of Transnational Corporations and Other Business Enterprises with Regard to Human Rights and the comparatively successful OECD Guidelines for Multinational Enterprises. U.N. Comm'n on Human Rights, Sub-Comm'n on the Promotion and Prot. Of Human Rights, Rep. on its 55th Sess., U.N. Doc. E/CN.4/Sub.2/2003/12/Rev.2 (August 26, 2003) [hereinafter U.N. Norms]. *See generally* Larry Catá Backer, *Multinational Corporations, Transnational Law: The United Nations' Norms on the Responsibilities of Transnational Corporations as a Harbinger of Corporate Social Responsibility in International Law*, 37 COLUM. HUM. RTS. L. REV. 287, 292–93 (2006) (discussing the U.N. Norms as evidence of an increasing trend toward mandatory Corporate Social Responsibility); Surya Deva, *UN's Human Rights Norms for Transnational Corporations and Other Business Enterprises: An Imperfect Step in the Right Direction?*, 10 ILSA J. INT'L & COMP. L. 493 (2004) (discussing the strengths and weaknesses of the U.N. Norms); David Weissbrodt & Muria Kruger, *Norms on the Responsibilities of Transnational Corporations and Other Business Enterprises with Regard to Human Rights*, 97 AM. J. INT'L L. 901, 901–08 (2003) (discussing the background and drafting history of the U.N. Norms). Most prior attempts to extend human rights law to transnational corporations were usually unsuccessful. Other attempts include, for example, the *1983 Draft United Nations Code of Conduct on Transnational Corporations*, the 1958 *Abs Shawcross Proposed Convention to Protect Foreign Investment*, and the 1957 *Draft International Convention for the Mutual Protection of Private Property Rights*, U.N. Comm'n on Transnat'l Corps., Draft U.N. Code of Conduct on Transnat'l Corps., E/1983/17/Rev.1 (1983). Arguably, one main reason these attempts failed is that they were trying to work within the *international* system, which suffers from an aggregated lack of incentives for states to enforce human rights laws as discussed above. As long as states have a disincentive to reform, such attempts face an uphill battle that almost dooms them to failure from the start.

248. Bernardo M. Cremades & Steven L. Plehn, *The New Lex Mercatoria and the Harmonization of the Laws of International Commercial Transactions*, 2 B.U. INT'L L.J. 317, 323 (1984) ("The model law process represents a compromise between the treaty process and purely unilateral action by nations."). However, variations to model laws should not be for the purpose of limiting their scope as is often the case with the use of reservations in the adoption of international treaties but rather, for example, for the purpose of integrating the model laws and the existing national legal system.

This proposal differs from earlier attempts in important ways. It is not set within the *international* political system, which suffers from an aggregated lack of incentives for states to enforce human rights laws as discussed above. Instead, it is structured as a *global* regulatory proposal. This will avoid some of the reasons that states have a disincentive to reform and sidestep the uphill battle that has doomed some prior efforts. This proposal envisions a global institution that, among other things, should be structured in a way to allow it to explicitly address the needs of developing countries from a developing country perspective, which is not the case for most international institutions.[249] This proposal draws on theories of Global Corporate Citizenship, which are a recent development of the past decade. However, Global Corporate Citizenship is currently voluntary and this Article is proposing mandatory regulations.[250] It also proposes universal civil jurisdiction for corporate-related human rights abuses, which extends beyond current understandings of universal civil jurisdiction.[251]

A. Global Law Commission

The first component of this proposal is a global institution, specifically, a Global Law Commission. This is an initial proposal that will be expanded and refined in a future article entitled *Global Law Commission: Institutionalizing Global Human Rights Regulation of Transnational Corporations* (working title). This Section sets out some of the options and challenges for a Global Law Commission.

A Global Law Commission would facilitate effective and simultaneous regulation of transnational corporations in multiple jurisdictions.[252] One goal of the Global Law Commission would be to provide countries with well-conceived, well-drafted legislation that brings clarity and stability to this critical area of the law. The Global Law Commission could contribute to protecting human rights and providing access to remedies for corporate-related human rights abuses. Among other things, the Global Law Commission should explicitly address the needs of developing countries. This Subsection sets out in broad brushstrokes basic contours of a proposal, including formation, membership, level of institutionalization, funding, selection of representatives, transparency of deci-

---

249. Kunibert Raffer, *Some Proposals to Adapt International Institutions to Developmental Needs*, *in* THE ROLE OF INTERNATIONAL INSTITUTIONS IN GLOBALISATION: THE CHALLENGES OF REFORM 81, 81 (John-ren Chen ed., 2003) ("Most international institutions were not drafted for the specific needs of so-called 'developing countries' . . . .").

250. Anderson, *supra* note 34, at 29.

251. *See generally* Donald Francis Donovan & Anthea Roberts, *The Emerging Recognition of Universal Civil Jurisdiction*, 100 AM. J. INT'L L. 142 (2006) (discussing universal civil jurisdiction).

252. *See* Mark Tushnet, *Keeping Your Eye on the Ball: The Significance of the Revival of Constitutional Federalism*, 13 GA. ST. U. L. REV. 1065, 1067 (1997) ("It seems obvious that only transnational political institutions are likely to be in a position to control transnational corporations.").

sion-making, and the role of the Global Law Commission as an advocate.[253]

1. Formation

The Global Law Commission should be formed in a manner that allows it to act independently of the international legal system while at the same time recognizing the interdependent nature of all legal systems. For example, like international financial regulatory organizations, the Global Law Commission could be informally constituted.[254] It should not be created through international treaties, as it is not intended to be a creature of international law.[255]

Instead, the constituting documents of the Global Law Commission could be a constitution (or similar document) and bylaws. This is a model that has been employed in the creation of successful international financial institutions including the Basle Committee on Banking Supervision and the International Organization of Securities Commissions.[256] Domestically, it is also the constituting structure of the U.S. National Conference of Commissioners on Uniform State Laws ("U.S. Uniform Law Commission").[257] This structure would facilitate the operation of the Global Law Commission as a global administrative and regulatory institution rather than a political institution. It would also allow the bylaws to be "broad and flexible."[258]

2. Membership

The membership of the Global Law Commission should be composed of commissioners who represent the people of the world. In the Global Law Commission, states would not be members.[259] This Subsec-

253. David Zaring, *Informal Procedure, Hard and Soft, in International Administration*, 5 CHI. J. INT'L L. 547, 569–71 (2005) (developing a model of regulatory cooperation on which these categories are based). Instead of Zaring's term "proselytizer," this Article uses the term "advocate."

254. *See id.* at 569–71 (discussing a model of regulatory cooperation based on common features of international law of the Basle Committee, the International Organization of Securities Commissions, and their successors).

255. *See id.*

256. *See* Zaring, *supra* note 244, at 282. International financial regulatory cooperation is one area in which institutions and rules have been successfully developed. *Id.* at 282; *see also* David Zaring, *Three Challenges for Regulatory Networks*, 43 INT'L LAW. 211, 211–17 (2009) (discussing the evolution and effectiveness of regulatory networks); Zaring, *supra* note 253, at 549 (discussing the Basle Committee on Banking Supervision and the International Organization of Securities Commissions).

257. The Uniform Law Commission has existed in the United States for over 100 years. *Unif. Law Comm'n.*, NCCUSL HOME, *at* http://www.nccusl.org/Update/ (last visited Oct. 31, 2010). The Uniform Law Commission "provides states with non-partisan, well-conceived and well-drafted legislation that brings clarity and stability to critical areas of the law." *Id.* Historically, the Uniform Law Commission has made numerous contributions including resolving conflict-of-law issues, facilitating commerce, and advancing reciprocity and remedies between states. Chato Hazelbaker, *National Legislation, One State at a Time*, 66-DEC BENCH & B. MINN. 22, 23 (2009).

258. Zaring, *supra* note 253, at 570 (stating that "many treaties that create traditional international organizations" are "specific, detailed, and constraining").

259. *See id.* at 569.

tion discusses initial proposals for the selection of members of the Global Law Commission. Other issues that will need to be addressed include determining criteria for membership and the size of the Global Law Commission, specifically the minimum and maximum number of members.

One option is to allow only national governments to appoint or nominate commissioners in a process that is similar to that used for the U.S. Uniform Law Commission or the International Law Commission.[260] A downside to this membership structure is individual countries' ability to interfere in the process by simply refusing to appoint or nominate commissioners. However, this could be remedied by giving the Global Law Commission the power to request the body that regulates the legal profession in that country to make appointments or nominations—assuming that the body regulating the legal profession is able to act independently. More significantly, it is inherently inconsistent and problematic to base membership selection on the international legal system since the global law system is intended to function in parallel with but not be dependent on the international legal system. If national governments appoint or nominate the commissioners, then they are arguably representatives of countries even if the Global Law Commission is structured to ensure them as much independence as possible.

Instead, one alternative for membership selection is proportional representation. The Global Law Commission would have a fixed or flexible—within a certain range—number of members. The commissioners could be allocated proportionally, for example, by continent according to population or other criteria and the allocation could be reviewed and revised at set intervals. This would reflect the nature of the Global Law Commission as a global institution rather than an international institution.

Commissioners could be nominated by national governments, NGOs, bodies that regulate the legal profession, as well other institutional actors. To be elected, nominees would need to meet the criteria set for membership in the Global Law Commission. Institutions that are given the right to nominate could also constitute the body that elects the commissioners from those nominated based on a one-entity, one-vote principle. One challenge associated with this option is determining which institutions and groups would have the right to make nominations. A benefit of this mode of selection would be to expand the pool of nominators and, thus, the pool of potential nominees. It would also serve to de-

260. UNIFORM LAW COMMISSION CONST. art. II, *available at* http://www.nccusl.org/Update/DesktopDefault.aspx?tabindex=3&tabid=18; UNITED NATIONS, THE WORK OF THE INTERNATIONAL LAW COMMISSION, 8–17 (Vol. I, 2007); JEFFREY S. MORTON, THE INTERNATIONAL LAW COMMISSION OF THE UNITED NATIONS 10-12 (2000); and SIR IAN SINCLAIR, THE INTERNATIONAL LAW COMMISSION 13-21 (1987).

crease the level of control by national governments but recognize the importance of national governments and the interconnection and interdependency of all legal orders of the world.

Another issue that will need to be addressed is the possibility of creating multiple forms of membership, such as *ex officio* membership in addition to elected membership, and the advantages and disadvantages of these alternative forms. Having different forms of membership would make it possible to select commissioners in more than one way. For example, some percentage of the commissioners could be nominated and elected among national governments, some could be nominated and elected by NGOs and other institutional actors, and some could be nominated and elected by bodies that regulate the legal profession. Allocating certain percentages of commissioners to certain interest groups such as national governments would help garner support from those groups for the Global Law Commission's work. However, this might expose the Global Law Commission's work to becoming more political than regulatory.

3. Institutionalization

The Global Law Commission should meet at regular intervals, for example annually, to draft model uniform laws in key human rights areas affected by transnational corporations. In addition, it is likely that the Global Law Commission would need to continue work on ongoing projects in committees. Committees could be determined based on substantive areas or type of tasks. Substantive committees offer an advantage because they allow the clustering of substantive expertise. However, task oriented committees may present a better option because the make-up of the committee would bring together people with varying substantive expertise who, together, have the skills and knowledge to achieve the committee's assigned tasks.

The Global Law Commission should have a permanent presence. Either a small permanent presence or a more substantial permanent presence would have their benefits. A small permanent presence would keep costs, and thus the need for funding, low.[261] A more substantial permanent presence would cost more but would allow the Global Law Commission to operate year-round. However, an increased need for funds amplifies questions about sources of funding and the level of influence of governments and intuitions providing those funds.

4. Funding

Funding will be a significant factor in the size and success of the Global Law Commission's administrative apparatus. The level of fund-

261. Zaring, *supra* note 253, at 571 (stating that international financial regulatory organizations "have little permanent presence" and "[t]heir annual budgets are minute").

ing required will depend in large part on whether the Global Law Commission has a large or small permanent presence. The larger the permanent presence, the greater the need will be for funding.[262] However, funding will be a less significant factor if the commissioners work on a volunteer basis and are not compensated. Nonetheless, some minimal funding would be necessary, even for volunteer commissioners, to defray expenses.

Contributions from interested parties such as national governments are one likely source of funding for the Global Law Commission. States would need to have an incentive to contribute to the Global Law Commission. One incentive for states would be that they could receive legal expertise in the form of drafting of uniform laws worth millions of dollars. Many individual states could not otherwise afford the cost of this type of legal expertise.[263]

5. Transparency

Ensuring a high level of transparency can help the Global Law Commission gain and maintain legitimacy. One question is whether meetings of the commission should be open to the public.[264] Open meetings may increase accessibility for those who have the time and financial resources to attend. However, on a global scale, open meetings may also create and legitimize preferential access for more influential actors and exclude interested parties with fewer resources. In addition, open meetings may stifle open and frank discussion.

Instead, the Global Law Commission should have a forum for publicizing drafts and accepting comments from interested parties.[265] The Global Law Commission should also commit to making a wide range of relevant documents available online.[266] Ensuring an opportunity for public comment on drafts via the Internet will increase transparency—assuming that the opportunity to comment is widely publicized in a manner that reaches interested parties.

---

262. *See id.* This discussion is intended only to raise the complex and fundamental issue of state support, in particular financial support. The question of securing state support will be analyzed in greater detail in a future article entitled *Global Law Commission: Institutionalizing Global Human Rights Regulation of Transnational Corporations* (working title).

263. *Financial Support*, NCCUSL, http://www.nccusl.org/Update/DesktopDefault.aspx?tabindex=0&tabid=11 (last visited on Oct. 31, 2010) (drawing from the Uniform Law Commission model).

264. This draws from the U.S. Uniform Law Commission model. UNIFORM LAW COMMISSION. art. XI. In contrast, international financial regulatory organization meetings are generally not open to the public. Zaring, *supra* note 253, at 571 ("[International financial regulatory organizations] generally operate through closed meetings . . . .").

265. *See* Zaring, *supra* note 253, at 571 ("[T]he Basle Committee and [International Organization of Securities Commissions] . . . make an effort to invite comment from interested parties on their most significant regulatory efforts.").

266. *Id.* at 571–72 (discussing a practice that has become common among international financial regulatory organizations).

6. Advocacy

Model laws drafted by the Global Law Commission can contribute to the development of both soft and hard law. Like other documents written by international organizations and NGOs, model laws drafted by the Global Law Commission may influence soft-law norms and contribute to the enactment of hard law. Model laws drafted by the Global Law Commission for domestic implementation would need to be passed by national governments to be enforceable in domestic courts. This could change if, at a later stage, a global court is established or countries sign onto conventions to enforce alternative dispute resolution or universal civil jurisdiction for claims related to corporate-related human rights violations.

Encouraging national governments to implement draft laws in their jurisdiction will be an important role for the Global Law Commission.[267] Like international financial regulatory organizations, the Global Law Commission would need to work actively to disseminate best practices. In addition, the Global Law Commission could advocate for and play a role in drafting conventions to enforce forms of alternative dispute resolution or universal civil jurisdiction as well as the creation of a global court.

Implementing the Global Law Commission will not be simple or without challenges. One of the key challenges is how to get countries on board with the Global Law Commission and enact the model laws developed by the Global Law Commission. However, there are examples that can provide insights into avenues that may lead to success. For example, forty-two countries adhere—more or less—to the OECD Guidelines for Multinational Enterprises.[268] An examination of the strengths and weaknesses of that process can provide valuable insights. Further, an analysis of why the United States became a member of the United Nations but not the League of Nations, and why the United States signed onto the General Agreement on Tariffs and Trade but not the International Trade Organization offers an opportunity to identify factors that may encourage countries like the United States to support—or at least accept—the Global Law Commission.

B. Global Laws and Regulations

The second component of this proposal is the development of global laws and regulations by the Global Law Commission informed by a the-

267. *Id.* at 572 (discussing the proselytizing roles played by international financial regulatory organizations).

268. Directorate for Fin. & Enter. Affairs, *The OECD Guidelines for Multinational Enterprises: Frequently Asked Questions*, ORG. FOR ECON. CO-OPERATION & DEV., http://www.oecd.org/document/58/0,3343,en_2649_34889_2349370_1_1_1_1,00.html (last visited on Oct. 31, 2010).

ory of Law and Global Corporate Citizenship.[269] This is an initial proposal that will be expanded and refined in a future article entitled *A Legal Theory of Global Corporate Citizenship: Reimagining Global Regulation of Transnational Corporations* (working title). Existing legal theory does not provide a viable framework in which to adequately address the effects of corporate activities on human rights. In addition to the weaknesses in international human rights law discussed above, theories of corporate law and foreign direct investment law are also inadequate to comprehensively theorize and regulate the intersection of transnational economic activity and human rights.[270] Because there is not an adequate theoretical framework to address issues raised by the effects of transnational corporations on individuals' and communities' enjoyment of their human rights, corporate-related human rights abuses remain on the periphery of law and theory.

In corporate law and theory in the United States, corporations are generally depicted as economic entities that should be regulated by the market through economic incentives and mechanisms.[271] This position is solidly established at the heart of American corporate legal theory.[272] However, there is a minority in the United States that argues that corporations are economic *and social* entities.[273] This minority often identifies with the Corporate Social Responsibility movement.[274] This line of argument does not limit the means of controlling corporations to economic incentives. Nonetheless, it remains on the periphery of U.S. corporate legal theory and is not likely to become a central tenet of U.S. corporate law and theory anytime in the near future.

Lacking traction in corporate law and theory, we might logically turn to the web of domestic laws, bilateral investment treaties, and other international law that govern foreign direct investment. However, foreign

---

269. Although the development of model global laws does not reduce the need for plaintiffs to have the financial capacity to bring suits, adherence to human rights-related laws and norms by transnational corporations can reduce the need for plaintiffs to bring suits by reducing the occurrence of corporate-related human rights abuses. *See* Backer, *supra* note 247, at 351–54.

270. Similar arguments can be made for other areas of law and theory, for example, law and development.

271. Backer, *supra* note 247, at 296–98 ("The state was to define the parameters within which this flexible framework could be effected and policed, but otherwise the market was to provide the mechanism for regulating corporate activity.").

272. *Id.* at 296–99 (tracing this position back to the acceptance of arguments made by Adolph Berle in a debate with E. Merrick Dodd in the 1930s); *see also* A. A. Berle, Jr., *Corporate Powers as Powers in Trust*, 44 HARV. L. REV. 1049, 1049 (1931) (discussing the impact of corporate actions taken serve to benefit its shareholders); E. Merrick Dodd, Jr., *For Whom are Corporate Managers Trustees?*, 45 HARV. L. REV. 1145, 1145–47 (1932) (discussing the purpose and function of the corporate structure).

273. *See* Cynthia A. Williams, *Corporations Theory and Corporate Governance Law: Corporate Social Responsibility in an Era of Economic Globalization*, 35 U.C. DAVIS L. REV. 705, 716 (2002).

274. The exact scope and contours of Corporate Social Responsibility are disputed within the U.S. legal discourse. *See id.* at 711–20 (categorizing positions taken on CSR in the U.S. legal discourse).

direct investment is generally thought of as a primarily economic activity and the laws governing such activities reflect those assumptions.[275] Purely domestic laws often fall prey to lax enforcement in the face of government goals to bring in funds and promote economic development through foreign investment.[276] Bilateral and other international agreements between countries tend to focus on protecting their own natural and legal citizens' foreign investments from expropriation and naturalization without appropriate compensation.[277] It is unlikely that the protection of the rights and interests of non-citizens will be given a higher priority than the protection of a state's own citizens and economic interests in foreign direct investment law. Thus, the regulation of the negative environmental and human rights effects of corporations is likely to remain on the periphery of foreign direct investment law and theory.

However, theoretical advancements have been made in other disciplines. Global Corporate Citizenship is a body of theoretical work initially developed in the management and business literature in the late 1990s that could provide a strong basis for developing global laws to regulate transnational corporations.[278] Over the past decade, a robust Global Corporate Citizenship literature has developed in the business and management fields.[279] This literature argues for voluntary initiatives by transnational corporations.[280] The rationale for transnational corporations

---

275. *See* Dr. Elfraim Chalamish, *The Future of Bilateral Investment Treaties: A De Facto Multilateral Agreement?*, 34 BROOK. J. INT'L L. 303, 328 (2009).

276. This problem should be addressed in part by universal civil jurisdiction, which would allow plaintiffs to bring cases in courts of countries with a record of more consistent enforcement. *See* Donovan & Roberts, *supra* note 251, 142–46 ("[U]niversal jurisdiction provides a mechanism for enhancing accountability.").

277. *See, e.g.*, Jason Webb Yackee, *Bilateral Investment Treaties, Credible Commitment, and the Rule of (International) Law: Do BITs Promote Foreign Direct Investment?*, 42 LAW & SOC'Y REV. 805, 808 (2008) ("Most BITs contain a common core of substantive promises to investors, including rights to some combination of 'most-favored nation,' national, 'non-discriminatory,' or 'fair and equitable' 'treatment'; rights to 'full protection and security'; rights to 'prompt, adequate, and effective' compensation in the event of expropriation or of government measures 'tantamount to expropriation'; and the right to transfer investment assets or proceeds out of the host state in convertible currency.").

278. *See* LOGAN, ROY & RUGELBRUGGE, *supra* note 96, at 13, 16.

279. *See, e.g.*, LOGAN, ROY & RUGELBRUGGE, *supra* note 96, at 6–7; James E. Post & Shawn L. Berman, *Global Corporate Citizenship in a Dot.com World: The Role of Organisational Identity*, *in* PERSPECTIVES ON CORPORATE CITIZENSHIP 66, 66–67 (Jörg Andriof & Malcolm McIntosh eds., 2001); Jeanne M. Logsdon & Donna J. Wood, *Business Citizenship: From Domestic to Global Level of Analysis*, 12 BUS. ETHICS Q., no. 2, 2002, at 155; James E. Post, *Global Corporate Citizenship: Principles to Live and Work By*, 12 BUS. ETHICS Q., no. 2, 2002, at 143, 148–49 [hereinafter *Principles to Live and Work By*]; James E. Post, *Moving from Geographic to Virtual Communities: Global Corporate Citizenship in a Dot.com World*, 105 BUS. & SOC. REV. 27, 27–29 (2000); Klaus Schwab, *Global Corporate Citizenship: Working With Governments and Civil Society*, 87 FOREIGN AFF. 107, 107–108 (2008); Sandra Waddock & Neil Smith, *Relationships: The Real Challenge of Global Corporate Citizenship*, 105 BUS. & SOC. REV. 47, 47–49 (2000); Grahame Thompson, *Tracking Global Citizenship: Some Reflections on 'Lovesick' Companies* 1–2 (Inst. for Int'l Integration Studies, Discussion Paper No. 192, 2006), *available at* https://www.tcd.ie/iiis/documents/discussion/pdfs/iiisdp192.pdf.

280. *See generally* Schwab, *supra* note 269, at 107–18 (discussing participation and benefits in engaging in global corporate citizenship and corporate social responsibility).

to undertake these initiatives is that they have economic benefits for the corporations in addition to being beneficial for society.[281]

However, one danger of relying on voluntary measures is that they themselves often rely on the flawed assumptions of a purely economic analysis of business management. Purely economic analyses of business choices are susceptible to over-reliance on quantitative tools.[282] Overreliance on quantitative tools is often accompanied by a disregard for moral and cultural factors, which artificially limits our understanding of issues and the range of corresponding regulatory options.[283] Instead, some scholars argue for a "deeply human-centered conception of business" that "depend[s] on the exercise of virtue, respect for dignity, and a shared sense of the common good."[284] Voluntary measures are also problematic because the self-interests of transnational corporations may not always align with societal interests. In such situations, voluntary measures are often—at most—a minimal but not dispositive incentive.

Global Corporate Citizenship literature builds on corporate citizenship literature by expanding it to the global sphere. Global Corporate Citizenship initiatives have garnered the attention and support of transnational corporations.[285] At the international level, the Global Compact, a United Nations initiative, has promoted a set of values that are aligned with theories of Global Corporate Citizenship.[286] However, the voluntary nature of these initiatives exposes them to the same weaknesses as other voluntary measures—transnational corporations can pick and choose among measures. From a business perspective, this is beneficial because it allows managers and directors to choose measures that suit their context and that fit their business modes.[287] From a human rights perspective, this is a disadvantage because voluntary norms and standards do not require transnational corporations to consider, recognize, or adhere to any domestic or international norms or standards such as international

281. *See Principles to Live and Work By*, *supra* note 279, 143–44, 148.

282. Kevin T. Jackson, *The Scandal Beneath the Financial Crisis: Getting a View from a Moral-Cultural Mental Model*, 33 HARV. J.L. & PUB. POL'Y 735, 746–47 (2010) (stating that the trend of increasing mathematization of science "is not as suitable for examining a broad range of phenomena—such as institutions, values, culture, and traditions—that clearly have an enormous bearing on economic life.").

283. *Id.* at 747 (quoting WILHELM RÖPKE, A HUMANE ECONOMY: THE SOCIAL FRAMEWORK OF THE FREE MARKET 247 (Elizabeth Henderson trans., 1960) (lamenting the decline of morality in economics)).

284. *Id.* at 752–53 ("[The] human-centered conception of business is supported by a long tradition of thought common to ancient cultures.") (citing SAMUEL GREGG, THE COMMERCIAL SOCIETY: FOUNDATIONS AND CHALLENGES IN A GLOBAL AGE 9 (2007)).

285. Anderson, *supra* note 34, at 22–23.

286. *The Ten Principles*, UNITED NATIONS GLOBAL COMPACT, http://www.unglobalcompact.org/AboutTheGC/TheTenPrinciples/index.html (last visited on Oct. 31, 2010).

287. *Cf.* Jackson, *supra* note 282, at 758 (stating that reactionary corporate social responsibility measures may be disadvantageous when companies are "placed under strict accountability, compliance, and enforcement demands in utter disregard of the type and character of the business at hand the conditions under which it might prosper").

human rights law.[288] Further, such voluntary measures are often reactionary and piecemeal rather than proactive and comprehensive.[289]

Although Global Corporate Citizenship is well developed in the management and business literature, it is not yet firmly established in legal scholarship.[290] In the business and management context, it is an evolving voluntary practice. Businesses are supposed to be good corporate citizens because it makes business sense—there are economic benefits. In the legal context, mandatory Global Corporate Citizenship is more appropriate when core values such as human rights are at issue. Mandatory laws and regulations are better suited to ensuring compliance and recourse to enforcement than voluntary ones. Thus, it is important to develop mandatory regulations in addition to voluntary initiatives.

There are four principles of Corporate Citizenship: *Minimize Harm*; *Maximize Benefit*; *Accountability and Responsiveness to Key Stakeholders*; and *Strong Financial Return*.[291] The first principle of Corporate Citizenship is *Minimize Harm*, which is defined as:

> Work[ing] to minimize the negative consequences of business activities and decisions on stakeholders, including employees, customers, communities, ecosystems, shareholders and suppliers. Examples include operating ethically, supporting efforts to stop corruption, championing human rights, preventing environmental harm, enforcing good conduct from suppliers, treating employees responsibly, ensuring the safety of employees, ensuring that marketing statements are accurate, and delivering safe, high-quality products.[292]

The second principle of Corporate Citizenship is *Maximize Benefit*, defined as:

> "Contribut[ing] to societal and economic well-being by investing resources in activities that benefit shareholders as well as broader stakeholders." Examples include participating voluntarily to help address social issues (such as education, health, youth development, economic development for low-income communities, and work force

---

288. Donovan & Roberts, *supra* note 251, at 145.

289. *See* Jackson, *supra* note 282, at 758 ("Under the 'corporate social responsibility' approach, the province of 'business ethics' gets denigrated to harum-scarum stratagems formulated as reactions to alarms sounded by 'stakeholders' that are in turn dictated by galleries of activists purporting to be their appointed representatives."). Jackson critiques such measures and argues that such measures are often aimed at "promoting politically correct agendas." *Id.*

290. *See* Anderson, *supra* note 34, at 23–24.

291. E-mail from Susan Thomas, Assistant Dir., Boston College Center for Corporate Citizenship, to Sarah Millard, Executive Editor, *Denver University Law Review* (Sept. 16, 2010, 15:59 MDT) (on file with *Denver University Law Review*) (This was originally posted in on the Boston College Center for Corporate Citizenship website as *What is Corporate Citizenship?*, BOSTON COLLEGE CENTER FOR CORPORATE CITIZENSHIP.) [hereinafter Boston College]. These principles are no longer listed on the Boston College Center for Corporate Citizenship website. However, I continue to use them here because I believe they set out a solid initial framework for corporate citizenship.

292. *Id.*

development), ensuring stable employment, paying fair wages, and producing a product with social value.[293]

The third principle of Corporate Citizenship is *Accountability and Responsiveness to Key Stakeholders*, which has two parts.[294] The first part of the third principle is "[b]uild[ing] relationships of trust that involve becoming more transparent and open about the progress and setbacks businesses experience in an effort to operate ethically."[295] The second part of the third principle is "[c]reat[ing] mechanisms to include the voice of stakeholders in governance, produce social reports assured by third parties, operate according to a code of conduct and listen to and communicate with stakeholders."[296] Examples include creating mechanisms to include the voice of stakeholders in governance, producing social reports assured by third parties, operating according to a code of conduct, and listening to and communicating with stakeholders.[297]

Finally, the fourth principle of Corporate Citizenship is a *Strong Financial Return*; this is defined as "return[ing] a profit to shareholders."[298] Business and management scholars have expressed the view that the responsibility of a company to return a profit to shareholders must always be considered part of its obligation to society.[299]

Developing uniform model laws informed by principles of Global Corporate Citizenship will require, among other things, transforming voluntary guidelines and standards into mandatory and enforceable rules and regulations.[300] The first principle of Corporate Citizenship, *Minimizing Harm*, requires transnational corporations to minimize negative effects on all categories of stakeholders.[301] In many ways, a legal standard of *Minimizing Harm* could parallel this standard as it is set out in the business and management literature. However, in a legal framework, *Minimizing Harm* and sector-specific standards would be legally mandated. These would need to be enforceable under law to encourage compliance and provide legal avenues for redress. Such standards already exist in a variety of areas such as the environment and labor law. In some ways, developing laws to address this element are the easiest because they fit easily into existing models that prohibit harmful behavior. Laws

293. *Id.*
294. *Id.*
295. *Id.*
296. *Id.*
297. *Id.*
298. *See id.*
299. H. Jeff Smith, *The Shareholders v. Stakeholders Debate*, 44 MIT SLOAN MGMT. REV. 85, 85–86 (2003) (discussing the rationale for profit in both shareholder and stakeholder theories of management).
300. Options for lawmaking will need to be worked through in detail but this goes beyond the scope of this paper.
301. Boston College, *supra* note 291.

in this category could include and build on concepts developed in the context of Corporate Social Responsibility.

As a legal standard, the second principle of Corporate Citizenship, *Maximizing Benefit*, would require by law affirmative steps on the part of transnational corporations to contribute to societal and economic well being. This would go beyond the investment of resources in activities that benefit various groups of stakeholders. Laws requiring *Maximizing Benefit(s)* would need to be informed by research that explores what kinds of laws create positive incentives. One way to maximize the benefits of foreign direct investment is for transnational corporations to exert positive influence over things like water and sanitation,[302] property rights,[303] and entrepreneurial skills and capabilities.[304] This may be one of the most controversial principles of corporate citizenship in a legal context, particularly from the perspective of U.S. scholars and decision makers.

The third principle of Corporate Citizenship, *Accountability and Responsiveness to Key Stakeholders* would require standards and transparency in a legal context.[305] It would differ from the business and management perspectives because it would not focus on developing trust, which is an essential goal of voluntary businesses relationships. Instead, it would stem from the idea that transnational corporations have duties and obligations to specific groups of stakeholders and to society as whole. In a mandatory legal framework, this principle could encompass the setting of standards for ethical business operation. It would also include core principles of corporate citizenship in the business and management literature, such as stakeholder participation and third party monitoring[306] Laws in this category could include and build on those developed in the context of Corporate Social Accountability.

For a mandatory legal framework, this Article proposes a substantive change to the fourth principle of Corporate Citizenship. Rather than a *Strong Financial Return*, a legal framework could instead require a *Competitive Financial Return*. This would be a shift from traditional profit-maximization models, which go further than the goal of a *Strong*

302. U.N. CONFERENCE ON TRADE AND DEVELOPMENT, WORLD INVESTMENT REPORT 2008: TRANSNATIONAL CORPORATIONS AND THE INFRASTRUCTURE CHALLENGE, at 140 , U.N. Sales No. E.08.II.D.23 (2008), *available at* http://www.unctad.org/en/docs/wir2008_en.pdf (discussing the effects of the lack of water and sanitation on child mortality, maternal health and gender equality).

303. *See, e.g.*, Leslie Kurshan, *Rethinking Property Rights as Human Rights: Acquiring Equal Property Rights for Women Using International Human Rights Treaties*, 8 AM. U. J. GENDER SOC. POL'Y & L. 353, 357–59 (2000).

304. *See* OECD, FOREIGN DIRECT INVESTMENT FOR DEVELOPMENT: MAXIMIZING BENEFITS, MINIMIZING COSTS 14–15 (2002) (discussing potential for multinational corporations to enhance entrepreneurial capacity in developing countries), *available at* http://www.oecd.org/dataoecd/47/51/1959815.pdf.

305. Boston College, *supra* note 291.

306. *See id.*

Financial Return.[307] Achieving a *Strong Financial Return* is a goal for a corporation and its officers and directors.[308] It is an internal guideline. Achieving a *Competitive Financial Return* is a standard that can be balanced against industry norms and practices as well as the other three principles of corporate citizenship. A *Competitive Financial Return* standard would allow corporations to consider the interests of stakeholders other than shareholders. In contrast to a *Strong Financial Return*, a *Competitive Financial Return* is an external standard that should be determined with consideration of external factors. In a legal context, appropriate legislation would need to be enacted to give transnational corporations an incentive to balance economic goals with socio-political goals.

This proposal is structured to be compatible with the U.N. Framework. Transnational corporations' effects on human rights have been the focus of recent work by John Ruggie's team under the auspices of the United Nations.[309] Ruggie and his team have developed a "protect, respect, and remedy framework," also known as the U.N. Framework, that can serve as a foundation for further work in this area.[310]

Integrating Global Corporate Citizenship into the U.N. Framework will allow for the development of a more robust and comprehensive system. Imagine then a matrix in which to classify global lawmaking. The x-axis would be populated by the categories of protect, respect, and remedy from the U.N. Framework. The y-axis would be populated by the four principles of corporate citizenship. Laws would be classified by their position at the intersection of the x- and y-axes.

Simply drafting model global laws has the potential to create soft law and encourage practices that reduce or prevent corporate-related human rights abuses. However, without hard law, potential plaintiffs would not be able to seek judicial or other enforcement of their rights. The next Section sets out an initial proposal for enforcement, the final element of this proposal.

---

307. *See* Smith, *supra* note 299, at 85 ("As Milton Friedman wrote, 'There is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it . . . engages in open and free competition, without deception or fraud.'" (alteration in original)).

308. *See* Boston College, *supra* note 291 (articulating principle of "support[ing] strong financial results" using normative rather than binding language).

309. Special Representative of the Secretary-General on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises, *Business and Human Rights: Further Steps Toward the Operationalization of the "Protect, Respect and Remedy" Framework*, U.N. Doc. A/HRC/14/27 (April 9, 2010) (by John Ruggie), *available at* http://www.reports-and-materials.org/Ruggie-report-2010.pdf; Special Representative of the Secretary-General on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises, *Protect, Respect and Remedy: A Framework for Business and Human Rights*, U.N. Doc A/HRC/8/5 (April 7, 2008) (by John Ruggie), *available at* http://www.reports-and-materials.org/Ruggie-report-7-Apr-2008.pdf.

310. *Id.*

C. Global Enforcement

The rights of victims of corporate-related human rights abuses are under-enforced and, therefore, the third component of this proposal is the expansion of adjudicatory and enforcement options.[311] This expansion should help compensate for the fact that some states enforce and adjudicate fewer claims or adjudicate less consistently than others. This proposal envisions adjudication of claims of corporate-related human rights abuses in domestic courts and universal civil jurisdiction for corporate-related human rights abuses.[312] However, it does not exclude the possibility of a global court or convention on enforcement of alternative forms of dispute resolution in the future.[313] This is an initial proposal that will be expanded and refined in a future article entitled *Universal Civil Jurisdiction: Adjudicating Corporate-Related Human Rights Abuses* (working title). Universal civil jurisdiction would allow civil claims of corporate-related human rights abuses to be brought in domestic courts regardless of the nationality of the parties or the location and effects of the conduct.[314]

Universal civil jurisdiction should be established for the limited scope of corporate-related human rights abuses. In a multinational context, jurisdiction is generally determined by domestic and international rules.[315] Universal civil jurisdiction would make it possible to bring a civil suit against a transnational corporation for corporate-related human rights abuses by any plaintiff in a domestic court in any country. One issue raised by universal civil jurisdiction is how to prevent unfairly overburdening the courts of any one jurisdiction.

---

311. This is an initial proposal that I am expanding and refining in a forthcoming article entitled *Toward Global Adjudication and Enforcement of Corporate-Related Human Rights Abuses* (working title).

312. *See* Beth Van Shaack, *Justice Without Borders: Universal Civil Jurisdiction*, 99 AM. SOC'Y INT'L L. PROC. 120, 122 (2005) (arguing that "[s]upranational mechanisms will never supplant domestic proceedings, so domestic courts will continue to play a central role in enforcing international law"); *see also* Dubinsky, *supra* note 225, at 306 (arguing that "[b]ecause much that is productive has come out of human rights adjudication to date, because the phenomenon is still fairly recent, and because there is not much that suggests it is harmful, we should go forward with a series of rebuttable presumptions: that victims of fundamental human rights treaty violations should not be beyond effective judicial protection; that alleged perpetrators should not enjoy impunity; that courts, as institutions best designed to provide fair process and decision-making insulated from political pressure, are appropriate for dealing with guilt, punishment, liability, and compensation").

313. *See* Dubinsky, *supra* note 225, at 307–12 (comparing the advantages and disadvantages of international vs. domestic courts).

314. *See* Menno T. Kamminga, *Universal Civil Jurisdiction: Is it Legal? Is it Desirable?*, 99 AM. SOC'Y INT'L L. PROC. 123, 123 (2005) (defining universal civil jurisdiction as "the principle under which civil proceedings may be brought in a domestic court irrespective of the location of the unlawful conduct and irrespective of the nationality of the perpetrator or the victim, on the grounds that the unlawful conduct is a matter of international concern").

315. *See generally* Born, *supra* note 202, at 11–22 (conducting comparative assessment of jurisdiction determinations among national courts facing transnational controversies).

Existing international law does not explicitly authorize universal civil jurisdiction.[316] However, universal civil jurisdiction also is not inconsistent with international law.[317] Thus, an expansion of universal civil jurisdiction is not precluded under international law. The Global Law Commission could contribute to ensuring the enforceability of universal civil jurisdiction by helping establish and advocating for soft-law norms that eventually could become customary international law. Another way to establish universal civil jurisdiction would be via an international convention. This would not undermine the development of global law but rather would reflect the interconnectedness and interdependency of global law and international law.

A Global Law Commission and the uniform model laws promulgated by the Global Law Commission should define the types of conduct to which universal civil jurisdiction would apply. The scope of relevant conduct should be broader than the types of tort claims that have been permitted, for example, under the U.S. Alien Tort Statute.[318] In addition, it should explicitly apply to conduct by and claims against transnational corporations.[319] However, it would not have to include every form of conduct that impinges on a human right enumerated in uniform model laws as drafted by the Global Law Commission. Further, universal civil jurisdiction should not prevent tort claims from being brought in proceedings based on universal criminal jurisdiction.

CONCLUSION

One of the most important issues in human rights law today is the role of corporations. The crisis of corporate-related human rights abuses demands a response. Corporate-related activities continue to cause harm

316. *See* Kamminga, *supra* note 314, at 123–24 (claiming that "[n]o rule of international law specifically authorizes let alone obliges the exercise of universal civil jurisdiction in respect of human rights offenders").

317. *See id.* at 124 (discussing the lack of objections to U.S. courts' exercise of universal civil jurisdiction for alien tort claims); *cf.* Shaack, *supra* note 312, at 120 (claiming that "[i]nternational law authorizes universal civil jurisdiction").

318. 28 U.S.C. § 1350 (2006); Sosa v. Alverez-Machain, 542 U.S. 692, 724–25 (2004) (requiring ATCA claims "based on the present-day law of nations to rest on a norm of international character accepted by the civilized world and defined with a specificity comparable to the features of the 18th-century paradigms [the Court has] recognized," namely violation of safe conducts, infringement of the rights of ambassadors, and piracy). See generally William S. Dodge, *Which Torts in Violation of the Law of Nations?*, 24 HASTINGS INT'L & COMP. L. REV. 351 (2001) for a discussion of the types of torts that are actionable under the Alien Tort Statute. See generally Richard M. Buxbaum & David D. Caron, *The Alien Tort Statute: An Overview of the Current Issues*, 28 BERKELEY J. INT'L L. 513 (2010) for issues in Alien Tort Statute jurisprudence.

319. *See* Kiobel v. Royal Dutch Petroleum Co., 621 F.3d 111, 149 (2d Cir. 2010) (finding that Alien Tort Claims Act does not apply to corporate defendants). *See generally* David A. Dana & Michael Barsa, *Three Obstacles to the Promotion of Corporate Social Responsibility by Means of the Alien Tort Claims Act: The Sosa Court's Incoherent Conception of the Law of Nations, The 'Purposive' Action Requirement for Aiding and Abetting, and The State Action Requirement for Primary Liability*, FORDHAM ENVTL. L. REV. (forthcoming), *available at* http://ssrn.com/abstract=1711934, for a discussion of obstacles inhibiting successful claims against transnational corporations.

around the world in every area of human rights. This Article has argued that the current state-centered international human rights regime is inadequate to respond comprehensively to this pressing problem.

At least three questions are central to determining whether corporate-related human rights abuses present a problem of sufficient importance and consequence that they can be adequately addressed only with a new legal and regulatory regime. First, is there a deficit in domestic and international regulation of transnational corporations? Second, what are possible explanations for this deficit? Third, what are the effects of this deficit? This Article has attempted to answer each of these questions as well as others. However, there are many questions that remain unanswered that will be explored in more depth in a series of future articles.

Understanding corporate-related human rights abuses, their causes, and the reasons they persist requires the analysis of a series of secondary questions. These include questions that focus on existing structures as well as questions that are forward looking. What is the current system for regulating transnational corporations, who makes the rules, and who controls implementation and monitoring? What are the goals of the current system and is it achieving those goals? What are the assumptions underlying the current system and are they accurate? How does the system work and what parts are working or not working? If the system is not working, why is it not working? Are the parts that are not working important, who does this affect, and how?

What needs to be done to improve the regulation of transnational corporations? Do we need new laws, a new system, or even to change the way we theorize the regulation of transnational corporations?

This Article has made two main proposals: (1) decouple and distinguish human rights law and international human rights law, and (2) create a global human rights law system. This Article argued that it is necessary to understand human rights law as a super-category and international human rights law as a sub-category rather than as synonyms. This reclassification highlights a deficiency in the development of human rights law and creates a space in which to generate and advance other areas of human rights law. This Article then proposed global human rights law as a new area of human rights law that is well suited to address corporate-related human rights abuses.

Of course, the creation of a new regulatory framework is neither simple nor easy. It requires consideration of multiple constituencies with a wide range of—sometimes competing—interests. Proposals for increasing regulation may be met with resistance from many directions, including states that fear the loss of capital and other benefits from foreign direct investment, bureaucrats who fear the loss of power, and companies that fear increased costs. Decision makers may worry about insti-

tutional competence, administrability of rules, and overburdening existing enforcement mechanisms.

This Article made several specific proposals for the development of a global human rights law regime: a Global Law Commission, uniform model laws and regulations, and universal civil jurisdiction. The emphasis of this proposal on global law rather than international law sets it apart from many earlier attempts to comprehensively regulate the activities of transnational corporations. Although implementation of this proposal would take time and would need to be achieved in stages, forward progress can be made in all three components of this proposal simultaneously.

One possible critique of this Article stems from the assumptions it makes about the role of morality and ethics in theorizing business practices, assumptions that are disputed in the U.S. legal academy, and in fact represent a minority perspective in the United States. However, part of the value of an academic discourse is the dialogue itself and the dissection, agreement, disagreement, and refinement of ideas. Therefore, if, in the United States, this Article simply sparks further discussion about global law and Global Corporate Citizenship among U.S. legal scholars and policy makers, it will have achieved one of its purposes.

Future articles will need to respond to a number of questions about prior and existing measures and attempts in greater detail. First, what attempts have been made to account for social considerations in the regulation of transnational corporations? Second, what can we learn from prior successes and failures? Third, are global law and Law and Global Corporate Citizenship new names for existing initiatives pulled together in a systematic way or do they represent completely new frameworks or something in between? This Article and earlier articles have begun answering some of these questions.

An exploration of the strengths and weaknesses of prior and existing regulatory efforts can be facilitated by answering another set of secondary questions. These include questions that focus on an assessment of specific measures. What is the status quo in academic literature and policy making? Is Global Corporate Citizenship already influencing law and policy and, if so, where and how? Which aspects are working or not working? What other types of measures and regulation have been tried in the past? Are there successful examples that we can draw from? Are there steps that we can take to avoid past failures? Where different approaches have been used, why were they chosen, what can be gleaned from a comparison? Can models that are developed and implemented in different countries be adapted to work in other countries or at a global level?

Finally, there are three questions that are arguably central to developing a new regulatory framework built on a foundation of global law

and Law and Global Corporate Citizenship. First, what should be the values and goals of a new legal regime? Second, what form should the implementation of a new framework take and who should play a role in the implementation? Third, how will the new laws and regulations be enforced? This Article proposed initial responses to these questions, which will be explored in more depth in a series of future articles.

Moving toward a more comprehensive regulatory framework will require the analysis of a final cluster of secondary questions. These are set out in sub-groups below. Some questions aim to flesh out values and goals. Who should decide what values should inform the new framework and from what sources can we draw such values? What steps should be taken in the short-, mid-, and long-term? Are there stages or can all parts of a new system be implemented simultaneously?

Other questions aim to assess what is needed to achieve the creation of a new regulatory regime. Is the existing literature sufficient or do we need more research? What types of data and studies would be useful? What is required to garner sufficient state support and participation in a global system? Who will fund a global legal system? Will this new global law system require new institutions and laws and, if so, what form should they take?

Some questions focus on roles and responsibilities. Who would be responsible for developing a Global Corporate Citizenship approach? What institutions will be involved? Is there a role for corporations in the development of a new framework and, if so, what is it?

Further questions would address the purpose of a new regulatory system and those who would benefit from or oppose such a framework. What would be the purpose and goals of a new institution and laws? Who would benefit from the new regulatory regime? Are there groups or institutions on which the new framework would have a detrimental effect? Who would be opposed to this approach and why?

Finally, it is also necessary to address questions of administrability. How would we administer and monitor the effectiveness of this approach? What structures should be set up to carry out this evaluation? How much will this approach cost and how can we achieve sustainability? What would be some of the stumbling blocks and how could we address them?

Some people may find the conclusion of this Article with more questions than answers unsatisfying. They may be even more unsatisfied to learn that the many questions noted here are not exhaustive. However, I believe this reflects the fact that this is the beginning of a large project and this Article does not aspire or attempt to comprehensively explore all relevant questions.

As is always the case when proposing a new or different regulatory regime, one opens oneself up to critique. For example, when one suggests revisiting the relationship between human rights law and international human rights law, as is the case with this Article–one question that does or should follow is, if we ask a different question, then are we asking the right one. This Article argues that the question of whether international human rights law is able to or even should be the sole or primary system of laws to govern corporate-related human rights abuses is pivotal—although there may also be other "right" questions that should be asked.

The goals of this Article are both ambitious and modest. They are ambitious because the proposal of a new or different regulatory regime can almost always be described as ambitious. At the same time, global law is attracting increasing attention in non-national legal scholarship. Global Corporate Citizenship is not new but was developed in other fields and in practice, and enjoys the support of international institutions. Nonetheless, the proposal to develop a systematic framework for regulating transnational corporations in the area of human rights is ambitious. Therefore, the proposal set out in this Article is necessarily rudimentary and will require fleshing out over time in future articles.

OUTSOURCING DEMOCRACY: REDEFINING PUBLIC-PRIVATE PARTNERSHIPS IN ELECTION ADMINISTRATION

GILDA R. DANIELS†

INTRODUCTION

"We are left with a system in which almost every state still outsources its elections to what are actually private organizations."[1]

In 2008, during a hotly contested presidential campaign, legions of private individuals, nonprofit organizations, political parties, and candidates embarked upon canvassing, registering and assisting citizens, and monitoring voting problems in a historic presidential election.[2] Certainly, the 2000 election debacle—in which the world watched as our election process seemed to implode with butterfly ballots, hanging chads, and disproportionate disenfranchisement in the minority community—prompted extreme criticism of how the United States handles elections.[3]

Since the 2000 election, we have witnessed an increase in private participation in elections. The level of private involvement spanned from registering voters, to getting citizens to the ballot box through get out the vote campaigns ("GOTV"), assisting limited English proficient ("LEP") citizens, and monitoring Election Day activities.[4] This increased in-

---

† Assistant Professor of Law, University of Baltimore School of Law, J.D. 1990, New York University School of Law. Thanks to Michele Estrin Gilman, Henry L. Chambers and Darren Hutchinson for their helpful comments; a special note of gratitude to my Research Assistants, Anne Wilkinson and Latoya Francis-Williams, for their incredible work. This Article was supported by a summer research stipend from the University of Baltimore School of Law.

1. Phil Keisling, *To Reduce Partisanship, Get Rid of Partisans*, N.Y. TIMES, March 22, 2010, at A27, *available at* http://www.nytimes.com/2010/03/22/opinion/22keisling.html?_r=1 (former Oregon Secretary of State arguing for the elimination of primaries and stating that "[w]ith the approval of the Supreme Court, the parties have the authority to exclude independent voters or other non-members who might seriously challenge their partisan shibboleths or taboos").

2. During the 2008 election, nonpartisan organizations chronicled numerous voting irregularities in voter registration, felon disenfranchisement, long lines at the polls, poll watcher challenges, and unwarranted challenges to student voters and deceptive practices. *See, e.g., Hearing on Lessons Learned from the 2008 Election Before the Subcomm. on the Constitution, Civil Rights, and Civil Liberties of the H. Comm. on the Judiciary*, 111th Cong. 1 (2009) (statement of Tova Andrea Wang, Vice President, Research, Common Cause); *Id.* (statement of Hillary O. Shelton, Washington Bureau Chief, NAACP); *see also Protecting the Right to Vote: Oversight of the Department of Justice's Preparations for the 2008 General Election: Hearing before the S. Comm. on the Judiciary*, 110th Cong. 1 (2008) (statement of Gilda R. Daniels, Assistant Professor, University of Baltimore School of Law).

3. *See, e.g.*, Bush v. Gore, 531 U.S. 98 (2000); Bush v. Palm Beach Cnty. Canvassing Bd., 531 U.S. 70 (2000); Gore v. Harris, 773 So. 2d 524 (Fla. 2000); Edward B. Foley, *Refining the* Bush v. Gore *Taxonomy*, 68 OHIO ST. L.J. 1035 (2007); Richard L. Hasen, *The Untimely Death of* Bush v. Gore, 60 STAN. L. REV. 1 (2007).

4. For example, nonpartisan organizations, such as Election Protection, self-described as "the nation's largest non-partisan voter protection coalition," used more than 10,000 volunteers and received over 200,000 calls to its hotline during the election season and more than 80,000 calls on

volvement has been criticized on many fronts; critiques range from the involvement of the media[5] to voting machine manufacture[6] to voter registration.[7] In some instances, government has outsourced its obligations to private entities.[8]

Private partisan involvement in elections has hit troubling levels—from partisan election administrators at the highest rank to federal government officials tasked with enforcing and protecting voting rights laws.[9] The government has outsourced its authority to private groups, particularly private partisans, who use public authority for private political gain. Governments outsource in a wide variety of areas, such as education, prisons, and the military.[10] The increase in private partisan

---

Election Day. Election Protection, *Election Protection 2008: Helping Voters Today, Modernizing the System for Tomorrow: Preliminary Analysis of Voting Irregularities* 1 (2008), *available at* http://www.866ourvote.org/tools/documents/files/0077.pdf; *About Us*, ELECTION PROTECTION, *http://www.866ourvote.org/about* (last visited Oct. 10, 2010).

5. James Brown & Paul L. Hain, *Private Administration of a Public Function: The News Election Service*, 2 INT'L J. PUB. ADMIN. 389, 389 (1980) ("In the case of election administration, however, the public officials have abdicated responsibility for election night aggregation of the national Vote totals to a private organization, News Election Service, which is owned by five major news organizations. This private organization performs without a contract, without public compensation, and without supervision by public officials. It makes decisions concerning its duties according to its own criteria. The questions of responsibility and accountability have not arisen in part because of the private organization's performance record and in part because the responsibility was assumed gradually over a lengthy period without ever being evaluated as an item on the public agenda.").

6. John C. Bonifaz, *Our Voting Re-Public*, *in* REBOOTING AMERICA: IDEAS FOR REDESIGNING AMERICAN DEMOCRACY FOR THE INTERNET AGE 133, 134 (Allison Fine, Micah L. Sifry, Andrew Rasiej & Josh Levy, eds., 2008), *available at* http://rebooting.personaldemocracy.com /files/JohnBonifaz.pdf ("[A] small handful of private companies (including, but not limited to, The Election Systems & Software Company, Diebold Election Systems, now known as Premier Election Solutions, and Sequoia Voting Systems) have gained enormous profit and influence marketing their electronic voting systems to states and municipalities as the answer to the 'hanging chad' fiasco in Florida.").

7. In recent years, state governments have placed limits on private entities, particularly in the area of third party voter registration and assistance at the polls. A 2008 Brennan Center report noted that:

> Americans have run voter registration drives to register their fellow citizens to vote for decades, but until very recently often had to be deputized by the state to do so. This "deputy registrar" system left drives at the mercy of county or state officials who could deny that deputization to disfavored groups, or restrict the number of people who could take part.

Restrictions on Community-Based Voter Registration Drives, Brennan Center for Justice 1 (Aug. 4, 2008), http://brennan.3cdn.net/ca85dbcf2d2ac045ff_awm6iviui.pdf. For suggestions on modernizing the voter registration process, see, e.g., Wendy Weiser, Michael Waldman & Renée Paradis, *Voter Registration Modernization: Policy Summary*, *in* VOTER REGISTRATION MODERNIZATION: COLLECTED BRENNAN CENTER REPORTS AND PAPERS 9, 9–15 (Wendy Weiser ed., 2008), *available at* http://brennan.3cdn.net/329ceaa2878946ba17_kwm6btu6r.pdf.

8. *See Developments in the Law—State Action and the Public/Private Distinction*, 123 HARV. L. REV. 1248, 1250–51 (2010) [hereinafter *Public/Private Distinction*]; *see also* Jody Freeman, *The Private Role in Public Governance*, 75 N.Y.U. L. REV. 543, 547 (2000).

9. *See generally* Daniel P. Tokaji, *The Future of Election Reform: From Rules to Institutions*, 28 YALE L. & POL'Y REV. 125 (2009) (suggesting ways to address partisan election administrators in the next wave of election reform); Pamela S. Karlan, *Lessons Learned: Voting Rights and the Bush Administration*, 4 DUKE J. CONST. L. & PUB. POL'Y 17 (2009) (arguing that the Bush administration's Department of Justice Civil Rights Division displayed partisan objectives in the enforcement of voting rights).

10. *See infra* Part I.

groups' participation in unique government roles follows the recent trend toward outsourcing and privatization in all aspects of government operations.[11] While private involvement often garners support for government ideals, such as promoting and encouraging voter participation, the use of private partisan voter challenges have different objectives: promoting one's candidate,[12] voter suppression,[13] or the less sinister but just as effective voter frustration and confusion.[14]

Private partisan activity, primarily voter caging[15] and voter challenges,[16] tends to undermine the very right it was designed to protect—namely, free and equal voter participation.[17] Since the states determine

---

11. *See Developments in the Law—State Action and the Public/Private Distinction*, 123 HARV. L. REV. 1248, 1250–51 (2010) [hereinafter *Public/Private Distinction*]; *see also* Jody Freeman, *The Private Role in Public Governance*, 75 N.Y.U. L. REV. 543, 547 (2000).

12. *See* Tiryak v. Jordan, 472 F. Supp. 822, 824 (E.D. Pa. 1979) ("The poll-watcher performs a dual function on Election Day. On the one hand, because he is designated and paid by a political party, his job is to guard the interests of that party's candidates. On the other hand, because exercise of his authority promotes an honest election, the poll-watcher's function is to guard the integrity of the vote. Protecting the purity of the electoral process is a state responsibility and the poll-watcher's statutory role in providing that protection involves him in a public activity, regardless of his private political motive.").

13. *See* Democratic Nat'l Comm. v. Republican Nat'l Comm., 671 F. Supp. 2d 575, 578–79 (D. N.J. 2009). After the 2008 election, the Democratic National Committee ("DNC") and the Republican National Committee ("RNC") sought to modify a consent decree limiting the RNC's "ballot security measures." *Id.* at 579. The court noted that, "The RNC produced thousands of pages of newspaper articles and other sources documenting alleged incidents of voter fraud over the past 27 years, while the DNC did the same with respect to voter intimidation." *Id.* at 578.

14. *See, e.g.*, WENDY WEISER & VISHAL AGRAHARKAR, BRENNAN CENTER FOR JUSTICE, BALLOT SECURITY AND VOTER SUPPRESSION: INFORMATION CITIZENS SHOULD KNOW 1 (2010), *available at* http://brennan.3cdn.net/e2d20eec819018aa49_xpm6iixxd.pdf.

15. *See* Chandler Davidson et al., *Vote Caging as a Republican Ballot Security Technique*, 34 WM. MITCHELL L. REV. 533, 537–38 (2008) ("[V]ote caging is defined as a three-stage process designed to identify persons in another political party or faction whose names are on a voter registration list, but whose legal qualification to vote is dubious, and then to challenge their qualification either before or on Election Day. Ostensibly, caging is an attempt to prevent voter fraud. In practice, it may have the effect of disenfranchising voters . . . ."). There are a few theories on the origin of the term, one being that caging lists seem to have originated in the direct-mail fund-raising business. Because the term is so new, a formal definition cannot be found in a traditional print dictionary but can be found on websites such as Wikipedia or the Double-Tongued Dictionary. *Id.* at 537. However, according to Wikipedia, fundraisers will hire third-parties to handle the processing of responses to direct mail, which may include "processing payments, compiling product orders, correcting recipient addresses, processing returned mail, providing lockbox services and depositing funds received into the hiring organization's bank account, and all of the associated data entry for each of these services." *Caging (Direct Mail)*, WIKIPEDIA, http://en.wikipedia.org/wiki/Caging_(direct_mail) (last modified July 17, 2010). The term "caging" itself may be a derivative of the financial teller cage, since a number of operations related to lockbox services involve the control and protection of funds. Another explanation involves the "old postal 'cages,' the hundreds of cubby holes that fronted postal desks for sorting." Ed Brayton, *GOP Has a History of Voter 'Caging,' According to Democrats' Lawsuit*, THE MICH. MESSENGER (Sept. 17, 2008, 8:10 PM), http://michiganmessenger.com/4479/gop-has-a-history-of-voter-caging-according-to-democrats-lawsuit.

16. *See infra* Parts I–IV.

17. *See* Mont. Democratic Party v. Eaton, 581 F. Supp. 2d 1077, 1082 (D. Mont. 2008) ("Eaton and the Montana Republican Party are abusing the process the State of Montana has provided to ensure the accuracy of voter rolls (indeed, they are using the process designed to protect the integrity of the political process to undermine it) . . . .").

the time, places, and manner for elections,[18] they also develop the rules for allowing voter challenges. No federal law governs voter challenges;[19] as such, laws explaining who can challenge and what authority they are given vary from state to state. Most states call the individuals who conduct voter challenges "poll watchers,"[20] "challengers,"[21] and/or "observers"[22] and allow them to observe the casting of ballots, the counting of absentee ballots, and in some instances, challenge the poll workers handling of the process.[23]

The often related voter caging or vote caging occurs when a political party sends registered mail to addresses of registered voters.[24] If the mail is returned as undeliverable—because, for example, the voter refuses to sign for it, the voter isn't present for delivery, or the voter is homeless—the party uses that fact to challenge the registration, arguing that because the voter could not be reached at the address, the registration is fraudulent.[25] The private group then prepares a caging list, which "is a list or database of addresses, updated after a mailing program is completed, with notations on responses received from recipients with corrections for addresses that mail has been returned as" undelivered or

18. U.S. CONST. art. I, § 4, cl. 1 ("The Times, Places and Manner of holding Elections for Senators and Representatives, shall be prescribed in each State by the Legislature thereof; but the Congress may at any time by Law make or alter such Regulations, except as to the Place of chusing Senators.").

19. *See* Frank Emmert, Chirstopher Page & Antony Page, *Trouble Counting Votes? Comparing Voting Mechanisms in the United States and Other Selected Countries*, 41 CREIGHTON L. REV. 3, 30–34 (2007). The authors argue that some of the problems in past U.S. elections occurred because of lack of uniform and transparent laws guiding elections at the national level and that the United States needs to make improvements in its election procedures and laws. *Id.*

20. *See, e.g.*, ARK. CODE ANN. § 7-5-312(a)(2) (2009) ("poll watcher" or "representative"); CAL. ELEC. CODE § 19362 (West 2003) ("poll watchers").

21. *See, e.g.*, LA. REV. STAT. ANN. § 18:427 (2004) ("watchers"); MICH. COMP. LAWS ANN. §§ 168.730–.734 (West 2005) ("challenger"); MO. REV. STAT. §§ 115.105–.107 (1997) ("challenger" and "watcher"); N.H. REV. STAT. ANN. §§ 666:4–:5 (2010) ("challenger"); N.J. STAT. ANN. § 19:7-5 (West 1999) ("challengers"); N.C. GEN. STAT. § 163-45 (2010) ("observers"); OHIO REV. CODE ANN. § 3505.21 (West 2010) ("observers").

22. *See, e.g.*, VA. CODE ANN. §§ 24.2-604(I) (2010) ("observers"); WASH. REV. CODE ANN. §§ 29A.04.630 (West 2005) ("political parties" and "elections observers").

23. A few states explicitly allow challengers to question an election official's performance. *See, e.g.*, LA. REV. STAT. ANN. § 18:427 (2004) (allowing a watcher to contest "any infraction of the law"); MICH. COMP. LAWS ANN. § 168.730 (West 2005) (allowing challengers to question election workers and procedures).

24. *See* Davidson et al., *supra* note 15, at 538.

25. *See* Jo Becker, *GOP Challenging Voter Registrations: Civil Rights Groups Accuse Republicans of Trying to Disenfranchise Minorities*, WASH. POST, Oct. 29, 2004, at A5. Jo Becker, a reporter for the Washington Post, stated that in 1981 "the [RNC] sent letters to predominantly" African-American neighborhoods in New Jersey. *Id.* When 45,000 of these letters were returned as undelivered, the RNC attempted to challenge the voters whose mail was returned and have his or her name removed from voter rolls. *Id.* In 1986, the RNC attempted to do the same to 31,000 people in Louisiana. *Id.* Similarly in 2004, the RNC challenged voters in Ohio for undeliverable mail. *Id.* In that same year, Republicans also sent mail to 130,000 voters in Philadelphia neighborhoods, whose residents are predominantly black and Democratic. *Id.* In reference to attempted voter purges and influencing voter rolls, Republicans said their actions had nothing to do with partisanship and had "nothing to do with race." *Id.*

forwarded.[26] A political party then uses the "caging list" to challenge the validity of a voter's registration.[27] In order for the voter's provisional ballot to be counted, the voter may have to defend and prove his or her eligibility.[28]

Voter caging and voter challenges have been used to target racial minorities and to otherwise disrupt the voting process. The use of abusive voter challenges was evident after the 2008 Presidential election when voters in New York suffered suppression of the right to vote.[29] Attorneys present at the polls "documented efforts to suppress minority voters."[30] The attorneys reported:

> Some of the suppressive actions . . . included an armed man with police shield escorting around alleged poll watchers throughout various polling sites during voting hours; repeated blanket challenges to minority voters at a particular polling location . . . and a widespread challenge to nearly 6,000 Democratic voters who allegedly did not live where their voter registration information claimed they did.[31]

A *New York Times* reporter characterized the suppression activities as "chaotic."[32] The New York legislature subsequently introduced legislation to outlaw these practices.[33] In another example, while the Republican National Committee (RNC) was defending a lawsuit in Louisiana, it

---

26. *Caging (Direct Mail)*, WIKIPEDIA, http://en.wikipedia.org/wiki/Caging_(direct_mail) (last modified July 17, 2010). Caging lists seem to have originated in the direct-mail fund-raising business. *See id.* Fundraisers will hire third-parties to handle the processing of responses to direct mail, "which may include processing payments, compiling product orders, correcting recipient addresses, processing returned mail, providing lockbox services and depositing funds received into the hiring organization's bank account, and all of the associated data entry for each of these services." *Id.* The term caging "may be a derivative of the financial teller cage, since a number of operations related to lockbox services involve the control and protection of funds." *Id.*

27. *See* Becker, *supra* note 25. During the 2008 presidential election, the Republican Party in Macomb County, Michigan proposed using foreclosure lists to challenge voters at the polls. Eartha Jane Melzer, *Lose Your House, Lose Your Vote*, MICH. MESSENGER (Sept. 10, 2008, 6:42 AM), http://michiganmessenger.com/4076/lose-your-house-lose-your-vote. The party quickly backtracked from its proposed strategy, but even the mere proposition that foreclosures would be used to intimidate and potentially disenfranchise voters is reprehensible and problematic. Eartha Jane Melzer, *Republicans Recant Plans to Foreclose Voters but Admit Other Strategies*, MICH. MESSENGER (Sept. 11, 2008, 3:41 PM), http://michiganmessenger.com/4231/republicans-recant-plans-to-foreclose-voters-but-admit-other-strategies.

28. *See, e.g.*, Act of June 20, 2005, ch. 277, § 24, 2005 Fla. Laws 2614, 2643 (allowing a voter to provide evidence of eligibility to the county canvassing board and providing that "a provisional ballot shall be counted unless the canvassing board determines by a preponderance of the evidence that the person was not entitled to vote"). State laws provide the opportunity for private actors to "challenge" a voter's eligibility to participate in the election process. *See, e.g.*, Act of June 20, 2005, ch. 277, § 27, 2005 Fla. Laws 2614, 2646. In order to exert a challenge, most states do not require any threshold finding of ineligibility. *See, e.g., id.* (requiring only a "good faith" belief).

29. *See* S. 2554, 232d Ann. Legis. Sess., Reg. Sess. (N.Y. 2009), *available at* http://www.open.nysenate.gov/legislation/bill/S2554B (giving the justification of the original bill).

30. *Id.*

31. *Id.*

32. *Id.*

33. *See id.*

released a memo from one of its party officials indicating its intentions to suppress the minority vote.[34] The pertinent part of the memo read: "I would guess that this program will eliminate at least 60–80,000 folks from the rolls . . . . If it's a close race . . . which I'm assuming it is, this could keep the black vote down considerably."[35]

Additionally, in a hotly contested 2003 mayoral election in Philadelphia, involving an African American incumbent, voters in predominantly African American areas of the city were systematically challenged by men carrying clipboards and driving a fleet of approximately 300 sedans with decals designed to look like law enforcement insignia.[36] During that election, local judges heard numerous allegations of voter intimidation and harassment and ultimately issued orders in two city polling places prohibiting Republican poll watchers from requesting registration and identification materials from voters.[37]

Voter challenges and vote caging serve as yet another new millennium mechanism to disenfranchise minority voters.[38] Historically, the disenfranchisement devices were clearly race based.[39] More contemporaneously, however, partisan actors have proclaimed that the challenges were meant to prevent voter fraud and target party affiliation, more so than race. Indeed, recent assertions merely use "partisan" as a proxy for race.[40] In spite of these assertions, whether the actions are motivated by

---

34. *See* Thomas B. Edsall, *'Ballot Security' Effects Calculated: GOP Aide Said Louisiana Effort 'Could Keep the Black Vote Down'*, WASH. POST, Oct. 25, 1986, at A1.

35. *Id.* (omissions in original).

36. PEOPLE FOR THE AM. WAY & NAACP, THE LONG SHADOW OF JIM CROW: VOTER INTIMIDATION AND SUPPRESSION IN AMERICA TODAY 4–6 (2004), *available at* http://www.pfaw.org/sites/default/files/thelongshadowofjimcrow.pdf.

37. Clea Benson, Cynthia Burton & Jacqueline Soteropoulos, *Chaotic Day Marks End of Tense Campaign*, PHILA. INQUIRER, Nov. 5, 2003, at A1.

38. This Article is the third that examines various new millennium mechanisms that are used to disenfranchise minority voters. *See also* Gilda R. Daniels, *Voter Deception*, 43 IND. L. REV. 343 (2010) (discussing the need for enforcement and penalties for voter deception); Gilda R. Daniels, *A Vote Delayed Is a Vote Denied: A Preemptive Approach to Eliminating Election Administration Legislation that Disenfranchises Unwanted Voters*, 47 U. LOUISVILLE L. REV. 57 (2008) (calling for a proactive approach to potentially disenfranchising election administration measures).

39. *See infra* note 107–09 and accompanying text.

40. The use of partisanship as a proxy for race is seen most often in the redistricting process. In *Bartlett v. Strickland*, Justice Souter wrote in his dissent:

> If districts with minority populations under 50% can never count as minority-opportunity districts to remedy a violation of the States' obligation to provide equal electoral opportunity under § 2, States will be required under the plurality's rule to pack black voters into additional majority-minority districts, contracting the number of districts where racial minorities are having success in transcending racial divisions in securing their preferred representation. The object of the Voting Rights Act will now be promoting racial blocs, and the role of *race* in districting decisions as a *proxy* for political identification will be heightened by any measure.

129 S. Ct. 1231, 1250 (2009) (Souter, J., dissenting) (emphasis added); *see also* Bush v. Vera, 517 U.S. 952, 958 (1996) (plurality opinion) ("Strict scrutiny does not apply merely because *redistricting* is performed with consciousness of race. . . . Electoral district lines are 'facially race neutral,' so a more searching inquiry is necessary before strict scrutiny can be found applicable in *redistricting* cases than in cases of 'classifications based explicitly on race.'" (quoting Adarand Constructors, Inc. v. Peña, 515 U.S. 200, 213 (1995) (Thomas, J., concurring in judgment) (emphasis added)).

racial animus or partisanship, the result is the same. Racial minorities are targeted for vote challenges and those private actions constitute unconstitutional state action.

When private partisans are allowed to challenge voters in the polling site, they are engaged in determining voter eligibility, which is a governmental function.[41] Moreover, when private partisan organizations are involved in the electoral process, their allegiance to a certain candidate or outcome is different than the governmental goal of ensuring access and applying the laws impartially, regardless of who the voter is perceived to support. Allowing private partisans to participate in voter eligibility decisions—which occur with the use of voter challenges—can confuse voters and blur the lines between appropriate public and private behavior.[42] In the election administration context, little has been written to assess whether the government has given too much of its authority to private actors at the expense of the exercise of the fundamental right to vote.[43] The abusive and intimidating use of voter challenges, however, has prompted legislation prohibiting or altering the practice.[44] Accordingly, the state statutes governing voter challenges and the private actor's access to Election Day voters require a closer look to ensure that they properly protect against discriminatory actions. A void currently exists in an analysis of this type of election administration activity. This Article attempts to fill that void with clear and thorough analysis of how voter challengers act as adjuncts of the state when issuing voter challenges and how courts should treat their actions as "state action."[45]

This Article explores the level of public-private "partnership" in election administration and takes a particular look at private partisan's manipulation of state voter challenge laws. This Article argues that the private partisans use of government authority in determining voter eligibility constitutes state action and could subject not only the private partisan—but also the state—to liability under the Fourteenth Amendment's Equal Protection Clause. In Part I, this Article explores the use of outsourcing government functions to private organizations and discusses how the public-private divide has a different dynamic than other outsourced areas, such as education or welfare reform, when juxtaposed against

41. *See* Tiryak v. Jordan, 472 F. Supp. 822, 824 (E.D. Pa. 1979) ("No activity is more indelibly a public function than the holding of a political election . . . .").

42. *See* Gillian E. Metzger, *Privatization as Delegation*, 103 COLUM. L. REV. 1367, 1369 (2003); *Public/Private Distinction*, *supra* note 12, at 1250–51.

43. *See, e.g.*, Barry H. Weinberg & Lyn Utrecht, *Problems in America's Polling Places: How They Can Be Stopped*, 11 TEMP. POL. & CIV. RTS. L. REV. 401, 434–36 (2002) ("[T]he states have delegated to the counties the responsibility for conducting the election and maintaining order in the polls, and by doing so, the states have abdicated their responsibility for preventing bad things from happening to voters in the polls on election day.").

44. *See, e.g.*, S. 6134, 59th Leg., Reg. Sess. (Wash. 2006) (authorizing canvassing boards to impose civil penalties for improper voter challenges).

45. *See supra* note 38 and accompanying text.

the fundamental right to vote. Part II examines the level of race-based targeting in voter challenges and vote caging. Part III discusses legal concerns and applies the state action framework to these new millennium mechanisms and cautions that states can be held liable for private partisan voter challenges. Scholars have argued both sides of the issue regarding whether poll watchers engage in state action.[46] This Article provides clarity to this discussion. Part IV recommends the elimination of voter challenges, or in the alternative, suggests a means to ensure that the private partisan's role does not cede government authority.

I. THE CEDING OF AUTHORITY: PRIVATE INVOLVEMENT IN PUBLIC FUNCTIONS

"Privatization is now virtually a national obsession. Hardly any domestic policy issue remains untouched by disputes over the scope of private participation in government . . . ."[47]

The public-private divide has become a blur in many areas. Private actors perform government functions in, *inter alia*, the military, education, welfare reform, and public safety.[48] Arguably, privatization is fueled by two primary arguments: efficiency and lower costs and a "democratizing" effect that returns the power to local communities.[49] One can readily make an argument that the government can cut its costs if it outsources the welfare benefits program to a private company, creates a charter school, or allows a private company to build and run a prison. For example, an argument for privatizing education centers on providing a choice for students from low-income families to apply public school funds to private school tuition.[50] Further, it is argued that the competition

46. *Compare* Jason Belmont Conn, *Of Challengers and Challenges*, 37 U. TOL. L. REV. 1021, 1032–33 (2006) (exploring the role of and the decisions of poll watchers in Ohio during the 2004 presidential election, and maintaining that the poll watchers were indeed state actors by virtue of their engaging in poll challenging, which is traditionally a state action), *with* Heather S. Heidelbaugh, Logan S. Fisher & James D. Miller, *Protecting the Integrity of the Polling Place: A Constitutional Defense of Poll Watcher Statutes*, 46 HARV. J. ON LEGIS. 217, 233–34 (2009) (arguing that placing poll watchers in polling sites is not state action but rather a licensing process).

47. Metzger, *supra* note 42, at 1369.

48. *See, e.g.*, Michele Estrin Gilman, *Legal Accountability in an Era of Privatized Welfare*, 89 CALIF. L. REV. 569, 571–73 (2001); Kenneth J. Saltman, *Putting the Public Back in Public Schooling: Public Schools Beyond the Corporate Model*, 3 DEPAUL J. FOR SOC. JUST. 9, 9–10 (2009) ("I consider how the corporatization of public schools redistributes economic control and cultural control from the public to private interests. I argue that these intertwined redistributions of power undermine public democracy (the possibilities for the development of a more participatory and deeper democracy), just social transformation, and critical citizenship while exacerbating material and symbolic inequality."); Ian Traynor, *The Privatisation of War*, THE GUARDIAN (Dec. 10, 2003, 2:28 AM), http://www.guardian.co.uk/world/2003/dec/10/politics.iraq/print.

49. *See* Gilman, *supra* note 48, at 596 ("Privatization advocates . . . contend that private companies can deliver services with greater efficiency and innovation than government at a lower cost. . . . Yet another strand of the privatization movement sees privatization as a democratizing force that returns power from the government to local communities and their mediating institutions, such as churches, neighborhoods, and voluntary organizations, which are better situated to address a community's needs.").

50. *See* Julie Huston Vallarelli, Note, *State Constitutional Restraints on the Privatization of Education*, 72 B.U. L. REV. 381, 395–96 (1992).

between private and public schools will create better schools for all students regardless of income.[51] Additionally, the private school's ability to provide a quality education is governed by its capacity to provide a competitive product.[52] Arguably, "privatization of welfare converges 'the free market ideology of the right and the citizen participation/empowerment objectives of the left.'"[53] While each of these examples are subject to strong criticism,[54] including whether the law can hold private companies accountable in the same manner as the government,[55] they tend to fit neatly into the costs-democratizing paradigm.

Most scholarship has focused on private contracts for public functions, such as Medicare and Medicaid services or building prisons and education.[56] In these service contracts, the debate has raged over whether the government has granted too much power to private parties and whether the public benefits from the agreement.[57] Nonetheless, private contractors, particularly not-for-profit entities, presumably have a similar goal as the government, i.e., to improve services or outcomes, such as improving the educational system.

Conversely, when private partisans are involved in the public function of election administration, it is somewhat of a misnomer to describe the relationship between the public governmental agencies and the private group as a partnership. Partnership is the wrong characterization, since it would normally connote two or more groups working towards the same goal.[58] In the election administration context, the government

---

51. *See* Saltman, *supra* note 48, at 27–28.

52. Vallarelli, *supra* note 50, at 384–85 ("To remain open, schools must attract student-consumers by providing equal or better services than competitors. Costs are reduced as each school responds to economic pressures and maximizes the services it can offer.").

53. Gilman, *supra* note 48, at 596.

54. *See, e.g.*, Daniel L. Hatcher, *Child Support Harming Children: Subordinating the Best Interests of Children to the Fiscal Interests of the State*, 42 WAKE FOREST L. REV. 1029, 1058–60 (2007) (arguing that allowing private companies to collect and disburse child support payments is in conflict with the best interests of the child); Saltman, *supra* note 48, at 9 (arguing that the privatization of schools is detrimental to education and democracy).

55. *See* Gilman, *supra* note 48, at 603 (analyzing the detrimental effects of privatizing welfare benefits on the recipient's due process rights).

56. *See id.* at 578–79 (examining private contracts for providing welfare).

57. *See, e.g.*, Daniel L. Hatcher, *Collateral Children: Consequence and Illegality at the Intersection of Foster Care and Child Support*, 74 BROOK. L. REV. 1333, 1377–79 (2009) (arguing for the elimination of foster care recovery programs, which are often administered by private entities, because they serve no benefit to the children in foster care).

58. The U.S. Election Assistance Commission's *Guidebook on Successful Practices for Poll Worker Recruitment, Training, and Retention* [hereinafter *Successful Practices*] lists the "Pitfalls and Challenges" of partnering with political parties to recruit poll workers as follows:

- Some political parties use Election Day poll service as a patronage job for the party faithful, not necessarily appointing those most qualified and willing to serve.
- Party representatives may be tempted to work for the success of a particular candidate.
- Political party lists may be submitted too late to be of use.
- Political parties often want members and others to serve as observers.
- Political party poll workers may want to work only in high-stakes elections and may not be reliable components of a long-term election team.

has a uniquely constitutional function to ensure that elections are held in a fair and impartial manner.[59]

Voting, however, is different from other outsourced functions—such as education and welfare.[60] Voting is a fundamental right,[61] and the United States Constitution mandates that the federal and state governments administer all aspects of elections. The democratic process is at its best when all three levels of government—federal, state, and local—work together to protect access and integrity.[62] No contractual relationship exists between the private groups and the governmental agencies, and in most instances, the private organizations are working to improve the democratic process by increasing the number of eligible voters through voter registration drives or delivering eligible voters to the polls.[63]

The political partisan, however, has a different goal in mind through his or her participation in the electoral process; to get a particular candidate elected or ballot measure passed.[64] Once the private partisans use statutory authority to achieve the goal of political gamesmanship and influence governmental decisions on voter eligibility, the government has ceded its authority to private partisans at the expense of the eligible voter.

In order to understand how private partisans are assuming governmental power, one must first understand what authority the government has under the Constitution and other federal statutes in regards to elections—as well as how the authority is divided amongst the federal, state, and local governments.

U.S. ELECTION ASSISTANCE COMM'N, GUIDEBOOK ON SUCCESSFUL PRACTICES FOR POLL WORKER RECRUITMENT, TRAINING, AND RETENTION 49-50 (2007) [hereinafter SUCCESSFUL PRACTICES], *available at* http://www.eac.gov/assets/1/AssetManager/Successful%20Practices%20for%20Poll%20Worker%20Recruitment%20Section%201%20Recruitment.pdf.

59. *See* Burson v. Freeman, 504 U.S. 191, 208–09 (1992) ("[A] government has . . . a compelling interest in securing the right to vote freely and effectively . . . .").

60. *See* Burdick v. Takushi, 504 U.S. 428, 433 (1992) ("It is beyond cavil that 'voting is of the most fundamental significance under our constitutional structure.'" (quoting Ill. Bd. of Elections v. Socialist Workers Party, 440 U.S. 173, 184 (1979))).

61. *See* Yick Wo v. Hopkins, 118 U.S. 356, 370 (1886) (proclaiming that voting is regarded as a fundamental political right).

62. *See generally* Jocelyn Friedrichs Benson, *One Person, One Vote: Protecting Access to the Franchise Through the Effective Administration of Election Procedures and Protections*, 40 URB. LAW. 269 (2008) (discussing ways that public and private entities can work together to ensure fair elections, focusing on advancing participation in the language minority community and Section 203 of the Voting Rights Act).

63. *Id.* at 278.

64. *See* Tiryak v. Jordan, 472 F. Supp. 822, 824 (1979) ("The poll-watcher performs a dual function on Election Day. On the one hand, because he is designated and paid by a political party, his job is to guard the interests of that party's candidates. On the other hand, because exercise of his authority promotes an honest election, the poll-watcher's function is to guard the integrity of the vote.").

A. The Federal Government

The federal, state, and local governments share in the administration of elections.[65] On the federal level, government agencies enforce various voting rights statutes[66] and Constitutional amendments.[67] Notwithstanding state authority to develop election administration laws governing "The Time[], Place[] and Manner of elections," Congress—under the Elections Clause—maintains authority to "make or alter" state regulation of federal elections.[68]

It is widely understood that the Constitution contemplated that federal and state governments would coordinate election administration issues, with any federal decisions reigning supreme.[69] Recent cases interpreting the Elections Clause reinforce Congress's broad authority to regulate all aspects of elections.[70] Accordingly, Congress has the power to develop and supersede state election regulations. Additionally, in *Smiley v. Holm*,[71] the Court noted that Congress's Elections Clause power allows it to "supplement . . . state regulations or . . . substitute its own."[72]

---

65. However, elections are primarily a local government function. *See* Note, *Toward A Greater State Role In Election Administration*, 118 HARV. L. REV. 2314, 2323 (2005) [hereinafter *Greater State Role*]; *see also* Weinberg & Utrecht, *supra* note 41, at 434–36 ("If the states do not assume the responsibility for conducting effective elections when the counties fail to do so, then the United States Congress should consider whether federal civil rights voting laws should be expanded to include the deprivations of voting rights at the polls . . . and should consider similarly expanding the unquestionably successful federal observer program.").

66. Federal agencies enforce various statutes that pertain to voting, such as the Voting Rights Act of 1965, the National Voter Registration Act, and the Help America Vote Act. 42 U.S.C. § 1973 *et seq.* (2006); 42 U.S.C. § 15511 (2006); 42 U.S.C. § 1973gg-9 (2006). Additionally, the Help America Vote Act spawned the Election Assistance Commission, which provides guidance and best practices for state and local election administrations and provides funds for improving the election process. 42 U.S.C. § 15381 (2006).

67. U.S. CONST. amend. XIV, § 1 ("No State shall . . . deny to any person within its jurisdiction the equal protection of the laws."); U.S. CONST. amend. XV, §§ 1–2 ("The right of citizens of the United States to vote shall not be denied or abridged by the United States or by any State on account of race, color, or previous condition of servitude. The Congress shall have power to enforce this article by appropriate legislation.").

68. U.S. CONST. art. I, § 4, cl. 1 ("The Times, Places and Manner of holding Elections for Senators and Representatives, shall be prescribed in each State by the Legislature thereof; but the Congress may at any time by Law make or alter such Regulations, except as to the Places of chusing Senators."). Congress can regulate the elections of Representatives and Senators. United States v. Gradwell, 243 U.S. 476, 482 (1917); *Ex parte* Siebold, 100 U.S. 371, 383–84 (1879); United States v. Manning, 215 F. Supp. 272, 287 (W.D. La. 1963); Commonwealth *ex rel.* Dummit v. O'Connell, 181 S.W.2d 691, 694 (Ky. 1944).

69. *Siebold*, 100 U.S. at 384 ("[T]he power . . . may be exercised as and when Congress sees fit to exercise it. When exercised, the action of Congress, so far as it extends and conflicts with the regulations of the State, necessarily supersedes them. This is implied in the power to 'make or *alter*.'").

70. Cook v. Gralike, 531 U.S. 510, 523–24 (2001) ("[The Elections Clause] encompasses matters like 'notices, registration, supervision of voting, protection of voters, prevention of fraud and corrupt practices, counting of votes, duties of inspectors and canvassers, and making and publication of election returns.'" (quoting Smiley v. Holm, 285 U.S. 355, 366 (1932))).

71. 285 U.S. 355 (1932).

72. *Id.* at 366–67. In *Smiley v. Holm*, Chief Justice Hughes wrote:

Together, the federal and state governments have the constitutional authority to regulate the election process, which includes voter eligibility. While Congress has used its constitutional power to implement laws that govern aspects of election administration,[73] it is the states' responsibility to implement those laws in a nondiscriminatory and fair manner.[74]

B. The State Government

In election administration matters, state and local governments bear the lion's share of the responsibility. Article I, Section IV of the United States Constitution grants the ability to administer elections to the states,[75] which provides extensive obligations to ensure that elections are conducted impartially and accurately.[76]

Perhaps one of the most powerful and important responsibilities of state administrators is to make an accurate and almost instantaneous determination of the eligibility of a citizen to cast a ballot. In most instances, a voter ID, signature, or other form of verification is all that is needed.[77] When a voter is challenged, the process changes tremendously, shifting the power from the state to the challenger and placing the voter on the defensive. The local administrator is tasked with the on the spot determination and the responsibility of communicating any problems to the county and state administrators, who, if needed, will also communicate with federal government agencies for assistance. Accordingly, federal offices also communicate any problems with the state, and local election administrators; in some instances, correcting for clear discriminatory practices, and in others, leaving the voter stranded on the road to full political participation.[78]

> It cannot be doubted that these comprehensive words embrace authority to provide a complete code for congressional elections, not only as to times and places, but in relation to notices, registration, supervision of voting, protection of voters, prevention of fraud and corrupt practices, counting of votes, duties of inspectors and canvassers, and making and publication of election returns; in short, to enact the numerous requirements as to procedure and safeguards which experience shows are necessary in order to enforce the fundamental right involved.

Id. at 366.

73. Congress utilized its Elections Clause authority when it enacted the NVRA and HAVA. *See* U.S. CONST. art. I, § 4, cl. 1.

74. The National Voter Registration Act demands that its measures are implemented in a nondiscriminatory manner. 42 U.S.C. § 1973gg-6(b)(1) (2006).

75. U.S. CONST. art. I, § 4, cl. 1; Storer v. Brown, 415 U.S. 724, 730 (1974); *Greater State Role*, *supra* note 65, at 2323.

76. *See Storer*, 415 U.S. at 730. *See generally* J. Kenneth Blackwell & Kenneth A. Klukowski, *The Other Voting Right: Protecting Every Citizen's Vote by Safeguarding the Integrity of the Ballot Box*, 28 YALE L. & POL'Y REV. 107 (2009); Jocelyn Friedrichs Benson, *Democracy and the Secretary: The Crucial Role of State Election Administrators in Promoting Accuracy and Access to Democracy*, 27 ST. LOUIS U. PUB. L. REV. 343, 355–61 (2008); Raquel A. Rodriguez, *Reflections of Another* Bush v. Gore *Lawyer*, 64 U. MIAMI L. REV. 631, 636 (2010).

77. *See, e.g.*, *Voter Identification Requirements*, NAT'L CONFERENCE OF STATE LEGISLATURES, http://www.ncsl.org/LegislaturesElections/ElectionsCampaigns/StateRequirementsforVoterID/tabid/16602/Default.aspx#in (last updated Sept. 30, 2010).

78. *Voting Rights Act: Sections 6 and 8 — The Federal Examiner and Observer Program:*

Most states refer to the persons primarily responsible for the Election Day experience as poll workers.[79] The poll workers, *inter alia*, open the polling sites, welcome voters, and provide the necessary voting materials—such as actual ballots or access to the electronic system employed in the jurisdiction.[80] Once the polls close, the poll workers must secure the actual ballots and/or the voting machines and deliver the machines and ballots to the central facility for the official Election Day count.[81] For all their hard work, poll workers are only paid approximately $100 for a more than 12-hour day in some cases.[82]

On the local level, poll workers protect the overall integrity of the voting process. A well-trained poll worker is aware of state and federal election regulations, as well as how to fairly administer and handle any potential disputes. Poll workers must consistently apply fair and egalitarian principles in a nondiscriminatory way and avoid the appearance of impropriety. They must also know clearly how to administer the ballot and provisional ballots, determine voter eligibility, and have a superior familiarity with voting procedures and voting technology.[83] In this new millennium, particularly with the advent of the electronic voting machine, computer literacy is crucial to a smooth voting process.

The public and private entities must work together in order to facilitate a smooth election process. It has become evident, however, that not all entities are working toward the same goal.[84] The unique placement of private partisans inside the polling place with the authority to challenge voters—and as such, determine whether a person is eligible to vote—aligns the private actions with government authority.[85]

Hearing Before the Subcomm. on the Constitution of the H. Comm. on the Judiciary, 109th Cong. 17–19 (2005) (statement of Barry H. Weinberg, Former Deputy Chief and Acting Chief, Voting Section, Civil Rights Division, U.S. Dep't. of Justice) (stating that federal observers are the "eyes and ears of the Justice Department"). For a discussion of the relationship between federal observers and the election process, see *infra* Part IV.C.

79. *See, e.g.*, CAL. ELEC. CODE § 12302 (West 2010) (allowing the state to recruit student poll workers). In New York, they are referred to as "inspectors" and "poll clerks." N.Y. ELEC. LAW § 3-420 (McKinney 2007). In Alabama, poll workers are referred to as "inspectors" and "clerks." ALA. CODE § 17-8-5 (2007).

80. *See* SUCCESSFUL PRACTICES, *supra* note 58, at 7.

81. *See generally id.*

82. In Alabama, clerks and inspectors are paid at least $75 and $100 per day respectively for statewide elections. ALA. CODE § 17-8-12 (2007); *cf.* N.Y. ELEC. LAW § 3-420 (McKinney 2007) (poll workers are paid $200 for working on election day and an additional amount for attending poll worker training); WIS. STAT. § 7.03(1)(a) (2004) (election officials are paid "a reasonable daily compensation"). In Madison, Wisconsin, "a reasonable daily compensation" is presently $11.66 per hour. MADISON CITY CLERK'S OFFICE, SERVING AS AN ELECTION OFFICIAL IN THE CITY OF MADISON, *available at* http://www.cityofmadison.com/election/pollWorkers/documents/ElectionOfficialBrochure.pdf (last visited Oct. 15, 2010).

83. *See* SUCCESSFUL PRACTICES, *supra* note 58, at 7.

84. *See* Emmert et al., *supra* note 19, at 8 (detailing the ineffective administration of elections in the United States).

85. *See id.* at 24–25.

C. Voter Challenge/Poll Watcher Statutes

State statutes determine who may remain inside a polling site to observe the electoral process.[86] Poll watcher eligibility is determined by state laws that allow persons of various positions, i.e., political party, political candidate, or concerned citizen, to view the actual voting process.[87] Some statutes are extremely liberal, allowing "[a]ny member of the public," except a candidate, to observe the election,[88] while others limit the observers to voters and election officials.[89]

Like most states, Florida law allows one poll watcher from each party and one poll watcher representing each candidate to view the voting process.[90] Further, a Florida poll watcher may challenge any voter's right to vote if done "in good faith,"[91] provided the challenger signs an oath that details the challenger's name, address, political affiliation, and reason for the challenge.[92] Once offered, the challenged voter is allowed to cast a provisional ballot.[93] Although filing a frivolous challenge is a first-degree misdemeanor in Florida, an exception is given for electors or poll workers making the challenge "in good faith."[94]

The process of issuing a challenge is rather involved. If a challenge is offered, the poll workers then engage in a time-consuming process, which takes them away from the business of processing eligible voters.[95] In some states, the challenger has to make his challenge in writing on a form that is provided for the purpose of challenging, and the challenged voter must respond with an affidavit.[96] In most instances, the challenged

---

86. *See, e.g.*, ALA. CODE § 17-8-7 (2007); FLA. STAT. § 102.031(3) (2008). Scholars have argued that states have failed to undertake their responsibilities in this arena by delegating the duty to maintain order at elections to counties. *See* Weinberg & Utrecht, *supra* note 41, at 434 (arguing that the states need to use their authority to regulate the activity of elections today to stop future problems).

87. *See, e.g.*, ALA. CODE § 17-8-7 (2007); FLA. STAT. § 102.031(3) (2008).

88. *See, e.g.*, WIS. STAT. § 7.41(1) (2004) ("Any member of the public may be present at any polling place . . . except a candidate whose name appears on the ballot at the polling place . . . .").

89. *See, e.g.*, W. VA. CODE § 3-31-37(a) (2010) ("[N]o person, other than the election officers and voters going to the election room to vote and returning therefrom, may be or remain within three hundred feet of the outside entrance to the building housing the polling place while the polls are open.").

90. FLA. STAT. § 101.131(1) (2008).

91. FLA. STAT. § 101.111(2) (2008).

92. FLA. STAT. § 101.111(1)(a) (2008).

93. FLA. STAT. § 101.111(1)(b)(1) (2008).

94. FLA. STAT. § 101.111(2) (2008) ("[E]lectors and poll watchers shall not be subject to liability for any action taken in good faith and in furtherance of any activity or duty permitted of such electors and poll watchers by law."). This exception essentially allows questionable challenges to stand without criminal repercussions. *Id.* For other examples of disenfranchising methods that avoid criminal prosecution, see generally Daniels, *supra* note 38.

95. *See, e.g.*, Crawford v. Marion Cnty. Election Bd., 553 U.S. 181, 223 (2008) (Souter, J., dissenting) (arguing that a voter ID law increases the likelihood of delay at the polls since any minor discrepancy between a voter's photo identification card and the registration may lead to a challenge).

96. *See, e.g.*, N.J. STAT. ANN. § 19:15-18.1 (West 1999) (permitting a challenged voter to establish a right to vote by a signed affidavit detailing voter qualifications or a presentation of a suitable identifying document); N.C. GEN. STAT. § 163-85(b) (2008) (requiring a challenge made in

voter may vote if it is determined by a majority vote of the poll workers that the voter is eligible.[97] Once a challenge is issued, a poll worker must resolve the challenge before the person can vote.[98] Thus, the poll workers decision regarding voter eligibility is reliant upon the private declaration of ineligibility. It then forces the voter to prove her identity. The inability to do so, could disenfranchise an eligible voter not because she is ineligible but for a myriad of reasons unassociated with eligibility, such as lack of an ID when the jurisdiction does not require identification.

D. Private Actors

Private partisan involvement in the exercise of the franchise is not new. Prior to the Civil War, most states did not require voters to prove their eligibility prior to Election Day.[99] One scholar has noted:

> Until 1888, political parties printed and distributed the ballots in each of the United States. Besides discouraging split-ticket voting and encouraging strong party organizations . . . the party ballot insured illiterates the right to vote. Nevertheless, reformers, who were more concerned with eliminating fraud than safeguarding the rights of illiterates, instituted the secret ballot . . . .[100]

Voters arrived at the polls with required documentation, including witnesses that could attest to their eligibility.[101] During this period, Republicans feared voter fraud promoted voter registration, while Democrats argued for extending voting hours.[102] During the Jim Crow era, the federal government had to step in to protect the right to vote as a result of the Black Codes, Jim Crow laws, and death threats.

As chronicled in *Burson v. Freeman*,[103] once the country moved from oral vote to paper ballots, private partisan involvement began to

---

writing, under oath, and on prescribed forms); N.C. GEN. STAT. § 163-86(d) (2008) (permitting a challenged voter to appear at a hearing in person or through affidavit).

97. *See, e.g.*, OHIO REV. CODE ANN. § 3505.20 (West 2006) (providing that the election judges shall provide a provisional ballot to a person whom a majority of the judges believe is not entitled to vote), *held unconstitutional by* Boustani v. Blackwell, 460 F. Supp. 2d 822 (N.D. Ohio 2006) (holding that provisions requiring that naturalized citizens present a certificate of naturalization upon challenge or otherwise cast a provisional ballot violates the Fourteenth Amendment).

98. *See, e.g.*, OHIO REV. CODE ANN. § 3505.19 (West 2006) (providing that if a challenged person shows up to vote and establishes to the satisfaction of the judges that he or she is entitled to vote, he or she shall be permitted to vote).

99. ALEXANDER KEYSSAR, THE RIGHT TO VOTE: THE CONTESTED HISTORY OF DEMOCRACY IN THE UNITED STATES 151 (2000). During the colonial era, the vote was reserved for white men and in some states only those men who owned property. *See id.* at 5. The later requirement of voter registration was due in part to the increase in urban areas where "voters were less likely to be known personally to election officials." *See id.* at 152.

100. J. MORGAN KOUSSER, THE SHAPING OF SOUTHERN POLITICS: SUFFRAGE RESTRICTION AND THE ESTABLISHMENT OF THE ONE-PARTY SOUTH, 1880-1910, at 51 (1974).

101. *See* KEYSSAR, *supra* note 99, at 151.

102. *See id.* at 152. In post Civil War Chicago, Republicans hired individuals to "check polling places" and offered a $300 reward to those who would assist in convicting persons who voted illegally. But, all those who were accused were acquitted. *Id.* at 153–54.

103. 504 U.S. 191 (1992).

overly influence and destroy the free exercise of the right to vote.[104] The Court wrote:

> Within 20 years of the formation of the Union, most States had incorporated the paper ballot into their electoral system. . . . Wishing to gain influence, political parties began to produce their own ballots for voters. These ballots were often printed with flamboyant colors, distinctive designs, and emblems so that they could be recognized at a distance. State attempts to standardize the ballots were easily thwarted—the vote buyer could simply place a ballot in the hands of the bribed voter and watch until he placed it in the polling box. Thus, the evils associated with the earlier *viva voce* system reinfected the election process; the failure of the law to secure secrecy opened the door to bribery and intimidation.[105]

While the practice of vote caging dates back to the 1950s, the use of race-based voter challenges can be traced to the post-Civil War era. Beginning in the 1860s, Republicans enacted vote challenging mechanisms with the mission of addressing fraud, while the Democrats responded with arguments that the Republican passed laws were "an act of hostility to the Democratic party."[106]

Still, the race-based voter targeting was both blatant and widespread. As an example, in 1865, the Florida legislature instituted measures designed for the sole purpose of denying the franchise to freed Blacks.[107] Likewise, after record black enfranchisement in 1867, white legislators adopted a statute that granted poll watchers the ability to challenge voter eligibility, stating:

> If any person offering to vote shall be challenged, as not qualified, by an inspector or by any other elector, one of the board shall declare to the person challenged the qualifications of an elector. If such person shall claim that he is qualified, and the challenge be not withdrawn, one of the inspectors shall administer to him an oath prescribed by law.[108]

In *Tiryak v. Jordan*,[109] the court found that, although the state is responsible for the administration of elections, "[t]he statutory scheme in certain instances delegates aspects of that responsibility to the political

104. *See id.* at 200–01.

105. *Id.*

106. KEYSSAR, *supra* note 99, at 155. Beginning in 1866, the Republican and Democratic parties set the stage for partisan battles over access and integrity for centuries to come. *See id.* at 152–53.

107. *See* ADVANCEMENT PROJECT, REPORT TO STATE AND LOCAL ELECTION OFFICIALS ON THE URGENT NEED FOR INSTRUCTIONS FOR PARTISAN POLL WATCHERS (2004), *available at* http://www.advancementproject.org/sites/default/files/AP-VSchallenge.pdf. Racial voting restrictions are as old as the Confederacy. *See id.* (citing FLA. CONST. of 1865, art. VI, § 1 (limiting the right to vote to white males)).

108. *Id.* at 108.

109. 472 F. Supp. 822 (1979).

parties. This delegation is a legislative recognition of 'the critical role played by political parties in the process of selecting and electing candidates for state and national office.'"[110] The "critical role" begins to impede upon government authority in the area of partisan poll watchers.

II. RACE BASED CHALLENGES AND VOTER CAGING

As discussed, *supra*, voter challenges have been used to target racial minorities. The practice of partisan voter challenges rests upon the discriminatory selection of voters based on geography, i.e., racially segregated neighborhoods and racial identification. As a first step, political operatives utilize voter caging to capture the voters and ultimately question their eligibility.

Although gaining recent prominence, caging is not new. The political tactic was used as a so-called "ballot security" measure as early as the 1950s.[111] Voter caging often goes hand in hand with voter intimidation[112] and deception tactics.[113] For instance, as discussed, *supra*, while the Republican National Committee was defending a lawsuit in Louisiana, a memo released from a party official indicated its intentions to suppress the minority vote. The relevant part of the memo read: "I would guess that this program will eliminate at least 60-80,000 folks from the rolls . . . . If it's a close race . . . which I'm assuming it is, this could keep the

110. *Id.* at 823–24 (quoting Marchioro v. Chaney, 442 U.S. 191, 195 (1979)).

111. *See* Davidson et al., *supra* note 15, at 559. Voter caging was recognized in Arizona in the 1950s and 1960s as a Republican political tool to exclude African Americans and Mexican Americans. *Id.* at 543. The Arizona example also implicated former Chief Justice William Rehnquist during his confirmation hearings. *See* JOHN W. DEAN, THE REHNQUIST CHOICE: THE UNTOLD STORY OF THE NIXON APPOINTMENT THAT REDEFINED THE SUPREME COURT 271 (2001). In 1962, the Republican National Committee expanded this approach and engaged in a national voter caging campaign named "Operation Eagle Eye," which targeted minority voters in urban areas in battleground states. *See* CHANDLER DAVIDSON ET AL., REPUBLICAN BALLOT SECURITY PROGRAMS: VOTE PROTECTION OR MINORITY VOTE SUPPRESSION—OR BOTH?: A REPORT TO THE CENTER FOR VOTING RIGHTS & PROTECTION 25 (2004). For example, in the 1960s in Chicago the RNC planned to have more than 10,000 poll-watchers for the 3,552 voting precincts. *Id.* at 26 n.5 (citing *250,000 Accredited to Watch City Polls*, CHICAGO DAILY NEWS, Nov. 2, 1964, at 1, 14).

112. Democratic Nat'l Comm. v. Republican Nat'l Comm., 671 F. Supp. 2d 575, 578–79 (D.N.J. 2009) ("Voter intimidation presents an ongoing threat to the participation of minority individuals in the political process, and continues to pose a far greater danger to the integrity of that process than the type of voter fraud the RNC is prevented from addressing by the Decree."); Jo Becker, *GOP Challenging Voter Registrations: Civil Rights Groups Accuse Republicans of Trying to Disenfranchise Minorities*, WASH. POST, Oct. 29, 2004, at A5 (introducing evidence that Republicans have "used tactics that were aimed at intimidating minority voters and suppressing their votes.").

113. *See* DAVIDSON ET AL., *supra* note 111 ("There are several noteworthy characteristics of these [vote caging] programs. They focus on minority precincts almost exclusively. There is often only the flimsiest evidence that vote fraud is likely to be perpetrated in such precincts. In addition to encouraging the presence of sometimes intimidating Republican poll watchers or challengers who may slow down voting lines and embarrass potential voters by asking those humiliating questions, these programs have sometimes posted people in official-looking uniforms with badges and side arms who question voters about their citizenship or their registration. In addition, warning signs may be posted near the polls, or radio ads may be targeted to minority listeners containing dire threats of prison terms for people who are not properly registered—messages that seem designed to put minority voters on the defensive."). *See generally* Daniels, *supra* note 38.

black vote down considerably."[114] Similarly, in 1981, "the Republican National Committee sent letters to predominantly black neighborhoods in New Jersey, and when 45,000 letters were returned as undeliverable, the [RNC] compiled a challenge list to remove those voters from the rolls."[115] More recently, prior to the 2004 presidential election the Republican National Committee used vote caging methods to compile a list of approximately 35,000 persons.[116] The RNC gathered its voter challenge list of 35,000 persons "by sending letters to registered voters in precincts with a high concentration of minorities, in this case inner-city areas in Cleveland, and recording the names of those voters for whom the letter was returned as undeliverable."[117]

In *Spencer v. Blackwell*,[118] a recent voter challenge case, African American voters sought and obtained a preliminary injunction against Ohio Secretary of State Kenneth Blackwell, seeking a court order prohibiting voter challenges inside the polling place.[119] The plaintiffs alleged that voter challenges would be used to discriminate against African American voters in Hamilton County, Ohio.[120] They asserted that the Hamilton County Board of Elections and the Hamilton County Republican Party were working together to discriminate against African American voters on Election Day.[121] The Hamilton County Republican Party had petitioned to have hundreds of poll challengers present during the 2004 presidential election to challenge voter eligibility and prevent fraud.[122] Tim Burke, Chair of the Hamilton County Board of Elections, testified that two-thirds of the poll challengers were designated for predominately African-American precincts.[123] The court found that the Ohio statutes governing voter challenges and the Secretary of State's guidance

---

114. *See* Edsall, *supra* note 34 (omissions in original).

115. Becker, *supra* note 112. The Democratic National Committee and Republican National Committee entered into a Consent Decree in 1982 in which the RNC agreed to:

> [R]efrain from undertaking any ballot security activities in polling places or election districts where the racial or ethnic composition of such districts is a factor in the decision to conduct, or the actual conduct of, such activities there and where a purpose or significant effect of such activities is to deter qualified voters from voting; and the conduct of such activities disproportionately in or directed toward districts that have a substantial proportion of racial or ethnic populations shall be considered relevant evidence of the existence of such a factor and purpose.

Consent Order at ¶ 2(e), Democratic Nat'l Comm. v. Republican Nat'l Comm., 671 F. Supp. 2d 575 (D.N.J. Nov. 1, 1982) (No. 81-3876).

116. Democratic Nat'l Comm. v. Republican Nat'l Comm., 671 F. Supp. 2d 575, 582 (D.N.J. 2009).

117. *Id.*

118. 347 F. Supp. 2d 528 (S.D. Ohio 2004).

119. *Id.* at 529. In Ohio, poll watchers must take an oath, be a resident of the state and a "qualified elector." OHIO REV. CODE ANN. § 3505.21 (LexisNexis 2010). *See Spencer*, 347 F. Supp. 2d at 529–30.

120. *Id.* at 529.

121. *Id.*

122. *See id.* at 530.

123. *Id.* The court heard evidence that "14% of new voters in a majority white location will face a challenger . . ., but 97% of new voters in a majority African-American voting location will see such a challenger." *Id.*

imposed a severe burden on Ohio voters and was not narrowly tailored to serve a compelling state interest.[124] The court opined that "questionable enforceability of the State's and County's policies regarding good faith challenges and ejection of disruptive challengers from the polls, there exists an enormous risk of chaos, delay, intimidation, and pandemonium inside the polls and in the lines out the door."[125]

In the 2010 midterm elections, reports of voter intimidation, voter suppression and increased challenges were numerous.[126] The outcome of these types of measures are a decrease in voter participation. The threat of voter challenges and voter caging can deter eligible voters from going to the polls.[127] Additionally, the heightened scrutiny associated with voter challenges also has a deterrent and intimidating effect.[128]

When election officials accept voter challenges that exhibit a pattern of race, ethnic, or national origin discrimination, the practice could be found volatile of the Equal Protection Clause and other statutes.[129] Much like discriminatory peremptory challenges in jury selection,[130] the current voter challenge system has created a sort of peremptory voter challenges. In this regard, voter challenges are similar to peremptory challenges used in criminal and civil trials and should be considered peremptory voter challenges because of their use to eliminate voters based solely on race and the ability to utilize the challenges without proffering a basis for the strike.

---

124. *Spencer*, 347 F. Supp. 2d at 536.

125. *Id.* at 535. The Court has repeatedly recognized that "as a practical matter, there must be a substantial regulation of elections if they are to be fair and honest and if some sort of order, rather than chaos, is to accompany the democratic process." *Id.* at 534 (quoting Timmons v. Twin Cities Area New Party, 520 U.S. 351, 358 (1997)).

126. *See, e.g.*, Krissah Thompson, *Some Complaints Surface Amid Stepped-Up Efforts to Monitor Voting Fraud*, WASH. POST, Nov. 3, 2010 ("In Minnesota, where conservative groups had run radio ads and offered $500 rewards to those willing to turn in anyone prosecuted for voter fraud, a few poll watchers aggressively challenged voters until they were confronted by election protection volunteers, lawyers monitoring the election reported."); *see also* Ian Urbina, *Reports of Intimidation and Electronic Problems Surface at Polls Across the U.S.*, N.Y. TIMES, Nov. 2, 2010, ("'One of the most worrisome things we're seeing is an uptick in voter intimidation and misinformation compared to prior elections,' said Wendy Weiser, director of the voting rights and elections project at the Brennan Center for Justice at New York University.")

127. *Voter Registration and List Maintenance: Hearing Before the Subcomm. on Elections of the H. Administration Committee*, 110th Cong. (2007) (statement of Joseph D. Rich, Director, Fair Housing Project) ("Targeted at traditionally disenfranchised voters, this practice relies on voter 'challenge' laws to blindly question the ability of eligible voters to cast a ballot. While dressed in the garb of protecting against the 'voter fraud,' caging is really a cynical way to undermine the most fundamental right of all Americans – the right to participate freely in our democracy – for partisan gain. It is especially pernicious because it has almost invariably been used to suppress the vote of minority voters.").

128. *See* Democratic Nat'l Comm. v. Republican Nat'l Comm., 671 F. Supp. 2d 575, 590 (D.N.J. 2009) (discussing partisan intimidation tactics including those associated with voter challenges).

129. *See generally id.*

130. Batson v. Kentucky, 476 U.S. 79, 82 (1986).

In *Batson v. Kentucky*,[131] the Supreme Court found that a prosecutor's use of peremptory challenges to exclude racial minorities from serving on a jury violated the Equal Protection Clause.[132] In a subsequent case, the Court focused its attention on whether the defendant, when exercising peremptory challenges, should be considered a state actor under the Fourteenth Amendment.[133] The court held that the defendant should be considered a state actor because the defendant used the peremptory challenge procedures with "the overt, significant assistance of state officials."[134] The dissenters in the case argued that the use of peremptory challenges equaled private action because the decision to use the peremptory challenge is left to the defendant's discretion.[135]

In many ways, voter challenges are similar to peremptory challenges due to the lack of an evidentiary standard before denying a fundamental right.[136] Once the government allows private actors to make an assertion regarding whether a citizen lives at a certain address or has a "belief" that the person is otherwise ineligible, it has outsourced its authority and granted the private actors governmental power.

III. PRIVATE PARTISAN ACTIONS AS STATE ACTION

"The timing of the challenges is so transparent that it defies common sense to believe the purpose is anything but political chicanery."[137]

In elections, private partisans participate to the extent that the law allows, i.e., laws governing electioneering, private access to the polls, and what voters can wear to the polls.[138] States should become wary of allowing voter challengers. In the exercise of race based or racially targeted voter challenges, where partisanship is used as a proxy for race, courts have implied liability for both public and private entities.[139] Private partisan racially targeted voter challenges could leave public authorities liable for their involvement in denying citizens the right to vote.

---

131. 476 U.S. 79 (1986).

132. *Id.* at 83–84.

133. *See* Edmonson v. Leesville Concrete Co., 500 U.S. 614, 620 (1991).

134. *Id.* at 622 (quoting Tulsa Prof'l Collection Servs. v. Pope, 485 U.S. 478, 485 (1988)).

135. *Id.* at 631–32 (O'Connor, J., dissenting). Justice O'Connor, Chief Justice Burger and Justice Scalia dissented arguing that the challenge allows private parties to exclude potential jurors and is left wholly within the discretion of the litigant, and considered it "an enclave of private action in a government-managed proceeding." *Id.* at 632–34.

136. *See supra* notes 134–35.

137. Mont. Democratic Party v. Eaton, 581 F. Supp. 2d 1077, 1081 (D. Mont. 2008).

138. Kimberly J. Tucker, *"You Can't Wear That to Vote": The Constitutionality of State Laws Prohibiting the Wearing of Political Message Buttons at Polling Places*, 32 T. MARSHALL L. REV. 61, 61–62 (2006) (arguing that the state statute that bans private partisans access to political buttons at polling places is unconstitutional because it violates free speech).

139. *See Mont. Democratic Party*, 581 F. Supp. 2d at 1078. The Montana Democratic Party sought a temporary restraining order to stop the Montana Republican Party from challenging 6,000 registered voters who were predominately young and registered Democrats. *Id.*

The Supreme Court has applied the notion that a private actor performing a governmental function can be deemed a state action in the area of election law—most notably in areas of race discrimination. In the first of the white primary cases,[140] the Supreme Court found that a Texas statute that excluded blacks from participating in the state primary that was run by the state Democratic Party constituted a state action in violation of the Equal Protection Clause.[141] In the landmark case *Smith v. Allwright*,[142] the Supreme Court stated, "[t]he [political] party takes the character as a state agency from the duties imposed upon it by state statutes; the duties do not become matters of private law because they are performed by a political party."[143]

A. Nexus between the State and the Private Action

The state's ceding of authority to private parties in the area of voter eligibility places governmental power in the private partisan and strips the challenged voter of her rights and the expectation of participating in the electoral practice free from discrimination. In determining whether state action has occurred, the Supreme Court looks at two distinct issues: whether "the deprivation . . . [was] caused by the exercise of some right or privilege created by the State or by a rule of conduct imposed by the state or by a person for whom the State is responsible," and whether "the party charged with the deprivation . . . [is] a person who may fairly be said to be a state actor."[144]

Under this test, the function of voter challenges can be deemed state action: It is the state statutes that provide for voter challenges, and, the action itself is derived from an inherently public-private relationship. In *Edmonson v. Leesville Concrete Co.*,[145] the Court found that "[a]lthough private use of state-sanctioned private remedies or procedures does not rise, by itself, to the level of state action, our cases have found state action when private parties make extensive use of state procedures with 'the overt, significant assistance of state officials.'"[146] In *Tiryak*, a § 1983 case involving voter challenges, the court found that "[n]o activity is more indelibly a public function than the holding of a political elec-

140. *See* Terry v. Adams, 345 U.S. 461, 467–70 (1953) (holding that excluding blacks from "pre-primary elections" constituted unconstitutional state action under the Civil War Amendments); Smith v. Allwright, 321 U.S. 649, 660 (1944) (the Court found that "delegation to a party of the power to fix the qualifications of primary elections is delegation of a state function that may make the party's action the action of the state"); Nixon v. Herndon, 273 U.S. 536, 540–41 (1927) (striking down the Texas statute that denied blacks an opportunity to participate in a state primary).

141. *Nixon*, 273 U.S. at 540–41.

142. 321 U.S. 649 (1944).

143. *Id.* at 663.

144. Lugar v. Edmondson Oil Co., 457 U.S. 922, 937 (1982).

145. 500 U.S. 614 (1991).

146. *Id.* at 622 (citation omitted).

tion . . . cases make it clear that the conduct of the elections themselves is an [e]xclusively public function."[147]

The crucial question is whether enough governmental involvement exists to convert private discrimination into state action. The Supreme Court has held that an activity constitutes state action when the State exercised "coercive power"[148] or when a private actor operates as a "willful participant in joint activity with the State or its agents."[149] Once the government allows private actors to make an assertion regarding whether a citizen lives in a certain address or "believes" that the person is otherwise ineligible, it has outsourced its authority and granted the private actors governmental power. It is the government's responsibility to determine whether a voter is eligible to vote, not a private party's.

Notwithstanding the use of peremptory electoral challenges, the state action doctrine should serve as a deterrent to those who wish to use public power in a discriminatory manner.[150] The Supreme Court has found that even peripheral involvement could equate to an Equal Protection violation.[151] The state's bestowal of authority, here, allows persons other than the state to determine the eligibility of voters. In the landmark case *Smith v. Allwright*, the Supreme Court found that "the [political] party takes its character as a state agency from the duties imposed upon it by state statutes; the duties do not become matters of private law because they are performed by a political party."[152] More contemporaneously, in *Brentwood Academy v. Tennessee Secondary School Athletic Ass'n*,[153] the Court found state action where the private actor received its power from the state.[154] Here, it is the state that allows private actors to serve as voter challengers and challenge voters—thus, granting power to the private actors.

147. Tiryak v. Jordan, 472 F. Supp 822, 824 (E.D. Pa. 1979) (internal quotation marks omitted).

148. Blum v. Yaretsky, 457 U.S. 991, 1004 (1982).

149. United States v. Price, 383 U.S. 787, 794 (1966)).

150. The Supreme Court has held that "the Equal Protection Clause of the Fourteenth Amendment restrains the States from fixing voter qualifications which invidiously discriminate." Harper v. Va. State Bd. of Elections, 383 U.S. 663, 666 (1966) (outlawing the poll tax).

151. United States v. Guest, 383 U.S. 745, 755–56 (1966) ("This is not to say, however, that the involvement of the State need be either exclusive or direct. In a variety of situations the Court has found state action of a nature sufficient to create rights under the Equal Protection Clause even though the participation of the State was peripheral, or its action was only one of several cooperative forces leading to the constitutional violation.").

152. Smith v. Allwright, 321 U.S. 649, 663 (1944).

153. 531 U.S. 288 (2001).

154. *Id.* at 290–91, 296 ("We have treated a nominally private entity as a state actor when it is controlled by an 'agency of the State,' when it has been delegated a public function by the State, when it is 'entwined with governmental policies,' or when government is 'entwined in [its] management or control.'") (alteration in original) (citations omitted).

In cases where the court has not found state action, it acknowledged the difficulty in making the designation.[155] The courts have neglected to find a state action violation primarily in the area of inner affairs of political parties.[156] The Supreme Court and lower courts have found that the political party's decision on the seating and selection of delegates does not constitute state action.[157] Here the discrepancy does not lie between the political party and its internal rules governing delegate selection or primary dates, but the party's manipulation of state laws to discriminate against voters on the basis of race. The use of the state statute to provide access to the voters and to allow the parties to challenge voters based on a list that was developed on a discriminatory premise, and the State's approval of the individual challenges such that voters eligibility is questioned and in some cases disallowed, clearly places the political party in the position of acting on behalf of the State and its actions can be attributed to the State.

B. Private Partisans Performing Governmental Functions

The Supreme Court has repeatedly determined that the act of regulating voting is a governmental responsibility.[158] The Federal Constitution grants the authority to administer elections to the federal and state government, which convenes the responsibility to register, monitor, and count ballots pursuant to the rules, regulations and statutes those governmental entities prescribe.[159] Clearly, administering the vote and access to the voting process is an exclusive governmental function.

Determining whether a particular voter is eligible, therefore, is likewise a governmental duty. The use of voter challenges outsources

---

155. *See, e.g.*, Evans v. Newton, 382 U.S. 296, 299 (1966) ("What is 'private' action and what is 'state' action is not always easy to determine. Conduct that is formally 'private' may become so entwined with governmental policies or so impregnated with a governmental character as to become subject to the constitutional limitations placed upon state action.") (citation omitted); *see also* Gilmore v. City of Montgomery, 417 U.S. 556, 571–74 (1974) (failing to find state action because the record was not conclusive on whether the nonexclusive use of recreational facilities by all-white private schools, private school-affiliated groups, and all-white non-school organizations was enough state action to warrant judicial intervention on constitutional grounds).

156. *See* Max v. Republican Comm. of Lancaster Cnty., 587 F.3d 198, 202 (3d Cir. 2009) (holding that a political party's involvement in a primary election did not constitute state action but instead constituted internal party affairs which are not subject to § 1983 enforcement).

157. Cousins v. Wigoda, 419 U.S. 477, 491 (1975) (holding that the First Amendment protects a national political party's delegate selection rules); Bachur v. Democratic Nat'l Party, 836 F.2d 837, 843 (4th Cir. 1987) (holding that a challenge to the gender allocation rule was not permitted for delegate selection); Ripon Soc'y., Inc. v. Nat'l Republican Party, 525 F.2d 567, 586 (D.C. Cir. 1975) (holding that the Equal Protection Clause was not violated when a national party did not make presidential nominations based on a defined constituency of one person-one vote).

158. United States v. Manning, 215 F. Supp. 272, 283, 285–86 (W.D. La. 1963) (holding that the "manner of holding elections" has been understood as covering the entire election process, from voter registration to the counting of ballots); *see also* Foster v. Love, 522 U.S. 67, 69 (1997) (finding Congress's Elections Clause authority "well settled . . . to override state regulations" involving federal election administration matters); *Ex parte* Coy, 127 U.S. 731, 755 (1888).

159. The Supreme Court described the state's abdication of responsibility as an "evil." Terry v. Adams, 345 U.S. 461, 477 (1953).

this responsibility to private actors. The voter challenger's use of discriminatory means to develop the list of voters to oppose promotes race discrimination and can make the state responsible for the private party's actions; thus, the private partisans are involved in a governmental function.

C. Scrutinizing Voter Challengers: Addressing the Vote Fraud Claim

While this Article fervently argues that the use of racially targeted voter challenges violates the Equal Protection Clause, it also recognizes that the level of scrutiny applied and the case-by-case analysis that the courts employ will determine whether a violation has occurred.[160] A governmental or private agency would argue that it narrowly tailored its voter challenge statutes and had a compelling state interest in allowing voter challengers in the polling place and that state and federal authorities would require adherence to the law.[161] Moreover, it would assert that the private party's actions did not constitute state action under the Equal Protection Clause. In this instance, even under the most forgiving standard, it is difficult to ascertain how the state could meet its burden.[162]

Additionally, proponents of voter challenges will argue that the challenges serve as a voter integrity measure that helps prevent fraud.[163]

---

160. In *Nat'l Collegiate Athletic Ass'n v. Tarkanian*, 488 U.S. 179, 179–81, 195 (1988), a state university imposed disciplinary sanctions against its basketball coach in compliance with National Collegiate Athletic Association rules and recommendations. The Court found that this action did not turn Association's otherwise private conduct into "state action," and thus Association could not be held liable for violation of coach's civil rights. In *San Francisco Arts & Athletics, Inc. v. U.S. Olympic Comm.*, 483 U.S. 522, 526–27, 547 (1987), the United States Olympic Committee (USOC) and the International Olympic Committee brought suit under the Amateur Sports Act against a California corporation and various individuals to restrain their use of the term "Olympics" to describe an athletic competition they sponsored. The Court said that the USOC was not a government actor.

161. *See* Burson v. Freeman, 504 U.S. 191, 199 (1992). "The Court also has recognized that a State 'indisputably has a compelling interest in preserving the integrity of its election process.'" *Id.* (quoting Eu v. San Francisco Cnty. Democratic Cent. Comm., 489 U.S. 214, 231 (1989)). "The Court has thus 'upheld generally applicable and evenhanded restrictions that protect the integrity and reliability of the electoral process itself.'" *Burson*, 504 U.S. at 199 (quoting Anderson v. Celebrezze, 460 U.S. 780, 788 n.9 (1983)). "In other words, it has recognized that a State has a compelling interest in ensuring that an individual's right to vote is not undermined by fraud in the election process." *Id.* The Court further held that "[t]o survive strict scrutiny, however, a State must do more than assert a compelling state interest-it must demonstrate that its law is necessary to serve the asserted interest." *Id.*

162. *See* Burdick v. Takushi, 504 U.S. 428, 434 (1992) (holding that during a challenge to a state election law a court "must weigh 'the character and magnitude of the asserted injury to the rights protected by the First and Fourteenth Amendments that the plaintiff seeks to vindicate' against 'the precise interests put forward by the State as justifications for the burden imposed by its rule,' taking into consideration 'the extent to which those interests make it necessary to burden the plaintiff's rights.'").

163. While voter fraud once was a distinct problem in this country, it has certainly become less of a threat to democracy. *See* KEYSSAR *supra* note 99, at 123. As early as 1934, noted election administration expert Joseph Harris found that voter fraud was disappearing and "[h]onest elections have become the established rule in most sections of the country." *See* JOSEPH P. HARRIS, ELECTION ADMINISTRATION IN THE UNITED STATES 317 (1934). In one instance, the Court held that voter fraud in Indiana has never been a problem. *See* Crawford v. Marion Cnty. Election Bd., 553 U.S. 181, 194 (2008).

To these arguments, courts have already spoken. For example, in *Democratic Nat'l Committee v. Republican Nat'l Committee*,[164] the court found that the effect of the alleged incidents of fraud "pale[d] in comparison to the damage that would likely result from allowing the types of ballot security initiatives that are currently prohibited by the Consent Decree."[165]

The strongest response to such arguments is that although permitting statutory voter challenges helps prevent voter fraud, the elimination of this ill is not the responsibility of private actors or partisan organizations. Instead, it is the government's sole responsibility to prescribe the manner of elections and to protect the process. Moreover, in the recent Supreme Court case *Crawford v. Marion Cnty. Election Board*,[166] the Court noted that the State of Indiana had not encountered any documented reports of voter fraud.[167] In fact, once implemented, the voter ID law that had the stated compelling reason of preventing fraud—much like the voter challenges—was more effective at eliminating eligible voters. Likewise, as other litigation has brought to light, so-called ballot security measures tend to interfere with the balloting process and do not promote voter participation and confidence.[168]

The precision that is used to eliminate, frustrate and intimidate eligible voters and, in most cases minority voters, should serve as a credible reason to consider curtailing, if not eliminating, the position of voter challengers. Or, at a minimum, to alter their authority and presence in the polling site. Their presence clutters the polling site, and distracts poll workers and election officials from the process of allowing citizens to vote by requiring that they comply with verification procedures that are based on a flawed discriminatory premise. Additionally, the means that private partisans have utilized to arguably prevent fraud is tainted with racial targeting and discrimination.

---

164. 671 F. Supp.2d 575 (D.N.J. 2009).

165. *Id.* at 610 ("Even if the Court were to assume that all 300 of the alleged incidents of fraud involved in-person misconduct at the polls, the effects of such fraud pales in comparison to the damage that would likely result from allowing the types of ballot security initiatives that are currently prohibited . . . . [Another] matter involved a voter challenge list that included 35,000 predominantly-minority individuals. If only one tenth of those individuals were deterred from voting by harassment at the polls, the effect would have been the disenfranchisement of 3,500 individuals-a number far greater than the 300 alleged incidents of voter fraud which the [Republican National Committee] points to in support of its claim.").

166. 553 U.S. 181 (2008).

167. *Id.* at 194.

168. *Democratic Nat'l Comm.*, 671 F. Supp. 2d at 610 ("The effects of ballot security initiatives . . . pose a far greater threat to the integrity of modern elections than in-person voter fraud. In fact, even a cursory investigation of the prevalence of voter intimidation demonstrates that ballot security initiatives have the potential to unfairly skew election results by disenfranchising qualified voters in far greater numbers of than [sic] the instances of in-person fraud that may occur during any given race.").

While the Supreme Court has found the State's interest in preventing voter fraud to be a constitutionally acceptable goal,[169] the levels of voter fraud are miniscule when compared to the imposition placed on voters to succumb to private inquisition.[170] When you compare the level of voter fraud, vote caging, and baseless voter challenges the voter is the victim.[171] The government's interest is to provide a fair and free process for citizens to access the ballot. It has become clear, however, that the goal of some private entities is to stop as many votes as possible under the guise of voter fraud prevention, ballot security or voter integrity.

IV. STATE ACTION SOLUTIONS

"As against the unfettered right, however, lies the [c]ommon sense, as well as constitutional law . . . that government must play an active role in structuring elections; as a practical matter, there must be a substantial regulation of elections if they are to be fair and honest and if some sort of order, rather than chaos, is to accompany the democratic processes."[172]

In *Crawford*, Justice Souter recognized that voting legislation has "two competing interests," the fundamental right to vote and the need for governmental structure in elections.[173] In a democracy, political participation serves as a fundamental component of its legitimacy.[174] Partisan voter challenges do not further these laudable principles. This allowance, however, is often used to "game" the system. Ballot security initiatives often use party as a proxy for race.[175] The presumption against the voter

---

169. *See, e.g.*, *Crawford*, 553 U.S. at 196 (stating that there is "no question" that prevention of voter fraud constitutes a legitimate and important state interest).

170. *See* David Schultz, *Less Than Fundamental: The Myth of Voter Fraud and the Coming of the Second Great Disenfranchisement*, 34 WM. MITCHELL L. REV. 483, 486 (2008) ("[V]oter fraud is used as a pretext for a broader agenda to disenfranchise Americans and rig elections.").

171. In a 1986 election, the RNC used vote caging to compile a list of voters, mostly black, that it attempted to have removed from the voter rolls. At the time, Kris Wolfe, the Republican National Committee Midwest political director, wrote Lanny Griffith, the committee's Southern political director, "I know this is really important to you. I would guess this program would eliminate at least 60-80,000 folks from the rolls . . . If it's a close race . . . which I'm assuming it is . . . this could keep the black vote down considerably." *See* Martin Tolchin, *The Political Campaign: Committees Prepare Negative Attacks in House Races*, N.Y. TIMES, Oct. 26, 1988, at B19 (discussing similar tactics in a 1984 election); TERESA JAMES, CAGING DEMOCRACY: A 50 YEAR HISTORY OF PARTISAN CHALLENGES TO MINORITY VOTERS 12 (2007), *available at* http://www.projectvote.org/images/publications/Voter%20Caging/Caging_Democracy_Report.pdf (discussing the vote caging activities of the 1986 Louisiana election). Following this caging scandal, both parties agreed to amend the original 1982 consent decree to require that the RNC would submit to the court any future ballot security plan for approval.

172. *Crawford*, 553 U.S. at 210 (Souter, J., dissenting) (alterations in original) (internal quotation marks omitted).

173. *Id.* at 210.

174. *See* Reynolds v. Sims, 377 U.S. 533, 555 (1964) ("The right to vote freely for the candidate of one's choice is of the essence of a democratic society, and any restrictions on that right strike at the heart of representative government.").

175. In his expert testimony in *Democratic National Committee*, Professor Chandler Davidson opined:

creates a barrier to participation.[176] Accordingly, voter challenges should be abolished, or at a minimum, altered to a more uniform and silent mechanism in the polling place.

In many instances, voter challenges are used for partisan gain, not election integrity or ballot security, and derive from race based caging schemes.[177] In all these situations, the law should provide some protections to ensure that only those ineligible are burdened with the responsibility of proving or disproving private allegations. In current practices, the presumption of eligibility has been removed from the voter and placed squarely in the hands of the private actor.[178] The following measures would provide a presumption of legitimacy to the voter and require the state to determine eligibility, not the private actors. Moreover, private actors would be more consistent with federal observers in their duties and Election Day characteristics, i.e., they would be allowed to enter and remain in the polls but not interject in the voting process.[179]

A. Eliminating Voter Challenges

The voter challenge process in its present form allows private parties the opportunity to disrupt the electoral process for partisan gain.[180] Because states, in large part, determine who has access to the polls, it would take a major effort to legislate the voter challenger's presence out of the voting precinct. However, the disruption that is caused and the discriminatory basis for the challenges—whether race, national origin or

---

> [T]he reason such programs are usually carried out by Republicans rather than Democrats may simply be a matter of statistics: minority voters—who are far more likely to be added to a challenge list because of irregularities in registration attributable to language barriers, lack of photo identification cards ("ID"), or changes of address—have historically tended to vote for Democratic candidates.

Democratic Nat'l Comm. v. Republican Nat'l Comm., 671 F. Supp. 2d 575, 590 (D.N.J. 2009).

176. *See* Richard L. Hasen, *The Democracy Canon*, 62 STAN. L. REV. 69, 71–79, 123 (2009) (arguing for expanding the use of the "Democracy Canon," which generally construes election statutes in the voter's favor).

177. *See* JAMES, *supra* note 171, at 5–6.

178. *See* Benson, *supra* note 62, at 271 (arguing for uniform election administration measures on basis that "the presumption that the voter is sovereign on Election Day, and that every policy and practice implemented on that day must consider the voter and his or her experience as paramount").

179. *See infra* Part IV.B–C.

180. In *Spencer v. Pugh*, 543 U.S. 1301, 1301–03 (2004), the State of Ohio allowed challengers to be stationed in polls. The Hamilton County Republican Party requested to have additional challengers beyond the typical filed list of precinct challengers present at polling places, assigned to predominantly African American precincts. *Id.* at 1301. The district court overruled the State's decision to allow the use of challengers on Election Day because it could create an extraordinary risk of intimidation and delay. *Id.* at 1301–02. However, the Sixth Circuit reversed the decision and the Supreme Court affirmed. *See id.* at 1302–03. Justice Stevens reviewed the case without referring the matter to the full Supreme Court. *Id.* at 1303. He noted in his brief opinion that although "the threat of voter intimidation is not new to our electoral system," there was simply no evidence that voter-targeted fraud would occur if challengers were permitted in the polling place. *Id.* at 1302. Emphasizing his "faith" that voter-targeted deception and intimidation would not occur and that "the elected officials and numerous election volunteers on the ground will carry out their responsibilities in a way that will enable qualified voters to cast their ballots," Justice Stevens upheld the Sixth Circuit's decision to allow challengers in the polling place. *Id.* at 1302–03.

language ability—as a proxy for party must be recognized and dealt with in either litigation, via the Fourteenth Amendment, or a legislative manner, via state statute. Some scholars would argue that the presence of voter challengers assist the electoral process as a deterrence factor. Such an argument, however, has little basis in fact,[181] and pales in comparison to the mammoth accounts of disruption and chaos.[182]

As discussed, partisans use voter challenges like illegal peremptory challenges. Armed with nothing more than an address, partisans make presumptions based on race, primarily that the voter will not vote for them thus they issue a challenge to frustrate the voter and the democratic process. Challenges make lines longer and cause substantial confusion. They can also intimidate voters and disrupt polling place procedures.[183] The primary detraction for voter challenges is that the practice casts dispersions on a wide swath of voters. The government could not disqualify a voter because of returned mail, yet, private partisans are given the ability to infringe upon a fundamental right.

While it is tempting to suggest that states eliminate the position of voter challengers in the polls to avoid the instances of partisan manipulation at the expense of the voter's fundamental right, a more balanced approach is warranted. The state rationale for massive voter challenges is to deter voter fraud. However, the basis for this assertion rests upon race discrimination.[184] The partisan is not seeking to eliminate voter fraud, but to eliminate voters whom it believes are predisposed to vote against its candidate and/or its ideals.

Granting the challenger the authority to question a voter's eligibility and requiring the voter to rebut the assertions of nongovernmental entities shifts the power from the voter to the partisan. This shifting is burdensome to the voter and requires the state to weigh the evidence and make a decision against the voter. The eligibility is questioned based on the challenge. If the voter cannot address the nature of the challenge, the State will, in some instances, not allow the citizen to vote or require them

181. *See supra* Part III.B–C.

182. *See* Summit Cnty. Democratic Cent. & Exec. Comm. v. Blackwell, 388 F.3d 547, 554 (6th Cir. 2004) ("The burden on the right to vote is evident. In this case, we anticipate the arrival of hundreds of Republican lawyers to challenge voter registrations at the polls. Behind them will be hundreds of Democrat lawyers to challenge these Challengers' challenges. This is a recipe for confusion and chaos.").

183. *See* Schultz, *supra* note 170, at 485 ("A second great disenfranchisement is afoot across the United States as, yet again, voter fraud is raised as a way to intimidate immigrants, people of color, the poor, and the powerless, and prevent them from voting."); *see also* Sherry A. Swirsky, *Minority Voter Intimidation: The Problem That Won't Go Away*, 11 TEMP. POL. & CIV. RTS. L. REV. 359, 361–65 (2002) (arguing that voter intimidation is a common campaign tactic and providing examples of "ballot security measures" and other measures used to deceive and intimidate).

184. *See supra* Part II.

to vote a provisional ballot.[185] The shifting of power and the presumption that the voter is in fact ineligible unduly burdens the voter.

As held in *Tiryak*, elections are public functions.[186] Allowing private parties to begin a process that could lead to racial targeting and racial discrimination manipulates a state's authority to stage elections. When this occurs, the private partisans "are abusing the process the State . . . has provided to ensure the accuracy of voter rolls (indeed, they are using the process designed to protect the integrity of the political process to undermine it)." [187]

If the poll worker allows the voter challenge to question a voter's eligibility in a racially targeted manner, that worker could subject himself to liability for allowing racially based discrimination.[188] In *Tiryak,* the court recognized the duality of poll watchers and the potential for state liability. The court opined that:

> The poll-watcher performs a dual function on Election Day. On the one hand, because he is designated and paid by a political party, his job is to guard the interests of that party's candidates. On the other hand, because exercise of his authority promotes an honest election, the poll-watcher's function is to guard the integrity of the vote. Protecting the purity of the electoral process is a state responsibility and the poll-watcher's statutory role in providing that protection involves him in a public activity, regardless of his private political motive.[189]

---

185. Additionally, an unfortunate consequence of the voter challenges has been the misuse of provisional ballots, which are mandated in the Help America Vote Act to ensure that eligible citizens are not turned away from the polls. *See* 42 U.S.C. § 15482 (2006). The administration of provisional ballots has been called into question for the myriad of ways that election administrators determine whether to issue and count the ballot. In 2004, nearly 1.9 million provisional ballots were cast and 1.2 million were counted, which left more than half a million people disenfranchised. U.S. ELECTION ASSISTANCE COMM'N, FINAL REPORT OF THE 2004 ELECTION DAY SURVEY, 6–5 (2005). Moreover, implementing the provisional ballot requirement left poll workers confused and many ballots uncounted for, creating even more disparities. *See* PEOPLE FOR THE AM. WAY ET AL., SHATTERING THE MYTH: AN INITIAL SNAPSHOT OF VOTER DISENFRANCHISEMENT IN THE 2004 ELECTIONS, 8 (2004), *available at* http://www.866ourvote.org/tools/publications_testimony/files/0002.pdf ("There was widespread confusion over the proper use of provisional ballots, and widely differing regulations from state to state—even from one polling place to the next—as to the use and ultimate recording of these ballots."); *see also* R. Bradley Griffin, Note, *Gambling with Democracy: The Help America Vote Act and the Failure of the States to Administer Federal Elections*, 82 WASH. U. L. REV. 509, 525–28 (2004) (arguing that HAVA provides states too much control over federal election procedures); Tokaji, *supra* note 9, at 129–33 (assessing the 2004 election and the failure of election reform to remedy election administration problems).

186. Tiryak v. Jordan, 472 F. Supp. 822, 824 (E.D. Pa. 1979) ("Protecting the purity of the electoral process is a state responsibility and the poll-watcher's statutory role in providing that protection involves him in a public activity, regardless of his private political motive.").

187. Mont. Democratic Party v. Eaton, 581 F. Supp. 2d 1077, 1082 (D. Mont. 2008) (discussing the Montana Democratic Party's pursuit of a temporary restraining order to stop the Montana Republican Party from challenging 6,000 registered voters who were predominately young and registered Democrats).

188. *See supra* Part II.

189. *Tiryak*, 472 F. Supp at 824.

In many instances, it is not the voter's ineligibility that is determined but his inability to produce supporting documents at the polling place that are not required. For example, a poll worker, in response to a voter challenge, could require a challenged voter to present valid photo identification. If the voter is unable to present the requested identification, his inability to do so does not prove that a citizen was attempting to commit fraud, but that he came to the polling place unprepared and for this, his right to vote in this election has been negated. In many instances, challenged voters receive a provisional ballot. The provisional ballots, however, are not counted unless the voter returns to a central registrar's office within a prescribed period of time to provide additional documentation to verify his identity. Without this additional step, the provisional ballot is not counted. In this situation, the voter challenge merely caused a frustrating delay at the polls and possibly the disenfranchisement of the voter if she does not ultimately provide the necessary documentation. Accordingly, the private partisan has effectively denied, with the state's endorsement and assistance, an eligible voter the opportunity to vote.

B. Need for Uniform Guidelines

The federal government does not administer elections. It can, as discussed *infra*, regulate elections through the passage of legislation. In fact, Congress has passed legislation to curb discriminatory practices in elections. Under the Voting Rights Act of 1965,[190] the federal government is tasked to ensure that the method of electing federal offices is not tainted with race, ethnic, national origin or language minority discrimination.[191] Other statutes, such as the National Voter Registration Act (NVRA)[192] and the Help America Vote Act (HAVA)[193] address the actual administration of elections, *inter alia*, providing funds for new voting machines, requiring the development of voter databases, and increas-

190. *See* Fannie Lou Hamer, Rosa Parks, & Coretta Scott King Voting Rights Act Reauthorization and Amendments Act of 2006, Pub. L. No. 109-246, 120 Stat. 577 (codified as amended at 42 U.S.C. § 1973). The Voting Rights Act's most prominent temporary provisions include Sections five and 203, which govern which jurisdictions must report all voting changes to the Attorney General and designates those jurisdictions required to provide election materials in certain minority languages. 42 U.S.C. § 1973c (2006) (Section 5); *id.* § 1973aa-1a (Section 203).

191. *See* discussion *supra* Parts I.A, II, III.A, IV.A (discussing the Fourteenth Amendment and Equal Protection Clause).

192. In enacting NVRA, Congress sought to increase voter registration and participation. 42 U.S.C. §§ 1973gg(b)(1)–(2) (2006). The NVRA requires states to register voters for federal elections through mail registration and when citizens apply for a driver's license or seek services from certain state agencies that receive federal funds, such as public assistance offices providing welfare and Medicaid, veteran's affairs offices, and libraries. *Id.* §§ 1973gg-2, 1973gg-5. *See supra* notes 73–74 and accompanying text.

193. In 2002, Congress passed the Help America Vote Act to provide funding to replace outdated voting machines and created the Election Assistance Commission to serve as a "clearinghouse" for election administration matters. Help America Vote Act of 2002 (HAVA), Pub. L. No. 107-252, 116 Stat. 1666 (codified as amended at 42 U.S.C. §§ 15301–15545). For a detailed description of HAVA's legislative history, see Leonard Shambon & Keith Abouchar, *Trapped by Precincts? The Help America Vote Act's Provisional Ballots and the Problem of Precincts*, 10 N.Y.U. J. LEGIS. & PUB. POL'Y 133, 160–65 (2006).

ing voter registration opportunities. Although both the NVRA and the HAVA are federal creations, the legislation was left almost exclusively to the states to implement.[194] In order to thwart the racially discriminatory practice of racially targeted voter challenges, federal agencies, such as the Election Assistance Commission should issue a Best Practices guide for states and encourage the use of uniform guidelines.[195] Congress could pass these guidelines, however, the lack of funding and the political response may seem too costly to implement. States should provide clear guidelines on the grounds for issuing challenges and limit those grounds to those that a challenger would have personal knowledge of—such as voting more than once, similar to the Vermont challenge statute.[196] States should not allow challenges based on residence, age or other voter qualification. It is the state that is in the best position to determine voter eligibility on these grounds. While state legislators provide the opportunity for partisans to challenge voters, it is the voter's fundamental right to participate in the electoral process. State laws should not allow challengers to jeopardize that right based on vote caging or intimidation schemes built to favor partisan outcomes instead of full participation in the electoral process or on other administrative grounds, such as a valid registration.

Likewise, because local election workers are primarily responsible for implementing federal and state election laws, it is crucial that poll workers receive proper and comprehensive training.[197] The federal government has made an effort to provide states with best practices recommendations for poll worker recruitment and training. Many nonprofit or-

194. Some states, like Tennessee, performed well, albeit only after federal intervention. *See* United States v. Tennessee, No. 3-02-0938 (M.D. Tenn. Sept. 27, 2002), http://www.justice.gov/crt/voting/nvra/tn_cd.pdf (consent decree); *see also* Jason Marisam, *Voter Turnout: From Cost to Cooperation*, 21 ST. THOMAS L. REV. 190, 202–04 (2009) (criticizing the NVRA for not dramatically improving voter turnout).

195. The EAC issued a Best Practices Guide for Poll Worker Training, but has not issued guidelines on poll watchers or voter challenges. EAC guidelines, however, are not met with the honor and reverence regularly afforded a federal government agency because of its lack of enforcement power. *See* Heather K. Gerken, *Shortcuts to Reform*, 93 MINN. L. REV. 1582, 1608 (2009) (criticizing the EAC's attempts to suggest Best Practices); Hasen, *supra* note 3, at 4 (noting that the EAC "has so far proven ineffective and now appears in danger of becoming a new site for partisan stalemate over election reform."); Leonard M. Shambon, *Implementing the Help America Vote Act*, 3 ELECTION L.J. 424, 428 (2004) ("The EAC was designed to have as little regulatory power as possible. . . . [A]nd for the most part it cannot 'issue any rule, promulgate any regulation, or take any other action' imposing a requirement on any state or unit of local government . . . .") (quoting HAVA, Pub. L. No. 107-252, § 209, 116 Stat. 1666, 1678 (2002)).

196. Karlan, *supra* note 9, at 19–24, 25–29 (arguing that the Bush administration treated vote fraud as a much larger problem than political exclusion). Karlan argues that "we need legislation that recognizes an official obligation to make sure all citizens who are eligible to vote are placed on the voting rolls and that elections run smoothly and accurately." *Id.* at 29. Also, Karlan believes "it is critical [in the area of election law] to make sure the rules are clear and clearly established before the election begins." *Id.* The Bush administration also politicized the Civil Rights Division (and the voting rights section in particular), and Karlan believes this department needs to be remade. *Id.* at 28–29.

197. Many states call those persons who administer elections on Election Day at the polling sites across the country election judges or poll workers.

ganizations also seek to assist county and state governments with poll worker training. Proper poll worker training can serve as the difference between a smooth election and a troubled one. It is difficult to overstate the importance of proper poll worker training and knowledge regarding how the machines operate, what items are needed to operate the machines, what question are appropriately asked, who can assist the voter, and how to avoid racial or ethnic discrimination during the exercise of the political process.[198]

C. From Poll Watchers to Silent Observers

On Election Day, certain jurisdictions are required either under the Voting Rights Act or a court order to allow federal observers inside the polling place to observe and document the election process.[199] The Voting Rights Act of 1965 instituted the advent of the federal observer, whose primary responsibility is to monitor the electoral process.[200] Federal observers are prohibited from interfering in the voting process or

198. See United States v. City of Hamtramck, No. 0073541 (E.D. Mich. Jan. 29, 2004) , http://www.justice.gov/crt/voting/sec_2/hamtramck_cd04.pdf (second amended consent order and decree requiring that the City appoint bilingual translators to protect against voter discrimination targeting Arab-Americans); United States v. Berks Cnty., 277 F. Supp. 2d 570, 575, 583 (E.D. Pa. 2003) (preventing race-based voter discrimination by ordering that, *inter alia*, all election-related materials to be printed in English and Spanish).

199. Various provisions in the Voting Rights Act authorize the Attorney General or a court to order or appoint federal observers. 42 U.S.C. §1973a (2006). For example, subsection (a) provides:

> Whenever the Attorney General or an aggrieved person institutes a proceeding under any statute to enforce the voting guarantees of the fourteenth or fifteenth amendment in any State or political subdivision the court shall authorize the appointment of Federal observers by the Director of the Office of Personnel Management in accordance with section 1973d of this title to serve for such period of time for such political subdivisions as the court shall determine is appropriate to enforce the voting guarantees of the fourteenth or fifteenth amendment . . . .

Id. The Attorney General may also assign Federal observers to a jurisdiction certified under 42 U.S.C. §1973f. The Attorney General has certified approximately 148 jurisdictions in Alabama, Arizona, Georgia, Louisiana, Mississippi, New York, North Carolina, South Carolina and Texas. *About Federal Observers and Election Monitoring*, DEPARTMENT OF JUSTICE: CIVIL RIGHTS DIVISION, http://www.justice.gov/crt/voting/examine/activ_exam.php (last updated Sept. 7, 2010). When determining where to send federal observers, the Attorney General assesses the following information:

- Any incidents of discrimination or interference with the right to vote in connection with upcoming or recent elections;
- Any complaints to local or state officials about the incidents and what, if anything, was done in response;
- Names and contact information for victims of discrimination or other violations of federal voting rights law;
- Names and contact information for any persons who have first-hand knowledge of the incidents;
- Names and contact information, if possible, for persons alleged to have engaged in discrimination or other violations of federal voting rights law;
- Locations where incidents have occurred.

See id. Thirteen jurisdictions in nine states have been ordered to allow federal observers to monitor elections. *Id.*

200. *See* §1973f(d)(1)–(2) (charging Federal observers with the duties of attending places where elections are being held or votes are being tabulated to "observe[] whether persons entitled to vote are being permitted to vote," and "whether votes cast votes cast by persons entitled to vote are being properly tabulated").

attempting to correct what may seem to constitute voting irregularities or illegalities. A federal observer documents what she observes from the beginning of the Election Day to the counting of the ballots.[201] Federal observers are merely eye witnesses, and are not allowed to interfere with the voting process. Unlike voter challenges they do not interfere with the voting process. They do not speak to nor direct poll workers or voters. At the end of the Election Day, federal observers draft a report that can be used, if need be, in any litigation if it is found that systematic discrimination in the operation of the polling place occurred. Likewise, voter challengers could serve as eyewitnesses and document what they observe throughout the Election Day experience. Federal observers help thwart potential voter intimidation. Their presence promotes the exercise of voter integrity and the elimination of voter fraud.

Instead of the voter losing her ability to vote or having that right called into question, if the voter challenger has credible documentation of voting irregularities, especially voter fraud, the voter challenger's Election Day report can serve as evidence in the litigation proceeding. Instead of forcing the voter to possibly lose her right to vote, the voter challenger preserves the right to challenge the election in future litigation. In this way, the presumption of eligibility is restored to the voter. More so, the burden to prove voter fraud or other irregularity is squarely placed on the partisan but does not disrupt the voting process nor strip the citizen's right to vote.[202]

This Article is not advocating that state governments treat voter challengers as government federal observers. Voter challenger statutes should include a crucial and vital characteristic that federal observers have, which is that they are neutral and impartial.[203] Just as the history of partisan involvement has caused the scaling back of partisan involvement in the election process, so here, an adjustment is needed to ensure the impartiality of the electoral system. When private partisan concerns are paramount to individual voters the process is skewed towards the partisan and questions the integrity of the system.

---

201. *See* James Thomas Tucker, *The Power of Observation: The Role of Federal Observers Under The Voting Rights Act*, 13 MICH. J. RACE & L. 227, 248 (2007) ("Federal observers are able to monitor [and document] every aspect of an election, from the time the voter enters the polling place to the moment that he or she casts her ballot, and even thereafter when the ballots are tabulated.").

202. See Steve Barber et al., Comment, *The Purging of Empowerment: Voter Purge Laws and the Voting Rights Act*, 23 HARV. C.R.-C.L. L. REV. 483, 483–84 (1988), for a discussion of the burdens on mistakenly purged voters. The authors assert that the voter purge laws place the burden of reregistering on the purged voters, which may "thwart political participation and place a disproportionate burden on minority voters." *Id.* at 483. The history of mechanisms such as voter purge laws have created the disillusionment of minority voters and has resulted in low political participation from such groups. *Id.* at 483–84.

203. Federal observers are Office of Personnel Management employees. They are nonpartisan and merely observe the election process without comment or interruption. Tucker, *supra* note 208, at 230, 241, 254.

CONCLUSION

While our democratic form of government thrives on political participation, the level of authority that government has given to private partisan organizations through the use of voter challenges may well threaten the heart and integrity of the democratic process. If private individuals are allowed to determine who is allowed to vote and in what manner, the damage to the integrity of the electoral process is deeply wounded, particularly when those challenges are based solely on geographical designation, physical hue or language ability. In actuality, these measures determined the eligibility of the voter and in many cases, prevent them from casting a ballot. A strong public-private alliance allows organizations, groups and individuals to express their collective political opinion. It is the ability, however, to silence those voices through the lack of confidence[204] in the system and particularly allowing partisans to determine voter eligibility that harms the system.

The State, which is given the authority to prescribe the requirements for voting, has surrendered that authority to private and often partisan individuals. More poll watchers, more litigation or stringent statutes are not the only answer to this perplexing problem. The solution lies in our willingness to prescribe measures that presume the validity and eligibility of voters and scale back the Jim Crow era disenfranchising methods that are becoming prevalent in this new millennium.

204. *See* Democratic Nat'l Comm. v. Republican Nat'l Comm., 671 F. Supp. 2d 575, 611 (D.N.J. 2009).

MELENDEZ-DIAZ V. MASSACHUSETTS: UPHOLDING THE GOALS AND GUARANTEES OF THE CONFRONTATION CLAUSE

INTRODUCTION

The Sixth Amendment's Confrontation Clause guarantees criminal defendants the right to confront witnesses who bring testimony against them.[1] The Clause is generally interpreted to afford criminal defendants the right to cross-examine adversarial witnesses live in court, under oath, and in the presence of the trier of fact.[2] Historically, accusatory testimony by absent witnesses was admissible in some situations, though with the adoption of the Sixth Amendment, the right to have an accuser present at a criminal trial became virtually absolute.[3] Nevertheless, as advances in technology ushered in new kinds of evidence, courts often disagreed about which of these novel categories were subject to the Clause.[4]

The admissibility of scientific evidence has certainly been no exception to this ongoing debate.[5] One issue stemming from this uncertainty—whether scientific evidence in the form of affidavits containing test results is admissible in court absent live testimony[6]—was answered by the United States Supreme Court in *Melendez-Diaz v. Massachusetts*.[7] Concluding such affidavits constitute testimonial evidence,[8] *Melendez-Diaz* rightly subjected scientific analysts to confrontation under the Sixth Amendment.[9]

Part I of this Comment gives a brief description of the history and case law behind the Supreme Court's Confrontation Clause jurisprudence. Part II summarizes the facts, procedural history, and opinions in *Melendez-Diaz*. Part III commends the *Melendez-Diaz* Court for upholding the purposes, guarantees, and historical intentions behind the Con-

1. U.S. CONST. amend. VI.
2. Maryland v. Craig, 497 U.S. 836, 845–46 (1990).
3. *See generally* Crawford v. Washington, 541 U.S. 36, 43–50 (2004) (discussing the history of the Confrontation Clause).
4. *See* Melendez-Diaz v. Massachusetts, 129 S. Ct. 2527, 2533 (2009) (commenting on conflicting state laws exempting scientific affidavits from confrontation).
5. *See* Lisa Gonzalez, *The Admissibility of Scientific Evidence: The History and Demise of* Frye v. United States, 48 U. MIAMI L. REV. 371, 371–72 (1993) (noting that "rapid developments in scientific knowledge" have forced courts to readdress the admissibility of scientific evidence).
6. *Melendez-Diaz*, 129 S. Ct. at 2530 (presenting this question).
7. 129 S. Ct. 2527 (2009).
8. *Id.* at 2530.
9. *See id.* at 2532 (holding certificates of analysis subject to the Confrontation Clause).

frontation Clause. This Comment concludes by noting that the practical implications of *Melendez-Diaz* are aligned with the overall goals of the Confrontation Clause.

I. BACKGROUND

A. The Confrontation Clause

The Sixth Amendment sets forth the constitutional rights afforded to defendants in "all criminal prosecutions."[10] The Confrontation Clause—made applicable to the states by virtue of the Fourteenth Amendment[11]—provides that all criminal defendants have the right "to be confronted with the witnesses against" them.[12] As interpreted by the Supreme Court, the Clause ensures that criminal defendants retain the right to be confronted face-to-face with sworn-in witnesses, in the presence of the trier of fact, by reserving the right to cross-examine witnesses that testify against them.[13] Designed to ensure the reliability of evidence,[14] the Clause acts as a safeguard for the basic rights guaranteed to all criminal defendants.[15]

Prior to 2004, out-of-court accusatory statements[16] were admissible in criminal trials if they bore an adequate indicia of reliability.[17] Today, however, the Confrontation Clause requires that a criminal defendant have the opportunity to cross-examine an accusatory witness who is unavailable for trial in order for his or her statements to be admissible.[18]

---

10. U.S. CONST. amend VI ("In all criminal prosecutions, the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed, which district shall have been previously ascertained by law, and to be informed of the nature and cause of the accusation; to be confronted with the witnesses against him; to have compulsory process for obtaining witnesses in his favor, and to have the Assistance of Counsel for his defence.").

11. Pointer v. Texas, 380 U.S. 400, 406 (1965) (holding that the Confrontation Clause applies to both state and federal governments through the Fourteenth Amendment).

12. U.S. CONST. amend. VI.

13. Maryland v. Craig, 497 U.S. 836, 845–46 (1990) (citing California v. Green, 399 U.S. 149, 158 (1970)).

14. Crawford v. Washington, 541 U.S. 36, 61 (2004) (Thomas, J., concurring).

15. *See id.* at 42 (majority opinion) (identifying the Confrontation Clause as a "bedrock procedural guarantee" and noting its importance).

16. *See* Ohio v. Roberts, 448 U.S. 56, 58–61, 77 (1980) (holding a transcript of a witness's preliminary examination testimony constitutionally admissible evidence), *abrogated by* Crawford v. Washington, 541 U.S. 36, 68–69 (2004).

17. *Roberts*, 448 U.S. at 66. "Indicia of reliability" can be defined as those factors or circumstances that allow a jury to hear an out-of-court statement, despite a lack of confrontation of the declarant. *See* Dutton v. Evans, 400 U.S. 74, 88–89 (1970) (finding an indicia of reliability in a statement that contained "no express assertion about past fact," contained a high level of personal knowledge, was not likely to be based on poor memory, and was made spontaneously).

18. *Crawford*, 541 U.S. at 68.

B. Ohio v. Roberts[19] *and the Now Abrogated "Reliability" Test*

In *Ohio v. Roberts*, the Court held that the admission of preliminary hearing testimony without the speaker's presence in court satisfied constitutional standards.[20] In *Roberts*, the defendant was charged with forgery and possession of stolen credit cards.[21] During the preliminary hearing, the defense unsuccessfully tried to elicit an admission from the victim's daughter that she provided the defendant with the checks and credit cards.[22] The victim's daughter never appeared at trial despite being served with five subpoenas.[23] At trial, the defendant testified that the victim's daughter had given him the checkbook and credit cards—along with permission to use them.[24] On rebuttal, the State offered the transcript of the daughter's preliminary hearing testimony as evidence to the contrary.[25] The trial court admitted the evidence and the defendant was convicted.[26]

After the Ohio Supreme Court vacated the conviction,[27] the United States Supreme Court granted certiorari and held that an unavailable witness's out-of-court statement was admissible if the statement bore an adequate "indicia of reliability."[28] The Court reasoned that the preliminary hearing testimony fit this reliability requirement because it "afforded the trier of fact a satisfactory basis for evaluating the truth of the prior statement."[29] The majority further held that reliability could be inferred "where the evidence falls within a firmly rooted hearsay exception"[30] or upon a "showing of particularized guarantees of trustworthiness."[31]

---

19. 448 U.S. 56 (1980).
20. *See id.* at 74–77.
21. *Id.* at 58.
22. *Id.*
23. *Id.* at 59.
24. *Id.*
25. *Id.* (relying on the Ohio statute that permitted the use of preliminary hearing testimony when a witness was unavailable).
26. *Id.* at 60.
27. *Id.* at 60–61.
28. *Id.* at 62, 66.
29. *Id.* at 73 (quoting Mancusi v. Stubbs, 408 U.S. 204, 216 (1972) (internal quotation marks omitted) (noting that "there was an adequate opportunity to cross-examine [the witness], and counsel . . . availed himself of that opportunity" (alterations in original) (internal quotation marks omitted)).
30. *Roberts*, 448 U.S. at 66. "Hearsay" is defined as "a statement, other than one made by the declarant while testifying at the trial or hearing, offered in evidence to prove the truth of the matter asserted." FED. R. EVID. 801(c). Hearsay exceptions include, among others, present sense impression, excited utterance, then existing mental, emotional, or physical condition, statements for purposes of medical diagnosis or treatment, recorded recollection, and records of regularly conducted activity. FED. R. EVID. 803(1)–(6).
31. *Roberts*, 448 U.S. at 66. The Court went on to explain that the purpose of cross-examination is "to challenge 'whether the declarant was sincerely telling what he believed to be the truth, whether the declarant accurately perceived and remembered the matter he related, and whether the declarant's intended meaning is adequately conveyed by the language he employed.'" *Id.* at 71 (quoting David S. Davenport, *The Confrontation Clause and the Co-Conspirator Exception in Criminal Prosecutions: A Functional Analysis*, 85 HARV. L. REV. 1378, 1378 (1972)).

In dissent, Justice Brennan criticized the majority for dispensing with the right of confrontation "so lightly."[32] The decision was also criticized for creating an unpredictable rule that produced results contrary to the intentions of the Confrontation Clause.[33] Abrogated by *Crawford v. Washington*[34] in 2004,[35] the *Roberts* reliability approach was rejected as a "malleable standard [that] often fails to protect against paradigmatic confrontation violations."[36]

C. Crawford v. Washington

In *Crawford v. Washington*, the defendant stabbed the victim after learning that the victim tried to rape his wife.[37] Officers took a recorded statement from the defendant's wife after the stabbing.[38] At trial, the prosecution introduced into evidence the wife's tape-recorded statement describing the stabbing to the police.[39] The defense did not have an opportunity to cross-examine the defendant's wife[40] and the defendant was convicted.[41] The Supreme Court overturned the conviction and created a new standard, holding that testimonial statements in a criminal prosecution are only admissible—absent live testimony—where the witness is unavailable[42] and where the defendant had a prior opportunity to cross-examine the witness.[43] Distinguishing between testimonial and non-testimonial evidence, the majority recognized the "core class of 'testimonial' statements"[44] as including affidavits, depositions, prior statements not subject to cross-examination, and statements "made under circumstances which would lead an objective witness reasonably to believe that the statement would be available for use at a later trial."[45]

The *Crawford* Court abrogated the *Roberts*[46] decision, relying heavily upon the text and history of the Confrontation Clause.[47] The majority

---

32. *Roberts*, 448 U.S. at 82 (Brennan, J., dissenting).

33. *See* Crawford v. Washington, 541 U.S. 36, 63 (2004) (criticizing the *Roberts* test for admitting "core testimonial statements that the Confrontation Clause plainly meant to exclude").

34. 541 U.S. 36 (2004).

35. *See id.* at 68–69.

36. *Id.* at 60.

37. *Id.* at 38.

38. *Id.* at 39–40.

39. *Id.* at 40.

40. *Id.* (explaining that the defendant's wife did not testify because of the state marital privilege law, which barred her from testifying without the defendant's consent).

41. *Id.* at 41.

42. Ohio v. Roberts, 448 U.S. 56, 74 (1980) (describing "unavailability" as the inability to procure a witness despite good-faith attempts to locate and present the witness).

43. *Crawford*, 541 U.S. at 59.

44. *Id.* at 51–52.

45. *Id.* (quoting Brief for the National Ass'n of Criminal Defense et al. as Amici Curiae Supporting Petitioner, Crawford v. Washington, 541 U.S. 36 (2004) (No. 02-9410), 2003 WL 21754961 at *3).

46. *Id.* at 63–64.

47. *See id.* at 60–63 (stating that the *Roberts* reliability approach departed from the historical principles of discouraging the use of *ex parte* evidence and the admission of testimonial statements

reasoned that the *Roberts* reliability approach operated on a model that was "amorphous, if not entirely subjective."[48] The *Crawford* Court went on to conclude that "[w]here testimonial statements are at issue, the only indicium of reliability sufficient to satisfy the constitutional demands is the one the Constitution actually prescribes: confrontation."[49]

D. Davis v. Washington[50]

In *Davis v. Washington*, the Court granted certiorari to the consolidated cases of Adrian Davis and Hershel Hammon in order to further define "testimonial" in the context of out-of-court statements made to law enforcement personnel.[51] Davis's case involved the admissibility of statements made during a 911 call related to an in progress domestic disturbance.[52] The victim made statements to a 911 operator both during the attack and after the defendant left the house.[53] Although,the victim did not appear for trial, the State successfully admitted a portion[54] of the victim's 911 call into evidence—leading to the defendant's conviction.[55] The Court, reviewed the constitutionality of the admitted evidence and held that the statements were nontestimonial, and therefore, immune from the reach of the Confrontation Clause.[56] The majority noted that because the statements were made "about events *as they were actually happening*" and were "necessary to be able to *resolve* the present emergency,"[57] they did not constitute testimonial evidence. Moreover, the Court reasoned that statements made during 911 calls or at crime scenes are generally nontestimonial in nature if "circumstances objectively indicat[e] that the primary purpose of the interrogation is to enable police assistance to meet an ongoing emergency."[58] The Court considered several factors in making this determination: the timing of the statement, the existence of an emergency, the primary purpose of the statement, and the level of formality surrounding the circumstances under which the statements were made.[59]

Hammon's case involved a domestic violence dispute and centered on the admissibility of the victim's written statements in an affidavit

---

not subjected to cross-examination because the test "often fail[ed] to protect against paradigmatic confrontation violations").

48. *Id.* at 63.

49. *Id.* at 68–69.

50. 547 U.S. 813 (2006).

51. *See id.* at 817.

52. *Id.*

53. *Id.* at 817–18.

54. *See id.* at 826-29 (stating that the admissible portion of the call consisted only of the victim's nontestimonial statements like those that relayed information vital to police intervention).

55. *Id.* at 819.

56. *Id.* at 828.

57. *Id.* at 827.

58. *Id.* at 822.

59. *See id.*

recorded by the police after the incident concluded.[60] Because the victim did not appear for the defendant's trial, the State admitted the affidavit as evidence of the defendant's guilt.[61] As a result, the defendant was convicted of battery.[62] The Supreme Court, reversing the conviction, reasoned that the lack of an ongoing emergency, in tandem with police questioning, constituted "part of an investigation into possibl[e] criminal past conduct," rendering the statements testimonial and subject to the Confrontation Clause.[63] The Court explained that a statement is testimonial when "the circumstances objectively indicate that there is no such ongoing emergency, and that the primary purpose of the interrogation is to establish or prove past events potentially relevant to later criminal prosecution."[64]

II. *MELENDEZ-DIAZ V. MASSACHUSETTS*

A. Facts

In 2001, Boston police responded to a tip reporting the suspicious behavior of Thomas Wright.[65] An informant alleged that Wright frequently received phone calls at work that prompted him to leave, wait for a blue sedan outside his place of employment, enter the car, and return in the same vehicle a short time later.[66] After observing this exact succession of events, officers arrested Wright and the two men in the blue sedan—one of whom was Luis Melendez-Diaz—on suspicion of drug possession.[67]

On the drive to the police station with the three men, the officers observed the men "fidgeting and making furtive movements in the back of the car."[68] A search of the police car revealed nineteen small plastic bags containing a substance resembling cocaine hidden in the back seat.[69] A similar substance was found during a personal search of Wright.[70] Police submitted the substance to a state laboratory for chemical analysis and identification.[71]

Laboratory analysts produced and swore to certificates of analysis stating the substance in the bags was found to be cocaine.[72] Melendez-

60. *Id.* at 820.
61. *Id.*
62. *Id.* at 821.
63. *Davis*, 547 U.S. at 829–30 (stating that "the product of the interrogation in *Hammon* is a much easier task, since they were not much different from the statements we found to be testimonial in *Crawford*").
64. *Id.* at 822.
65. Melendez-Diaz v. Massachusetts, 129 S. Ct. 2527, 2530 (2009).
66. *Id.*
67. *Id.*
68. *Id.*
69. *Id.* (noting that the bags were hidden in a partition between the front and back seats).
70. *Id.*
71. *Id.*
72. *Id.* at 2531 (mentioning that the certificates were sworn to before a public notary).

Diaz was charged with distributing and trafficking cocaine.[73] Three certificates of analysis were submitted into evidence at trial, along with the bags seized from Wright and the police car.[74]

B. Procedural History

At trial, Melendez-Diaz argued that the Supreme Court's recent Confrontation Clause decision, *Crawford v. Washington*, required the state lab analyst to offer live testimony.[75] Despite this argument, the trial court admitted the scientific certificates into evidence as "'prima facie evidence of the composition, quality, and net weight of the narcotic . . . analyzed.'"[76] Consequently, the jury convicted Melendez-Diaz of distributing and trafficking cocaine.[77]

On appeal, Melendez-Diaz again asserted a violation of his Sixth Amendment right to confrontation.[78] The Appeals Court of Massachusetts held the admission of the certificates constitutional, relying on a Massachusetts Supreme Judicial Court case,[79] which exempted certificates of analysis from confrontation under the Sixth Amendment.[80] After the Supreme Judicial Court of Massachusetts denied review, the United States Supreme Court granted certiorari.[81]

C. Majority Opinion

In a 5–4 decision authored by Justice Scalia, the Supreme Court reversed the decision of the Massachusetts appellate courts, finding the admission of the certificates unconstitutional.[82] The majority, applying *Crawford*,[83] held that because the certificates of analysis constituted affidavits, they fell within the "core class of testimonial statements" covered by the Confrontation Clause.[84] The Court further held that because Melendez-Diaz was never provided with an opportunity to cross-examine the analyst who prepared the certificates, his right to confront the witness under the Sixth Amendment had been violated.[85]

---

73. *Id.* at 2530.
74. *Id.* at 2530–31.
75. *Id.* at 2531 (noting Melendez-Diaz's objection to the admission of the certificates at trial).
76. *Id.* (alteration in original) (quoting MASS. GEN. LAWS ch. 111, § 13 (2006)).
77. *See id.*
78. *Id.*
79. *Id.*
80. *Id.*
81. *Id.*
82. *Id.* at 2542.
83. *Id.* at 2532–33 (stating that the decision is a "rather straightforward application" of *Crawford*).
84. *Id.* at 2532. The forensic reports contemplated in *Melendez-Diaz* fall firmly within the testimonial category; and therefore the emergency/non-emergency distinction in *Davis* is not useful in the discussion of scientific analysts. *Id.*
85. *Id.*

The majority went on to reject various arguments raised by the State of Massachusetts and the dissent.[86] First, the Court rejected the argument that the non-accusatory nature of the certificates rendered them nontestimonial.[87] Instead, the Court concluded that the certificates themselves constituted the "witnesses against [the defendant],"[88] whom Melendez-Diaz had a right to confront.[89] Second, relying primarily on *Davis*, the Court found that the scientific nature of the testimony was not grounds for removing analysts from the coverage of the Confrontation Clause.[90] Third, the Court dismissed the respondent's arguments that the neutral, scientific nature of the testimony excluded analysts from confrontation as "little more than an invitation to return to our overruled decision in *Roberts*."[91] The majority also held that a defendant's right to subpoena the analyst who prepared the report was not a substitute for his or her right to confrontation.[92]

Finally, the Court reasoned that the demands of the Confrontation Clause may not be relaxed in order to expedite the judicial process.[93] While acknowledging the increased burden its holding places on prosecutors, the Court ultimately rejected the dissent's claim that requiring analysts to testify imposed too high a burden on criminal prosecutions.[94] In doing so, the "simplest form" of notice-and-demand statutes[95] were upheld as a constitutional way to expedite trials of this nature.[96]

D. Justice Thomas's Concurring Opinion

In his concurrence, Justice Thomas clarified his position that out-of-court statements governed by the Confrontation Clause were limited to those "contained in formalized testimonial materials, such as affidavits, depositions, prior testimony, or confessions."[97] Justice Thomas joined the opinion of the Court because the certificates of analysis were "quite plainly affidavits" and therefore, subject to the right of confrontation.[98]

---

86. *Id.* at 2532–42 (responding to "a potpourri" of arguments advanced by the dissent and respondent).

87. *Id.* at 2533.

88. *Id.*

89. *Id.* at 2532.

90. *Id.* at 2535.

91. *Id.* at 2536. The Court relied on *Crawford*'s holding that a statement's purpose will determine whether it is testimonial or nontestimonial. *Id.* at 2532.

92. *Id.* at 2540.

93. *Id.*

94. *Id.*

95. *Id.* at 2541 ("In their simplest form, notice-and-demand statutes require the prosecution to provide notice to the defendant of its intent to use an analyst's report as evidence at trial, after which the defendant is given a period of time in which he may object to the admission of the evidence absent the analyst's appearance live at trial.").

96. *Id.*

97. *Id.* at 2543 (Thomas, J., concurring) (quoting White v. Illinois, 502 U.S. 346, 365 (1992)).

98. *Id.*

E. Dissent

The dissent, authored by Justice Kennedy and joined by Chief Justice Roberts, Justice Breyer, and Justice Alito, criticized the majority for discarding "an accepted rule governing the admission of scientific evidence."[99] Justice Kennedy offered numerous arguments as to why the introduction of scientific analysis into evidence should be allowed without an analyst's testimony.[100] First, he noted that the Court's failure to define the word "analyst" could lead to confusion and inconsistencies as to *which* scientists must testify at trial.[101] Next, Justice Kennedy argued that a number of suitable alternatives were available for defendants who wished to challenge scientific evidence brought against them.[102] These options included serving subpoenas, seeking independent scientific tests, forming opposing arguments, and objecting to the admission of evidence.[103] The dissent further stated that laboratory analysts were not the kind of conventional witnesses subject to the Confrontation Clause.[104] Justice Kennedy distinguished between analysts and conventional witnesses by reasoning that analysts record near-contemporaneous events as opposed to past observations,[105] they do not testify "against the defendant,"[106] and they do not respond to questions under interrogation.[107]

The dissent also listed a number of adverse results that could stem from the majority's decision[108]—including unjust, technical dismissals on account of analysts unavailable to give testimony for reasons such as death, illness, or travel.[109] He further argued that requiring analysts to testify at trial could result in an increase of "not guilty" verdicts,[110] an increase in administrative costs, the creation of a "new prosecutorial duty,"[111] a substantial burden on analysts' time, and an inundation of cases requiring the presence of analysts in court.[112]

---

99. *Id.* (Kennedy, J., dissenting).

100. *See id.* at 2543–53.

101. *See id.* at 2544 (noting that many people play a role in preparing a test's results and the majority opinion does not clarify which of these people the defendant has a right to confront).

102. *See id.* at 2547–49.

103. *Id.* at 2547.

104. *Id.* at 2550–51.

105. *Id.* at 2551.

106. *Id.* at 2552 (noting that analysts rarely have knowledge of the defendant's identity or "of an aspect of the defendant's guilt").

107. *Id.*

108. *See id.* at 2549–50.

109. *Id.* at 2550.

110. *Id.* ("The result, in many cases, will be that the prosecution cannot meet its burden of proof, and the guilty defendant goes free on a technicality that, because it results in an acquittal, cannot be reviewed on appeal.").

111. *Id.* at 2556.

112. *See id.* at 2549–50.

III. ANALYSIS

The *Melendez-Diaz* Court properly aligned its decision with the goals and guarantees of the Confrontation Clause. First, the limitations and deterrents created by *Melendez-Diaz* furthered the fundamental goals of ensuring the reliability of evidence[113] and avoiding the use of *ex parte*[114] examinations as evidence against criminal defendants.[115] Second, the guarantees of face-to-face confrontation[116] and cross-examination[117] remain strong after the Court's decision. Furthermore, *Melendez-Diaz* advanced the Framer's intent of securing rights for criminal defendants by firmly reinforcing the underlying principles behind the Confrontation Clause.

A. Upholding the Fundamental Goals of the Confrontation Clause

The holdings of *Melendez-Diaz* and *Crawford* reveal two primary goals of the Confrontation Clause: (1) ensuring the reliability of evidence,[118] and (2) preventing the use of *ex parte* examinations and accusatory out-of-court statements where the defendant is unable to cross-examine.[119] The Supreme Court recognized the reliability of evidence as the Clause's "ultimate goal."[120] *Melendez-Diaz* maintained this goal by establishing a deterrent for admitting unreliable evidence and broadening the class of testimony subject to the Confrontation Clause.[121] Similarly, *Crawford* identified the use of *ex parte* examinations as the "principal evil" the Confrontation Clause was intended to prevent.[122] The Court's decision in *Melendez-Diaz* advanced this goal by moving Confrontation Clause jurisprudence in a direction that avoids the admission of *ex parte* testimony.

1. Reliability

The *Melendez-Diaz* Court identified the Confrontation Clause's "ultimate goal" as ensuring the reliability of evidence, which is accomplished by requiring that the evidence undergo cross-examination.[123] The Court found that there was nothing uniquely reliable about scientific evidence and it is subject to cross-examination under the Sixth Amend-

113. Crawford v. Washington, 541 U.S. 36, 61 (2003) (Thomas, J., concurring) (identifying reliability of evidence as the Confrontation Clause's "ultimate goal").

114. *See id.* at 49 (describing *ex parte* depositions as statements made by witnesses who have not been subject to cross-examination).

115. *Id.* at 50.

116. Maryland v. Craig, 497 U.S. 836, 847–48 (1990) (discussing the importance of face-to-face confrontation to the Confrontation Clause).

117. *See Crawford*, 541 U.S. at 53–56 (discussing the importance of cross-examination as it relates to the Confrontation Clause).

118. Melendez-Diaz v. Massachusetts, 129 S. Ct. 2527, 2536 (2009).

119. *Id.* at 2531.

120. *Id.* at 2536 (quoting *Crawford*, 541 U.S. at 61).

121. *See Melendez-Diaz*, 129 S. Ct. at 2532.

122. *Crawford*, 541 U.S. at 50.

123. *Melendez-Diaz*, 129 S. Ct. at 2536 (quoting *Crawford*, 541 U.S. at 61).

ment.[124] The *Melendez-Diaz* Court stated that "[c]onfrontation is one means of assuring accurate forensic analysis."[125] Confrontation is meant to expose both fraudulently and honestly produced erroneous evidence.[126] Finally, the right to confrontation does not vanish because other state or federal procedures, such as the ability to subpoena analysts, are aimed at ensuring reliability.[127] Instead, the Confrontation Clause stands as a fundamental safeguard in criminal prosecutions against unreliable evidence.[128]

The Court's decision in *Melendez-Diaz* furthers the goal of ensuring the reliability of evidence by reducing opportunities for false information to go unnoticed and by providing criminal defendants every opportunity to expose fraudulent data.[129] After *Melendez-Diaz*, prosecutors are on notice that their forensic evidence will be subject to the Confrontation Clause.[130] The case, therefore, acts as a deterrent to introducing unreliable evidence in the first place.[131] For example, knowledge that lab analysts are subject to cross-examination will likely compel prosecutors to investigate the reliability of any forensic analysis before trial. Consequently, *Melendez-Diaz* results in more reliable evidence because it discourages the prosecution from initially introducing inadequate evidence.[132]

Furthermore, *Melendez-Diaz* broadened the categories of evidence requiring cross-examination and provided criminal defendants with an opportunity to expose certain weaknesses in the evidence brought against them.[133] As a result, *Melendez-Diaz* ensures that more reliable evidence will be available to the trier of fact by allowing them to consider a broader range of facts surrounding the evidence presented.[134] For example, before *Melendez-Diaz*, some juries would only be presented with the fact that a laboratory affidavit confirmed the identification of a substance.[135] Now, juries will receive information regarding an analyst's proficiency, a machine's calibration, and a lab's reputation.[136] Confrontation may even lead to the discovery that no testing was ever performed

124. *See Melendez-Diaz*, 129 S. Ct. at 2536 (explaining how forensic evidence is not immune to manipulation because forensic scientists often need to answer precise questions regarding a specific case and may feel pressured to compromise methodology for the sake of expediency).

125. *Id.*

126. *Id.* at 2537 ("Confrontation is designed to weed out not only the fraudulent analyst, but the incompetent one as well.").

127. *Id.* at 2540.

128. *See id.* at 2536.

129. *Id.* at 2536–37.

130. *See id.* at 2532.

131. *See id.* at 2537.

132. *See id.*

133. *See id.* at 2532.

134. *See id.*

135. *See id.* at 2554–55 (Kennedy, J., dissenting).

136. 5 CLIFFORD S. FISHMAN & ANNE T. MCKENNA, JONES ON EVIDENCE § 34:27.35 (7th ed. 2010).

on the substance at issue.[137] Conversely, cross-examination may reveal that a lab had a 99.9% accuracy rate and employed the most esteemed analysts in the country. Either way, the trier of fact will gain more information about the evidence brought against the defendant, and will therefore have a deeper understanding about that evidence's reliability.

Finally, had the *Melendez-Diaz* Court affirmed the lower court's decision, the reliability of evidence in criminal prosecutions would have deteriorated.[138] This is true because many kinds of evidentiary facts would be held to very low standards of accountability.[139] The trier of fact would never have the opportunity to learn how some evidence was collected, analyzed, or handled.[140] Juries and judges would lack crucial knowledge about the reliability of certain facts and the truth regarding weaknesses in the prosecution's case.[141] Furthermore, analysts that produce the evidence ultimately used against a defendant at trial might never be called to testify regarding the grave impact of their work product.[142] Had the *Melendez-Diaz* court held differently, the decision would not promote reliable evidence.

2. Avoiding *Ex Parte* Examinations as Evidence

The decision in *Crawford* centered on avoiding the use of *ex parte* evidence.[143] Instead of shifting the goal of the Confrontation Clause away from reliability,[144] *Crawford* simply shed light on another purpose of the Clause. In *Crawford*, the Court stated that "the principal evil at which the Confrontation Clause was directed was the civil-law mode of criminal procedure, and particularly its use of *ex parte* examinations as evidence against the accused."[145] In this statement, the *Crawford* Court referred to the practice of admitting out-of-court statements into evidence without providing the defendant with an opportunity to cross-examine the speaker.[146] *Ex parte* evidence has been deemed "utterly incompe-

137. *See Melendez-Diaz*, 129 S. Ct. at 2536–37 (majority opinion) (discussing "drylabbing"—a procedure in which some analysts prepare scientific reports authenticating results of tests that were never performed).

138. *Id.* at 2536 (discussing how confrontation assures accurate forensic analysis by exposing and deterring fraudulent scientific reports).

139. *Id.* (stating that cross-examination may be the only way to challenge the results of scientific tests that cannot be repeated—like autopsies and breathalyzers).

140. *Id.* at 2537 ("Like expert witnesses generally, an analyst's lack of proper training or deficiency in judgment may be disclosed in cross-examination.").

141. *Id.* at 2535. For example, Melendez-Diaz's trier of fact would be unaware of the fact that the scientific affidavits were completed almost a week after the tests were performed. *Id.*

142. *Id.* at 2533–34.

143. *See* Crawford v. Washington, 541 U.S. 36, 68 (2004).

144. Christopher B. Mueller, *Cross-Examination Earlier or Later: When is it Enough to Satisfy* Crawford?, 19 REGENT U. L. REV. 319, 320 (2006) (advancing the argument that *Crawford* properly shifted the goal of the Confrontation Clause away from reliability).

145. *Crawford*, 541 U.S. at 50.

146. *Id.* (referring specifically to *ex parte* statements wrongly admitted at Sir Walter Raleigh's trial).

tent"[147] and abhorrent[148] because of its unreliable and malleable nature.[149] The Confrontation Clause is a crucial safeguard against this kind of practice.[150]

The holding in *Melendez-Diaz* advanced the goal of avoiding the use of *ex parte* testimony as evidence. By holding that forensic analysts are subject to cross examination, the Court made it clear that Confrontation Clause jurisprudence should move toward, and not away from, allowing confrontation. As the Court reasoned, "the paradigmatic case identifies the core of the right to confrontation, not its limits."[151] *Melendez-Diaz* established these limits in a way that avoids a slippery-slope toward admission of *ex parte* examinations.

B. Upholding the Fundamental Guarantees of the Confrontation Clause

Not only did *Melendez-Diaz* promote the goals at which the Confrontation Clause is aimed, it also upheld the guarantees promised to individual criminal defendants. It is generally accepted that the Clause guarantees the right of face-to-face confrontation and the right to cross-examination.[152] The Court's *Melendez-Diaz* decision guarded these two guarantees, which are fundamental to the right of confrontation.

1. Face-to-Face Confrontation

As the *Crawford* Court noted, the right to face one's accusers dates back to early Roman and English common law.[153] This right allows the trier of fact to observe the witness under questioning,[154] provides an opportunity for an accuser to tell the truth under oath,[155] and gives the defendant an opportunity to face his or her accuser.[156] *Melendez-Diaz* rightly granted defendants the opportunity to come face-to-face with the scientific analysts who produced the evidence used against them at trial. As Justice Scalia opined, "the analyst who provides false results may, under oath in open court, reconsider his false testimony."[157] The re-

147. *Id.* at 49.
148. *Id.* at 48.
149. *Id.* at 50.
150. *Id.*
151. Melendez-Diaz v. Massachusetts, 129 S. Ct. 2527, 2534 (2009) (referring again to Sir Walter Raleigh's trial).
152. Marc C. McAllister, *The Disguised Witness and* Crawford*'s Uneasy Tension with* Craig*: Bringing Uniformity to the Supreme Court's Confrontation Jurisprudence*, 58 DRAKE L. REV. 481, 519 (2010) (referring to "Face-to-Face Confrontation and Cross-Examination as the Two Pillars of Confrontation"). *See generally* Crawford v. Washington, 541 U.S. 36 (2004) (discussing throughout the importance of face-to-face confrontation and cross-examination).
153. *Crawford*, 541 U.S. at 43.
154. *See id.*
155. *See Melendez-Diaz*, 129 S. Ct. at 2536–37.
156. Jeffery L. Fisher, *Preface: Reclaiming Criminal Procedure*, 38 GEO L.J. ANN. REV. CRIM. PROC., at iii, ix (2009) (expounding on the importance of face-to-face confrontation in the criminal justice system).
157. *Melendez-Diaz*, 129 S. Ct. at 2537.

quirement of live confrontation will cause analysts to be more careful in their work and more conscious of its effect at trial. Furthermore, the Court ensured that the right to face-to-face confrontation would not be chipped away at over the years. The majority did this by holding forensic test results to the same standard as any other testimonial evidence offered against criminal defendants.[158] Had the Court held otherwise, it would have undermined the fundamental guarantee to confront one's accuser face-to-face.

2. Cross-Examination

The right to cross-examination is so fundamental that the Supreme Court held testimonial evidence inadmissible if the defendant was not afforded such a right.[159] Cross-examination allows the defense and prosecution alike to show weaknesses in the opposing arguments and reveal evidentiary facts. As the *Melendez-Diaz* Court noted, "[A]n analyst's lack of proper training or deficiency in judgment may be disclosed in cross-examination."[160] By subjecting analysts to cross-examination, *Melendez-Diaz* furthered this protection of the integrity of the adversarial process.

The Confrontation Clause is a constitutional guarantee that is meant to prevent the conviction of the innocent. The dissent in *Melendez-Diaz* worried that subjecting analysts to cross-examination would result in "guilty" defendants going free on purely technical grounds.[161] However, this apprehension is misguided because every defendant is afforded a presumption of innocence,[162] which is removed only when the prosecution proves beyond a reasonable doubt that the defendant is guilty. If, through the "crucible of cross-examination,"[163] among other means, the prosecution fails to meet this burden, the defendant remains innocent. This may result in the dismissal of some cases on technical grounds. However, that result is better than stripping criminal defendants of their constitutional rights based on a presumption of guilt. The *Melendez-Diaz* decision correctly operated under the premise that revoking a defendant's right to cross-examination based on a presumption of guilt would reorder the prosecutorial system.

Additionally, the dissent's "parade of horribles,"[164] if it does ensue, will be short lived. If courts do become swamped with requests for live

158. *Id.* at 2532.
159. *Crawford*, 541 U.S. at 68.
160. *Melendez-Diaz*, 129 S. Ct. at 2537.
161. *Id.* at 2550 (Kennedy, J., dissenting).
162. Coffin v. United States, 156 U.S. 432, 453 (1859) ("The principle that there is a presumption of innocence in favor of the accused is the undoubted law, axiomatic and elementary, and its enforcement lies at the foundation of the administration of our criminal law.").
163. *Crawford*, 541 U.S. at 61.
164. *Melendez-Diaz*, 129 S. Ct. at 2542 (characterizing the dissent's numerous arguments against the majority decision as a "parade of horribles").

testimony of analysts, it will not last for long. Once analysts are called to the stand to testify regarding their work, labs will begin to increase training, accuracy, and accountability.[165] Defense lawyers will soon realize that calling a credible, careful, well-trained analyst is of no help to their case; ultimately resulting in more reliable data, less court costs, and a protection of the criminal defendant's right to cross-examination.

C. Upholding the Historical Intentions Behind the Confrontation Clause

Not only did the *Melendez-Diaz* decision uphold the purposes of the Confrontation Clause, it also upheld the historical principles[166] and intentions behind the Sixth Amendment. The Bill of Rights was drafted as a guarantee of individual rights and freedoms. Each specified right is one the Founders considered important enough to enumerate. The Sixth Amendment is an enumeration of the rights guaranteed to criminal defendants.[167] The Confrontation Clause, specifically, was enacted in response to the concern that evidence would be admitted unjustly against criminal defendants.[168]

By interpreting the Confrontation Clause as a text that imparts, rather than limits, rights of criminal defendants, the *Melendez-Diaz* Court upheld the historical purposes and intentions behind the Clause. The Sixth Amendment grants privileges, it does not limit them. If scientific data was not subject to cross examination, the rights of criminal defendants, as the Founders desired, would deteriorate. As the Founders intended, *Melendez-Diaz* protected the rights of defendants.

D. Pendergrass v. State[169]*: Begging an Unanswered Question of Melendez-Diaz*

Eventually, the Supreme Court will need to address some unanswered questions created by *Melendez-Diaz*. Which analyst is required to testify is one such uncertainty.[170] While the Court's *Melendez-Diaz* decision did "not mean that everyone who laid hands on the evidence must be called,"[171] it did not specify any requirements for selecting which analyst should testify. Many phases are required for the production of most

---

165. Mark Hansen, *Taking Techs to Trial: Two Terms in a Row, Justices Weigh Bringing Lab Analysts into Court*, 96 JAN. A.B.A. J. 17, 18 (2010) ("Stanford University law professor Jeffery L. Fisher, who represented Melendez-Diaz, says the decision will help ensure that analysts will be careful when they do their testing and will be held accountable when they make mistakes.")

166. *See* Justin Chou, *Melendez-Diaz v. Massachusets: Raising the Confrontation Requirements for Forensic Evidence in California*, 14 BERKLEY J. CRIM. L. 439, 442–43 (2009) (discussing various historical principles behind the Confrontation Clause).

167. U.S. CONST. amend. VI.

168. *See Crawford*, 541 U.S. at 49 (describing concerns of some early Americans regarding the admission of evidence at trial and noting that "[t]he First Congress responded by including the Confrontation Clause in the proposal that became the Sixth Amendment").

169. 913 N.E.2d 703 (Ind. 2009), *cert. denied*, 130 S. Ct. 3409 (2010).

170. *See Melendez-Diaz*, 129 S. Ct. at 2544 (Kennedy, J., dissenting).

171. *Id.* at 2532 (majority opinion).

scientific reports submitted at trial. For example, DNA test results alone go through several phases of analysis and exchange hands many times before the results are finalized. Each report involves the work product and supervision of numerous scientists.

The Supreme Court recently denied certiorari to *Pendergrass v. State*,[172] a case involving DNA certificates of analysis.[173] In *Pendergrass*, the prosecution submitted DNA and paternity test results as evidence against the defendant.[174] Pursuant to *Melendez-Diaz*, a supervisor employed by the laboratory that produced the DNA results offered live testimony at trial. The supervisor, Lisa Black, testified to the certificates of analysis produced by her employers and general lab procedures, among other things.[175] Ms. Black did not personally conduct the DNA testing.[176] The defendant was convicted and appealed based on a Confrontation Clause violation, arguing that *Melendez-Diaz* required the analyst who actually produced the results to offer live testimony.[177] The Supreme Court of Indiana upheld the conviction and the United States Supreme Court denied certiorari.

Pendergrass was correctly decided in light of the goals of the Confrontation Clause. The analyst offering live testimony should be one who is in a position to testify to the general and specific scientific procedures that lead to the data submitted as evidence against the defendant. The purpose behind the analysts live testimony is to ensure reliable evidence through cross examination. Subjecting Ms. Black to cross-examination accomplished this goal. First, she was able to testify to the general testing procedures that were used in conducting the scientific tests. Second, she was able to testify to the specific procedure that occurred in order to produce the certificates of analysis that were ultimately submitted at trial. In doing so, she offered the trier of fact information regarding the general reliability of DNA testing in her laboratory and specific information regarding the reliability of the evidence presented against the defendant.

For now, courts are left to their own discretion in determining which analyst's testimony is sufficient to satisfy the Confrontation Clause requirements set forth in *Melendez-Diaz*. These determinations must be made in light of the goal of achieving reliability through cross-examination. However, the Supreme Court may eventually need to address this unanswered question.

172. 913 N.E.2d 703 (Ind. 2009), *cert. denied*, 130 S. Ct. 3409 (2010).
173. *See Pendergrass*, 913 N.E.2d at 704–05.
174. *Id.*
175. *Id.*
176. *Id.* at 705.
177. *Id.* at 708.

CONCLUSION

As new technologies expand, courts must determine what role, if any, these scientific developments will play in criminal prosecutions. These determinations must be made in light of and in a manner consistent with constitutional guarantees and limits. Accordingly, the *Melendez-Diaz* holding was decided correctly in light of the goals and guarantees of the Confrontation Clause. The decision promotes the reliability of evidence and opposes the use of *ex parte* examinations; it upholds the fundamental Confrontation Clause guarantees of face-to-face confrontation and cross-examination. Finally, the *Melendez-Diaz* Court upheld the historical goals behind the Sixth Amendment by protecting the rights of criminal defendants. The long-term, practical implications of *Melendez-Diaz* remain to be seen. If negative conditions do result, these consequences will be short lived. The true long-term effects of *Melendez-Diaz* will be more reliable evidence, more accountability in the criminal justice system, more information available to the trier of fact, and a strong guarantee of the rights of criminal defendants.

Laura Bowzer[*]

* J.D. Candidate, 2012, University of Denver Sturm College of Law. I would like to thank the *Denver University Law Review* board and staff along with Professor Shelia Hyatt for their valuable time and edits in developing this comment. Thanks also to Michael Dougherty and Larry Adkisson for taking the time to teach and inspire me; and to Matt Arentsen for supplying me with endless encouragement and comedy. Above all, I would like to thank my family, especially my mother Janice, my father Jim, and my brother Kyle who taught me to work hard, have integrity, and laugh.

MARYLAND V. SHATZER: STAMPING A FOURTEEN-DAY EXPIRATION DATE ON *MIRANDA* RIGHTS

INTRODUCTION

Over forty years ago, the United States Supreme Court established a suspect's right to be informed of his rights to counsel and silence in *Miranda v. Arizona*.[1] Today, *Miranda* rights inundate American televisions, movie screens, and perceptions of criminal justice.[2] In its controversial decision, the *Miranda* Court used the Self-Incrimination Clause of the Fifth Amendment as the foundation for rights of the accused.[3] Although *Miranda* has been a part of American culture since the decision was handed down in 1966, the Court is still fine-tuning the application of *Miranda* rights.[4] Most recently, the Court held in *Maryland v. Shatzer*[5] that a suspect's invocation of the right to counsel is only powerful enough to prevent further questioning by law enforcement for fourteen days after the suspect's release from custody.[6]

This Comment explores the flaws, inconsistencies, and impact of the *Shatzer* fourteen-day rule. Overall, the *Shatzer* Court lost sight of the prophylactic ideas of *Miranda* in its quest for an easy standard—jeopardizing not only the accused's right to counsel but also his right to remain silent. Furthermore, even though these constitutional rights are more valuable to suspects today than they were at the time *Miranda* was decided, the *Shatzer* fourteen-day rule continues the Court's pattern of gradually deteriorating suspects' *Miranda* rights.

Part I of this Comment briefly describes the Court's development and clarification of *Miranda* rights, highlighting the topics most altered by *Shatzer*. Part II summarizes the facts, procedural history, and opinions of *Shatzer*. Part III asserts four propositions: (1) *Shatzer* continues the Court's retreat from the prophylactic principles of *Miranda*, further compromising the right to remain silent and the right to counsel; (2) the Court's retreat wrongly abandoned prophylactic measures in favor of efficiency; (3) the Court's fourteen-day rule compromises a suspect's *Miranda* rights at a time when those rights are increasingly valuable and

1. Miranda v. Arizona, 384 U.S. 436, 444 (1966).

2. *See* Dickerson v. United States, 530 U.S. 428, 443 (2000) ("*Miranda* has become embedded in routine police practice to the point where the warnings have become part of our national culture.").

3. *Miranda*, 384 U.S. at 458, 467.

4. *See, e.g.*, Davis v. United States, 512 U.S. 452, 461–62 (1994) (holding that a suspect's request for counsel must be unambiguous); Edwards v. Arizona, 451 U.S. 477, 484–85 (1981) (holding that after a suspect has invoked his right to counsel, questioning cannot resume until the suspect has obtained counsel or the suspect initiates discussion).

5. 130 S. Ct. 1213 (2010).

6. *Id.* at 1223.

decisive to his case; and, (4) the Court's focus on the *Edwards* rule, instead of *Miranda* rights, makes it easier for the Court to continue to curtail the rights of the accused. This Comment concludes that *Shatzer* was wrongly decided, and that the decision will have a detrimental effect on the fair administration of criminal justice in America.

I. BACKGROUND

Prior to *Miranda*'s landmark ruling in 1966, the only way for a defendant to attack the prosecution's use of his confession made before indictment or the filing of charges was by bringing a due process claim.[7] A defendant cannot rely on the Sixth Amendment to challenge confessions made during initial interrogations because the right to counsel—guaranteed by the Sixth Amendment—attaches only when prosecution formally commences, "whether by way of formal charge, preliminary hearing, indictment, information, or arraignment."[8] Although constitutional challenges under the Due Process Clause remained available to defendants, the prophylactic measures established by the *Miranda* Court provided additional protections to ensure that a suspect's constitutional rights were fully honored.

However, because the rights guaranteed by *Miranda* were not explicitly found in the text of the Constitution,[9] its holding has been subject to several challenges—resulting in numerous exceptions to the *Miranda* holding. Although the Court has kept *Miranda*'s mandate alive, its pattern of fashioning exceptions to *Miranda*'s application has slowly deteriorated the rights that it previously found indispensible to suspects in custodial interrogation.

A. Traditional Constitutional Challenges to the Admissibility of Statements

The Self-Incrimination Clause of the Fifth Amendment provides that "[n]o person . . . shall be compelled in any criminal case to be a witness against himself."[10] Prior to 1966, the United States Supreme Court interpreted this right literally to mean that a criminal defendant cannot be compelled to testify in his own criminal proceeding.[11] Therefore, the

7. *Dickerson*, 530 U.S. at 433 ("[F]or the middle third of the 20th century our cases based the rule against admitting coerced confessions primarily, if not exclusively, on notions of due process.").

8. Brooks Holland, 99 J. CRIM. L. & CRIMINOLOGY 381, 390 (2009) (quoting McNeil v. Wisconsin, 501 U.S. 171, 175 (1991)).

9. Miranda v. Arizona, 384 U.S. 436, 490 (1966). Although there is no right to be advised of one's rights in the Constitution, the Constitution's relation to *Miranda* warnings has been debated by the Court. *Compare Dickerson*, 530 U.S. at 444 (concluding "that *Miranda* announced a constitutional rule that Congress may not supersede legislatively"), *with Dickerson*, 530 U.S. at 454 (Scalia, J., dissenting) (accusing the majority of playing "word games" to make *Miranda* a constitutional mandate).

10. U.S. CONST. amend. V.

11. William T. Pizzi & Morris B. Hoffman, *Taking* Miranda*'s Pulse*, 58 VAND. L. REV. 813, 814–15 (2005) (explaining that prior to the *Miranda* decision, the United States Supreme Court held

Self-Incrimination Clause cannot single-handedly protect suspects from police coercion during interrogations outside of any criminal proceedings or prevent the admission of any evidence obtained from such coercion at trial.[12] Similarly, the Fifth Amendment could not provide suspects with counsel during interrogation because the Sixth Amendment alone governed the right to counsel.[13]

In the mid-twentieth century, the Court became increasingly concerned about coerced confessions obtained through dishonest and threatening police interrogations.[14] Although the bare text of the Self-Incrimination Clause did not encompass police tactics prior to trial, the Court sought to condemn and prevent involuntary confessions[15] by looking to the Due Process Clause of the Fifth and Fourteenth Amendments.[16] Unlike voluntary confessions, coerced confessions offend due process by forcing an individual to incriminate himself, thereby preventing him from attaining a fair trial.[17] Framing the issue around due process, the Court established a totality of the circumstances inquiry to determine the voluntariness of a confession in *Johnson v. Zerbst*.[18] This analysis evaluated whether a confession was truly voluntary by determining "whether the defendant's will was overborne at the time he confessed."[19]

While the due process approach afforded defendants broader rights than the Self-Incrimination Clause, the Due Process Clause had its own set of limitations.[20] Notably, it did not affirmatively protect suspects from coercion, but only afforded defendants the opportunity to challenge any involuntary statements at trial.[21] Furthermore, because interrogations

---

that a "case" meant the actual criminal proceeding, and "compelled" applied only to a defendant's right to not be held in contempt for refusing to testify at the proceeding).

12. *See* Michael J. Zydney Mannheimer, *Ripeness of Self-Incrimination Clause Disputes*, 95 J. CRIM. L. & CRIMINOLOGY 1261, 1317–18, 1323 (2005) (arguing that the United States Supreme Court violates the Constitution by hearing claims based on the Self-Incrimination Clause before charges have been filed or a criminal proceeding has commenced because such claims are not yet ripe for adjudication under the plain language of the Fifth Amendment).

13. U.S. CONST. amend. VI ("In all criminal prosecutions, the accused shall enjoy the right . . . to have the Assistance of Counsel for his defence.").

14. Dickerson v. United States, 530 U.S. 428, 434–35 (2000) ("In *Miranda*, we noted that the advent of modern custodial police interrogation brought with it an increased concern about confessions obtained by coercion.").

15. Miranda v. Arizona, 384 U.S. 436, 458–59 (1966) (tracking the Court's historical disapproval of coerced confessions, which finds its roots in the Star Chamber Oath).

16. *Dickerson*, 530 U.S. at 433 ("[F]or the middle third of the 20th century our cases based the rule against admitting coerced confessions primarily, if not exclusively, on notions of due process."); *see, e.g.*, Haynes v. Washington, 373 U.S. 503, 515 (1963); Lynumn v. Illinois, 372 U.S 528, 537 (1963); Payne v. Arkansas, 356 U.S. 560, 568 (1958).

17. *See Haynes*, 373 U.S. at 515; *Lynumn*, 372 U.S. at 534, 537.

18. 304 U.S. 458, 464 (1938).

19. *Lynumn*, 372 U.S. at 534.

20. Chavez v. Martinez, 538 U.S. 760, 796 (2003) (Stevens, J., concurring in part and dissenting in part).

21. Note, *Procedural Protections of the Criminal Defendant – A Reevaluation of the Privilege Against Self-Incrimination and the Rule Excluding Evidence of Propensity to Commit Crime*, 78 HARV. L. REV. 426, 430–31 (1964) (asserting that only preventing the admission of involuntary

were often conducted secretly, defendants would struggle to prove the use of coercive interrogation tactics in court over the contradictory testimony of law enforcement officers.[22] As a result, only extremely visible instances of police brutality and deceit were found to violate due process.[23] Due to concerns about more subtle and sophisticated police tactics that made coercion often difficult to ascertain, the Court sought additional protections for suspects in custodial interrogation.[24]

B. Additional Safeguards to the Right against Self-Incrimination: Miranda v. Arizona

In an "unprecedented stretch of the language of the Self-Incrimination Clause," the *Miranda* Court imposed an affirmative obligation on law enforcement to prevent the occurrence of involuntary statements.[25] In *Miranda,* Ernesto Miranda was accused of kidnapping and rape.[26] Miranda was taken into custody and questioned without first being advised that he had a right to have an attorney present.[27] After being interrogated, Miranda eventually confessed.[28] At trial, the State presented evidence of Miranda's confession over his objection.[29]

In considering whether Miranda's confession was properly admitted at trial, the Supreme Court addressed "the necessity for procedures which assure that [an] individual is accorded his privilege under the Fifth Amendment to the Constitution not to be compelled to incriminate himself."[30] To set the stage for its landmark holding, the *Miranda* Court depicted the nature and setting of in-custody interrogations, focusing on recent studies and case law revealing police brutality during such ques-

---

statements at trial does not provide enough protection to defendants because juries may assume the parties are debating over a confession when the defense objects to evidence the prosecution attempts to admit, and the prosecutor's knowledge of a confession, admissible at trial or not, could heavily influence his decision whether or not to press charges).

22. *Id.* at 431.

23. *See, e.g.*, Leyra v. Denno, 347 U.S. 556, 558–59 (1954) (finding the defendant's confession involuntary because defendant had been questioned on different days for eight hours, fourteen hours, and twenty-three hours respectively, and during the last session a police psychiatrist, posing as a medical doctor to treat the suspect's sinus infection, attempted to hypnotize the suspect); Brown v. Mississippi, 297 U.S. 278, 284 (1936) (finding the defendant's confession involuntary because defendant had been whipped and tortured over several days).

24. *See, e.g.*, Escobedo v. Illinois, 378 U.S. 478, 490–91 (1964) (holding that absent the right to counsel and an opportunity to remain silent, any incriminating statements obtained from a suspect in custody were inadmissible at trial). The defendant in *Escobedo* was accused of murder, held in custody, and not advised of his constitutional rights. *Id.* at 479, 481. The *Escobedo* Court found that any rights lost during interrogation were irrevocably lost and therefore interfered with any rights guaranteed to the accused during trial. *See id.* at 486. In other words, the *Escobedo* Court found that although the Constitution's language only applied to the courtroom, certain rights could be curtailed before a defendant ever reached the courtroom. *See id.*

25. Pizzi & Hoffman, *supra* note 11, at 815–16.

26. Miranda v. Arizona, 384 U.S. 436, 491 (1966).

27. *Id.* at 492.

28. *Id.*

29. *Id.*

30. *Id.* at 439.

tioning.[31] The Court was especially concerned with incriminating statements made by defendants who faced more subtle police tactics and "inherently compelling pressures"[32] of custodial interrogation that were not egregious enough to warrant protections under the Due Process Clause. Thus, the Court returned to the Self-Incrimination Clause of the Fifth Amendment in order to provide a more effective set of protections to defendants in these circumstances.[33]

Although the Fifth Amendment does not provide a textual right to counsel, the *Miranda* Court considered the right necessary to secure the explicit privilege in the Self-Incrimination Clause to remain silent.[34] Pursuant to the Fifth Amendment, the right to remain silent was an established principle at the time of *Miranda*.[35] The *Miranda* Court—aiming to ensure that a suspect's choice to communicate with the police was voluntary throughout the interrogation process—reasoned that the presence of an attorney would make a suspect more confident and able to remain silent if desired.[36] The Court found that in order to give meaning to a suspect's right to silence, "the compulsion inherent in custodial interrogation [had] to be diffused by warning the suspect not only of his right to silence, but of his right to an attorney" as well.[37] For this reason, the Court held that the right to counsel, though not a textual right in itself, was an indispensable companion to the fundamental right to remain silent.[38]

C. *Effects of* Miranda

Miranda's mandate was clear: prior to any questioning, the authorities must warn a suspect that he has the right to remain silent and the right to an attorney.[39] The Court made equally clear that any questioning must automatically cease once a suspect invokes his right to remain silent.[40] Similarly, when a suspect invokes his right to counsel, the interrogation cannot continue until counsel is present.[41] If a suspect cannot ob-

---

31. *Id.* at 445–47; *see also id.* at 446 n.7 (citing multiple cases involving police brutality). The Court was particularly disturbed by police manuals that described in detail how to psychologically disadvantage suspects and extract confessions. *Id.* at 448–55; *see also id.* at 449 n.8 (referring to several of the manuals then in use by the police).

32. *Id.* at 467.

33. *See id.* at 442.

34. *Id.* at 466; Marcy Strauss, *Understanding* Davis v. United States, 40 LOY. L.A. L. REV. 1011, 1015 (2007).

35. Malloy v. Hogan, 378 U.S. 1, 8 (1964) (holding that a person has the right "to remain silent unless he chooses to speak in the unfettered exercise of his own will, and to suffer no penalty . . . for such silence").

36. *Miranda*, 384 U.S. at 469–70.

37. Donald P. Judges & Stephen J. Cribari, *Speaking of Silence: A Reply to Making Defendants Speak*, 94 MINN. L. REV. 800, 812 (2010).

38. *Id.* at 812–13.

39. *Miranda*, 384 U.S. at 444.

40. *Id.* at 473–74.

41. *Id.* at 474.

tain an attorney on his own, law enforcement must either accept his decision to remain silent or provide him with counsel.[42]

Adherence to these rules has become a prerequisite for the admissibility of any statement made by a defendant at his criminal trial.[43] The *Miranda* holding "drastically overhauled the law of police interrogations"[44] by imposing a "positive obligation on police to advise suspects of a given litany of rights before any custodial interrogation could begin."[45] Accordingly, most of *Miranda*'s critics attack the decision for meddling with law enforcement procedures,[46] asserting that it prevents the admission of voluntary confessions in criminal trials.[47] Specifically, extending the application of *Miranda* rights may deter police from trying to obtain voluntary confessions, which are "essential to society's compelling interest in finding, convicting, and punishing those who violate the law."[48] Although a suspect may waive his *Miranda* rights, the government has the heavy burden of demonstrating that the defendant knowingly and intelligently waived his privilege.[49] The Court openly acknowledged that it imposed a heavy burden on the government, but averred that because the government is in a position of authority throughout the interrogation process, that "burden is rightly on its shoulders."[50] Furthermore, the American criminal justice system places the burden of proof wholly on the government for every element of a crime, including proof of the voluntariness of any confession offered as evidence.[51]

D. A "Second Layer of Prophylaxis": Edwards v. Arizona[52]

Fifteen years after *Miranda*, the Court buttressed the accused's right to counsel in *Edwards v. Arizona* by clarifying that a custodial interrogation cannot be resumed until the protections articulated in *Miranda* have been provided.[53] In *Edwards*, the defendant was charged with robbery, burglary, and first-degree murder.[54] Pursuant to an arrest warrant, Edwards was detained and interrogated by law enforcement officials after

42. *Id.*
43. *Id.* at 476.
44. Strauss, *supra* note 34, at 1014.
45. Pizzi & Hoffman, *supra* note 11, at 817.
46. *See id.* at 817 n.20.
47. *See* Maryland v. Shatzer, 130 S. Ct. 1213, 1221–22 (2010).
48. *Id.* at 1222 (quoting McNeil v. Wisconsin, 501 U.S. 171, 181 (1991)).
49. *Miranda*, 384 U.S. at 475 (citing Johnson v. Zerbst, 304 U.S. 458, 465 (1938) (holding that the Constitution "imposes the serious and weighty responsibility upon the trial judge of determining whether there is an intelligent and competent waiver by the accused")).
50. *Miranda*, 384 U.S. at 475.
51. *Id.* at 460; *see also* Judges & Cribari, *supra* note 37, at 806 (noting that in the late eighteenth century to early nineteenth century, American criminal justice switched from an accused-speaks model to a testing-the-prosecution model).
52. 451 U.S. 477 (1981). The *Edwards* holding was labeled a "second layer of prophylaxis." *McNeil v. Wisconsin*, 501 U.S. 171, 176 (1991).
53. *Edwards*, 451 U.S. at 484–85.
54. *Id.* at 478.

being properly informed of his *Miranda* rights.[55] Questioning quickly ceased after Edwards denied involvement in the crimes and requested to speak with an attorney.[56] However, when officers visited Edwards at the county jail the next morning, the jail's guard told Edwards that he "had to" speak with them.[57] As a result, Edwards spoke to officers and implicated himself in the crimes, even after officers again informed him of his *Miranda* rights.[58]

In its review, the *Edwards* Court considered whether the defendant had voluntarily waived his right to counsel by speaking with law enforcement at the second interrogation.[59] Relying heavily on its rationale in *Miranda*, the Court held that the waiver was involuntary and that the confession was inadmissible at trial.[60] The Court reasoned that once a suspect initially invokes his right to counsel, any subsequent waivers of that right are presumed involuntary because such waivers are likely the result of police coercion, badgering, or dishonesty.[61] Therefore, when Edwards asserted his right to an attorney on the night of his arrest, the police were required to honor his desire to communicate with law enforcement only through counsel for the remainder of the investigation.[62]

However, the *Edwards* decision permitted questioning to resume if the suspect initiated the discussion with law enforcement.[63] The Court reasoned that where the accused initiated the discussion, the risk of any police coercion was minimal and the presumption of involuntariness no longer applied.[64] The Court found that the effect of a suspect's assertion of his right to counsel differed from a suspect's invocation of the right to remain silent, which only temporarily paused the interrogation.[65]

Although a seemingly bright-line rule, the Court has since been inundated with proposed exceptions to the *Edwards* application. In nearly all of the challenges to *Edwards*, "[T]he Court was concerned with preserving the clear, bright-line nature of the *Edwards* decision."[66] For ex-

55. *Id.*
56. *Id.* at 479.
57. *Id.*
58. *Id.*
59. *Id.* at 482–84.
60. *Id.* at 487.
61. *See id.* at 484–85.
62. *Id.* at 484–86.
63. *Id.* at 484–85.
64. *See id.* at 484–86 & n.9.
65. Marcy Strauss, *The Sounds of Silence: Reconsidering the Invocation of the Right to Remain Silent Under* Miranda, 17 WM. & MARY BILL RTS. J. 773, 818–19 (2009) (asserting that invoking the right to counsel has more dire consequences for law enforcement than does asserting the right to remain silent). *Compare Edwards*, 451 U.S. at 484–85 (holding that when a suspect invokes the right to counsel, the interrogation must cease and cannot resume until counsel is made available or the suspect initiates discussion), *with* Michigan v. Mosley, 423 U.S. 96, 102–04 (1975) (holding that although an interrogation must immediately cease upon assertion of the right to remain silent, it does not follow that law enforcement may not resume questioning two hours later).
66. Strauss, *supra* note 34, at 1022.

ample, the Court applied the *Edwards* rule to interrogations concerned with unrelated offenses, forbidding police from questioning a suspect if he had asserted his right to counsel during a prior interrogation for an unrelated offense.[67] Additionally, the Court applied the *Edwards* rule when a suspect had the opportunity to consult with counsel, but did not have counsel present for questioning.[68] For about a decade, the right to counsel was a powerful and effective protection against deceitful interrogation techniques. Despite several challenges and critiques, *Edwards* rendered any police-initiated confessions made after an assertion of the right to counsel per se involuntary.

E. The Court's Gradual Retreat from Additional Prophylaxes

Although *Edwards* secured a suspect's right to counsel, the Court crafted various exceptions to other aspects of *Miranda*'s application in the decades following the advent of *Miranda* rights.

Davis v. United States[69] established a notable limitation on *Miranda*'s application. In *Davis,* the defendant was accused of murder and initially waived his *Miranda* rights during an interview.[70] However, an hour-and-a-half into the interview, Davis stated that he might want to speak with a lawyer.[71] The defendant's interviewers testified that they asked Davis if he meant that he wanted a lawyer, to which Davis allegedly answered, "No, I'm not asking for a lawyer."[72] After a short break, the interview continued for another hour until Davis stated, "I think I want a lawyer before I say anything else."[73] At trial, Davis moved to suppress statements made during the interview.[74] Specifically, Davis claimed that his statement, "Maybe I should talk to a lawyer," constituted an invocation of his right to counsel, and that based on *Miranda* and *Edwards*, the interrogation should have ceased until that right was fully honored.[75]

In its review, the *Davis* Court considered how law enforcement officers should respond when a suspect makes a reference to counsel that is insufficiently clear to invoke the *Edwards* prohibition on further questioning.[76] Stressing the need for effective law enforcement, the Court held that interrogations may continue unless a suspect clearly and une-

---

67. Arizona v. Roberson, 486 U.S 675, 687 (1988); *see also* Thomas N. Radek, Note, Arizona v. Roberson*: The Supreme Court Expands Suspects' Rights in the Custodial Interrogation Setting,* 22 J. MARSHALL L. REV. 685, 686 (1989).

68. Minnick v. Mississippi, 498 U.S. 146, 153 (1990).

69. 512 U.S. 452 (1994).

70. *Id.* at 454.

71. *Id.* at 455. Specifically, Davis said, "Maybe I should talk to a lawyer." *Id.*

72. *Id.*

73. *Id.*

74. *Id.*

75. *See id.* at 459.

76. *See id.* at 454.

quivocally requests an attorney.[77] This caveat that a request for counsel must be unambiguous introduced an element of uncertainty to the *Edwards* rule and limited its reach.[78] Paradoxically, the Court highlighted the importance of *Miranda* rights,[79] while making the invocation of those rights difficult.

Other notable exceptions to *Miranda* provided loopholes for admitting statements obtained without advising a suspect of his rights, as required by *Miranda*. In *Michigan v. Tucker*,[80] the Court held that the exclusion of the "fruits" of a *Miranda* violation—the statement of a witness whose identity the defendant revealed while in custody—was not required.[81] Additionally, the Court ruled in *Oregon v. Haas*[82] that voluntary statements obtained without advising a suspect of his *Miranda* rights could be used to impeach a defendant at trial.[83] In *New York v. Quarles*,[84] the Court created a "public safety" exception that freed law enforcement from *Miranda* requirements if questioning needed to occur quickly to secure the safety of the public.[85] Despite these exceptions, in *Dickerson v. United States*,[86] the Court asserted the continued importance and survival of *Miranda*'s core holding when it invalidated an act of Congress meant to overrule *Miranda* because *Miranda* was itself a "constitutional" holding.[87] In its ruling, the Court rejected the idea that the advisement of *Miranda* rights was merely a factor for a court to consider in determining the voluntariness of a statement.[88]

Although the Court has declined to overrule *Miranda*, it has begun to limit the application of the *Edwards* rule. Prior to its decision in *Maryland v. Shatzer*, the Court had only broadly recognized a time limit to the application of the *Edwards* case, noting, in dicta, that *Edwards* applied "assuming there has been no break in custody."[89] The Court would next have to determine what exactly constituted a "break in custody." The

77. *Id.* at 460–61. In contrast to *Miranda* and *Edwards*, the *Davis* Court appeared to value law enforcement efficiency more than rights of the accused. According to *Davis*, the primary benefit of *Miranda* was the *advisement* of rights. It was then up to the suspect to unambiguously invoke those rights. *See id.*

78. *See* Strauss, *supra* note 34, at 1027–28.

79. *Davis*, 512 U.S. at 458.

80. 417 U.S. 433 (1974).

81. *Id.* at 450–52.

82. 420 U.S. 714 (1975).

83. *Id.* at 723–24.

84. 467 U.S. 649 (1984).

85. *Id.* at 653. In a grocery store, police apprehended a rape suspect known to be carrying a gun, did not find the gun on his person, and then asked him where he had put the gun. *Id.* at 652. The suspect answered, "[T]he gun is over there." *Id.* The statement was ruled admissible under a "public safety" exception. *Id.* at 659–60.

86. 530 U.S. 428 (2000).

87. 18 U.S.C. § 3501 (2000), *invalidated by* Dickerson v. United States, 530 U.S. 428 (2000).

88. *Dickerson*, 530 U.S. at 442–44.

89. Maryland v. Shatzer, 130 S. Ct. 1213, 1220 (2010) (quoting McNeil v. Wisconsin, 501 U.S. 171, 177 (1991)).

answer revealed just how far the Court was willing to extend the prophylactic measures it instituted in *Miranda* and solidified in *Edwards*.

II. *MARYLAND V. SHATZER*

In *Maryland v. Shatzer*, the United States Supreme Court considered whether a suspect's invocation of his right to an attorney indefinitely shields the suspect from further questioning until he hires or is provided an attorney. In *Shatzer*, the defendant was re-interrogated for the same charge two-and-a-half years after he asserted his right to an attorney concerning that charge. Although the entire Court agreed that two-and-a-half years was a sufficient time period for the suspect's invocation of *Miranda/Edwards* rights to expire, the majority insisted on pinpointing exactly how long law enforcement must honor a suspect's request for an attorney. Despite criticism from two concurring Justices, the majority held that a suspect's assertion of his right to an attorney guarded the suspect from further interrogation without an attorney present for only fourteen days.

A. Facts

In *Shatzer*, a detective initially visited the defendant, Michael Shatzer, Sr., in 2003 to question him about allegations of sexually abusing his own son.[90] At the time, Shatzer was serving a sentence for an unrelated child sexual-abuse offense at the Maryland Correctional Institution-Hagerstown.[91] When he learned the reasoning behind the detective's visit, Shatzer declined to speak to the detective without an attorney present.[92] The detective then ended the visit and closed the case.[93]

Two-and-a-half years later, the case was re-opened based on additional evidence obtained from Shatzer's son.[94] Investigators visited Shatzer at the Roxbury Correctional Institute, where Shatzer had been transferred.[95] This time, Shatzer waived his *Miranda* rights and consented to a polygraph examination.[96] During the interview, he "admitted to masturbating in front of his son at a distance of less than three feet."[97] Later, after failing the polygraph, Shatzer told police that he "didn't force" his son to perform fellatio on him, thereby admitting that the act had occurred.[98] Following this admission, Shatzer requested an attorney and the detectives ended the interrogation.[99]

---

90. *Id.* at 1217.
91. *Id.*
92. *Id.*
93. *Id.*
94. *Id.* at 1217–18.
95. *Id.*
96. *Id.* at 1218.
97. *Id*
98. *See id.*
99. *Id.*

B. Procedural History

"The State's Attorney for Washington County, Maryland, charged Shatzer with second-degree sexual offense, sexual child abuse, second-degree assault, and contributing to conditions rendering a child in need of assistance."[100] In response, Shatzer argued that the *Edwards* protections rendered his 2006 waiver involuntary and moved to suppress his statements from that day.[101] Shatzer pled not guilty and waived his right to a jury trial.[102]

The trial court denied Shatzer's motion and found him guilty of sexually abusing his son.[103] The Court reasoned that *Edwards* did not apply because the two-and-a-half-year time period separating the two interrogations constituted a sufficient break in custody to allow his previously asserted *Miranda* rights to expire.[104] The Court of Appeals of Maryland reversed and remanded, holding that: (1) the passage of time alone was insufficient to end *Edwards* protections; and (2) if a break-in-custody exception to *Edwards* existed, Shatzer's release back into prison did not constitute such a "break in custody."[105] The United States Supreme Court granted certiorari to determine whether, and at what point, "a break in custody ends the presumption of involuntariness established in *Edwards*."[106]

C. Majority Opinion

Justice Scalia wrote the opinion of the Court, with Chief Justice Roberts and Justices Kennedy, Ginsburg, Breyer, Alito, and Sotomayor joining in the decision.[107] The majority's main concern was that without some time limit on *Edwards*'s protections, the effect of a suspect's invocation of the right to counsel would be "eternal," and therefore an acute burden on the administration of justice.[108] Accordingly, the Court sought to place an objective, predictable limit on the applicability of *Edwards* by employing a cost-benefit analysis of the indefinite protection it provided.[109]

According to the Court, the primary benefit of *Edwards* was "measured by the number of coerced confessions it suppresse[d] that otherwise

100. *Id.*
101. *See id.*
102. *Id.*
103. *Id.* at 1218 & n.1 (discussing Maryland's filing of a *nolle prosequi* to the second-degree sexual offense charge, and consenting to dismissal of the misdemeanor charges barred by the statute of limitations).
104. *Id.* at 1218.
105. Shatzer v. State, 954 A.2d 1118, 1131 (Md. 2008), *rev'd*, 130 S. Ct. 1213 (2010).
106. *Shatzer*, 130 S. Ct. at 1217.
107. *Id.*
108. *See id.* at 1222.
109. See *id.* at 1220.

would have been admitted" at trial.[110] On the other hand, any voluntary confessions withheld from evidence constituted a cost paid by society.[111] The Court elaborated on the costs by arguing that because *Edwards* protections apply even where a subsequent interrogation concerns a different crime,[112] or is conducted by a different law enforcement agency,[113] a repeat offender may escape conviction because he remains protected after a single unrelated invocation of his rights.[114] The Court concluded that with no set limitations, the costs of the *Edwards* rule outweighed its benefits.[115]

In Shatzer's case, the Court found that the two-and-a-half year break in custody was sufficient to make his subsequent waiver voluntary.[116] However, the Court then questioned whether a period of one year or one week would have been sufficient.[117] The Court held that it would be impractical to leave these answers unresolved and established a clear-cut rule that a fourteen-day break in custody was sufficient to end the presumption of involuntariness established in *Edwards*.[118]

Although the fourteen-day limitation appeared with little explanation, the Court justified the rule in two ways. First, the Court reasoned that the need for the *Edwards* protections lessened where a suspect returned to "normal life."[119] A return to normal life, the Court noted, increased the likelihood that the suspect would have consulted with friends, family, or an attorney, and decreased the likelihood that a waiver was the result of badgering or coercion by law enforcement officers.[120] The Court found that two weeks was a sufficient amount of time to constitute a return to normal life.[121] Second, the Court asserted that a suspect would still be protected under *Johnson v. Zerbst*,[122] which mandated a totality of the circumstances inquiry into the voluntariness of a confession.[123] While the Court acknowledged that it was unusual for the Court to set precise limits governing police action, it asserted its prerogative to clarify its own legal mandate[124] and instituted its fourteen-day rule over the biting criticism of Justice Thomas[125] and Justice Stevens.[126]

---

110. *Id.*
111. *See id.*
112. *See* Arizona v. Roberson, 486 U.S. 675, 687–88 (1988).
113. *See* Minnick v. Mississippi, 498 U.S. 146, 153–54 (1990).
114. *Shatzer*, 130 S. Ct. at 1222.
115. *See id.*
116. *See id.*
117. *Id.*
118. *Id.* at 1223.
119. *Id.* at 1221.
120. *Id.*
121. *Id.* at 1223.
122. *Id.* at 1223 n.7.
123. Johnson v. Zerbst, 304 U.S. 458, 464 (1938).
124. *Shatzer*, 130 S. Ct. at 1220 ("We have frequently emphasized that the *Edwards* rule is not a constitutional mandate, but judicially prescribed prophylaxis.").
125. *See infra* Part II.D.

The Court also considered judicial and law enforcement efficiency in its determination.[127] The *Edwards* bright-line rule conserved judicial resources that would otherwise be dedicated to determining the voluntariness of a suspect's waiver.[128] By establishing another clear-cut rule, the *Shatzer* Court was able to maintain efficiency while restricting *Edwards*'s application.[129] The Court strengthened its holding by identifying those hardships on law enforcement that the fourteen-day rule would alleviate, increasing the admissibility of voluntary confessions.[130] Specifically, the Court reasoned that police investigations are more effective if officers "know, with certainty and beforehand, when renewed interrogation is lawful."[131]

Last, the Court addressed whether release back into the general prison population constituted a release from custody for *Edwards* and *Miranda* purposes.[132] Because prisoners retain some control over their lives, are often able to communicate with others, and the interrogator has no power over the incarceration, the Court answered the question in the affirmative.[133] Thus, the Court defined "normal" as merely returning to the state of life enjoyed by the suspect immediately before the interrogation. The Court held that as long as the suspect was not in "interrogative custody," meaning isolated with his accusers, release back into the general prison population constituted a break in custody for purposes of *Miranda* and *Edwards*.[134]

D. Justice Thomas's Concurring Opinion

Justice Thomas, concurring in part and in the judgment, criticized the majority's bright-line fourteen-day rule. Thomas immediately made clear his disagreement with any extension of the *Edwards* rule beyond the narrow facts of that case.[135] He then argued that even if *Edwards* applied in Shatzer's case, the majority's rule was arbitrary, incomplete, and inefficient.[136]

Furthermore, Justice Thomas maintained that the new fourteen-day rule was unnecessary because *Zerbst* mandated a totality of the circum-

126. *See infra* Part II.E.
127. *Shatzer*, 130 S. Ct. at 1220.
128. *Id.*
129. *Id.* at 1223–24 ("Now, in cases where there is an alleged break in custody, they simply have to repeat the inquiry for the time between the initial invocation and reinterrogation. In most cases that determination will be easy. And when it is determined that the defendant pleading *Edwards* has been out of custody for two weeks before the contested interrogation, the court is spared the fact-intensive inquiry into whether he ever, anywhere, asserted his *Miranda* right to counsel.").
130. *Id.* at 1223.
131. *Id.* at 1222–23.
132. *Id.* at 1224.
133. *Id.*
134. *Id.* at 1224–25.
135. *Id.* at 1227 (Thomas, J., concurring).
136. *See id.* at 1227–28.

stances test, which accounted for any time lapse, to determine the voluntariness of a waiver.[137] In addition, Justice Thomas disagreed with the majority's conclusion that the fourteen-day rule would aid police investigations.[138] Specifically, Justice Thomas stated, "Determining whether a suspect was previously in custody, and when the suspect was released, may be difficult without questioning the suspect, especially if state and federal authorities are conducting simultaneous investigations."[139] Last, Justice Thomas accused the majority of valuing certainty and ease of application over well-reasoned, substantive conclusions.[140]

E. Justice Stevens's Concurring Opinion

Justice Stevens, also concurring in the judgment, attacked the fourteen-day rule mainly on public policy concerns. He asserted that any bright-line rule was unsatisfactory because "[n]either a break in custody nor the passage of time ha[d] an inherent, curative power" to establish genuine voluntariness.[141] Justice Stevens argued that a suspect may assume that his requests for counsel have been ignored if he is reinterrogated after two weeks without having obtained counsel, and may assume he has no choice but to submit to the interrogation.[142] Moreover, Justice Stevens maintained that the police will be motivated "to delay formal proceedings, in order to gain additional information by way of interrogation after the time limit lapses."[143]

Justice Stevens also addressed the dangerous implications of the fourteen-day rule for suspects already in prison. First, Justice Stevens argued that because prisoners are summoned by guards to interrogation, they may assume that the guards and police are not independent, and feel forced to surrender to the questioning.[144] Next, Justice Stevens asserted that the fourteen-day rule could encourage officers or guards to badger imprisoned suspects, who will not have the opportunity to overcome the pressures from the interrogation.[145] Although Shatzer did not claim any disparate treatment by prison officials or guards between his two interrogations,[146] Justice Stevens was concerned with this "troubling set of incentives for police."[147] Last, because a suspect is already in custody, po-

137. *Id.* at 1227 n.1 (citing Miranda v. Arizona, 384 U.S. 436, 475 (1966)).
138. *See Shatzer*, 130 S. Ct. at 1228 n.2.
139. *Id.*
140. *Id.* at 1228.
141. *Id.* at 1234 (Stevens, J., concurring).
142. *Id.* at 1229.
143. *Id.* at 1231.
144. *Id.* at 1233 ("Prisoners are uniquely vulnerable to the officials who control every aspect of their lives; prison guards may not look kindly upon a prisoner who refuses to cooperate with police. And cooperation is frequently relevant to whether the prisoner can obtain parole.").
145. *See id.* at 1232 (asserting that a prisoner's freedom is "severely limited," making it unlikely that a suspect in prison has communicated with friends, family, or an attorney within fourteen days after questioning).
146. *Id.* at 1225 (majority opinion).
147. *Id.* at 1233 n.13 (Stevens, J., concurring).

lice have no need to formally place the suspect under arrest and can "comfortably bide their time, interrogating and reinterrogating their suspect" with little or no evidence of guilt, until the suspect surrenders and incriminates himself.[148]

III. ANALYSIS

The *Shatzer* fourteen-day rule confirms the Court's retreat from the prophylactic measures established in *Miranda* and *Edwards*. This retreat jeopardizes a suspect's rights to counsel and to remain silent. With *Shatzer*, the Court continued its gradual abandonment of *Miranda*'s protections by valuing efficiency above individual rights. Distressingly, this abandonment arrives at a time when Fifth and Sixth Amendment rights are increasingly more valuable to suspects. And while suspects are most in need of those rights, the Court's recent focus on *Edwards* makes it easier to curtail suspects' rights, making the *Shatzer* opinion more detrimental to suspects today. The Court's continued limitation of *Miranda* rights is logically unsound, wrongly focused, and inconsistent with the modern realities of criminal justice.

A. The Court's Retreat from the Prophylactic Ideals of Miranda *and* Edwards

The Court's recent retreat from prophylactic tenants overlooks the general concerns that guided the *Miranda* Court forty years ago. Relying on *Miranda*'s assertions, the *Shatzer* Court noted that a set of prophylactic measures was necessary to protect suspects from the "'inherently compelling pressures' of custodial interrogation."[149] "Inherently compelling" pressures denoted an inescapable characteristic of interrogation that generated psychological pressures and uncomfortable experiences for individuals under interrogation. Accordingly, the Court recognized some degree of psychological pressure present in *all* custodial interrogations.[150]

In *Edwards*, the Court recognized that these inherent pressures build with subsequent interrogations.[151] As such, *Edwards* held that waivers of the right to counsel occurring after a previous invocation of that right are presumed involuntary.[152] Concerned with genuine voluntariness, the *Edwards* Court likely declined to place a limit on the time between interrogations because every person will react to, and overcome, any coercive techniques differently and within varying timeframes. *Edwards* adhered to the *Miranda* Court's concerns about inherent pressures by holding that

148. *Id.*
149. *Id.* at 1219 (majority opinion) (quoting Miranda v. Arizona, 384 U.S. 436, 467 (1966)).
150. *See Shatzer*, 130 S. Ct. at 1219.
151. *See* Edwards v. Arizona, 451 U.S. 477, 483–84 (1981).
152. *See id.* at 484–85.

an inherent characteristic does not fade with time, and no generalizations about individual triumph over coercion would prove effective.

However, the *Shatzer* Court imported its own determination of the time it takes a suspect to overcome coercive effects of an interrogation and make a voluntary waiver: fourteen days.[153] That is all it takes to eliminate coercive psychological pressures, according to the Court.[154] The Court admits that pressures will still exist during subsequent interrogations, but assumes that the degree of pressure felt by a suspect after a two-week break in custody will never be more than the pressure felt at any prior custodial interrogations.[155] This assumption ignores the likely possibility that an individual will feel more pressured after a second, third, or tenth interrogation because he feels hunted and badgered by police.

Because pressure naturally builds in this way, the Court's estimation that pressure will only increase in "narrow circumstances" where no break in custody has occurred is flawed.[156] Though the Court contends that repeated interrogation attempts will increase the likelihood that a suspect will again assert his right,[157] it is more likely that a suspect will feel his requests have been ignored and he has no option but to talk.[158] Feeling that his rights have been ignored naturally increases pressure because the suspect will feel that he cannot trust his questioners. Logically, a break in custody will not always place a suspect in the same, or better, position than he was at the initial meeting.

Not only did the *Shatzer* Court discount and misapprehend the meaning of inherent pressures, it also erroneously failed to account for variances in individual personalities, experiences, and understandings of the criminal justice system. *Miranda* sought to provide "individuals the tools to counter inherently coercive pressure by asserting their right not to deal with the police alone."[159] However, although every suspect is given the same "tool" by being read the same rights, every individual has varying capacities to use this tool. For example, providing every American with a fishing pole does not mean that every American eats fish for dinner that night. Some Americans will have no clue what to do with the contraption, others will be scared of the sharp hook and live bait, and

153. *Shatzer*, 130 S. Ct. at 1223.

154. *Id.* at 1222–23.

155. *Id.* at 1223 ("It seems to us that period is 14 days. That provides plenty of time for the suspect to get reacclimated to his normal life, to consult with friends and counsel, and to shake off any residual coercive effects of his prior custody.").

156. *Id.* at 1226; *see id.* at 1231–32 (Stevens, J., concurring).

157. *Id.* at 1226 (majority opinion) (arguing that if a break in custody has not changed the suspect's mind about having counsel present, he will know from experience that he need only ask for counsel for the interrogation to cease).

158. *Id.* at 1229 (Stevens, J., concurring).

159. Strauss, *supra* note 65, at 815.

others still will be physically unable to maneuver the device because of age or disability.

A rule based on blanket generalizations is directly opposed to the core of *Miranda*'s analysis. The *Miranda* Court strongly asserted that "the privilege against self-incrimination applies to *all individuals*."[160] In fact, *Miranda* declared that the Fifth Amendment privilege is so fundamental that the defendant's "age, education, intelligence, or prior contact with authorities" should have no bearing on his ability to exercise his rights.[161] This reasoning accorded with the Court's earlier holding in *Zerbst* that "[t]he determination of whether there has been an intelligent waiver of right to counsel must depend, in each case, upon the particular facts and circumstances surrounding that case, including the background, experience, and conduct of the accused."[162]

Interestingly, the Court made the same mistake in *Davis* when it declared that a suspect's invocation of the right to counsel must be unambiguous.[163] Numerous scholars argue that the *Davis* rule will have a disproportionate effect on females and minorities.[164] Specifically, women and minorities are far more likely to use indirect speech patterns such as "maybe" and "I think."[165] Therefore, by failing to account for linguistic variances in certain segments of the population, *Davis* arbitrarily denied some individuals the right to counsel.[166]

Despite precedent that acknowledged and protected individual abilities, the Court ignored this principle in *Davis* and *Shatzer*. With its fourteen-day rule, the *Shatzer* Court took *Miranda*'s and *Edwards*' concern with genuine, individual voluntariness and replaced it with a blanket generalization about human reaction to subsequent or repeated interrogation techniques. The *Shatzer* Court expressed this generalization as the suspect returning to "normal life."[167] However, the emphasis on a return to normalcy is troublesome because even if a suspect is placed back at equilibrium, *inherent* pressures will still revisit him during subsequent interrogation. And if he felt unwilling or unable to communicate to his interrogators without counsel the first time, the return of these pressures will probably restore, or even enhance, that feeling.

---

160. Miranda v. Arizona, 384 U.S. 436, 472 (1966) (emphasis added).
161. *Id.* at 468–69.
162. Johnson v. Zerbst, 304 U.S. 458, 464 (1938).
163. Davis v. United States, 512 U.S. 452, 459 (1994).
164. Strauss, *supra* note 34, at 1030 (citing David Aram Kaiser & Paul Lufkin, *Deconstructing* Davis v. United States*: Intention and Meaning in Ambiguous Requests for Counsel*, 32 HASTINGS CONST. L.Q. 737, 759 n.69 (2005) ("[T]he actual linguistic practices of many women and minorities preclude them from meeting the standard of clarity demanded by *Davis*.")).
165. Strauss, *supra* note 34, at 1030–31.
166. Strauss, *supra* note 34, at 1031.
167. Maryland v. Shatzer, 130 S. Ct. 1213, 1221 (2010) (majority opinion).

The Court's retreat from prophylactic measures compromises not only a suspect's right to counsel, but also his right to remain silent. "The right to counsel exists only to protect the right to remain silent," as the *Miranda* Court considered the former as a means to protect the latter.[168] Consequently, any limitations on the right to counsel also limit the right to remain silent. *Shatzer*'s limitation of the right to counsel may also spur nationwide decisions similarly limiting the right to remain silent. Although courts have largely considered the two rights as separate and distinct standards,[169] nine out of eleven circuits and the District of Columbia have applied the *Davis* standard for invoking the right to counsel to the right to remain silent.[170] Similarly, *Shatzer*'s limitation of the right to counsel may be applied in cases concerning the right to remain silent. Admittedly, because the right to remain silent is already quite limited,[171] *Shatzer*'s fourteen-day rule would actually bolster that right. But the underlying trend of *Shatzer* and *Davis*—limiting the prophylactic protections awarded by *Miranda* and *Edwards*—is a dangerous ideal to transport into cases involving the right to remain silent. Courts may use *Shatzer*'s fourteen-day rule to proportionally limit the time lapse required to spoil a suspect's invocation of the right to remain silent. And any additional limitations on the right to remain silent may evaporate the right entirely.

B. *The* Shatzer *Court Wrongly Abandoned Prophylactic Measures in Favor of Efficiency*

The *Shatzer* Court held that it would be "impractical" to leave *Edwards*'s application open for clarification on a case-by-case basis, partly for judicial efficiency[172] and partly to ensure that law enforcement officers know for certain when renewed interrogation is lawful.[173] Though the importance of judicial efficiency is debatable, it is beyond the scope of this Comment.[174] Nevertheless, the Court placed too much emphasis

168. Strauss, *supra* note 65, at 817.

169. *See supra* note 65 and accompanying text.

170. Strauss, *supra* note 65, at 784–85 (citing Valle v. Sec'y for the Dep't of Corr., 459 F.3d 1206, 1213–15 (11th Cir. 2006); United States v. Nelson, 450 F.3d 1201, 1211–12 (10th Cir. 2006); United States v. Sherrod, 445 F.3d 980, 982 (7th Cir. 2006); McGraw v. Holland, 257 F.3d 513, 519 (6th Cir. 2001); Simmons v. Bowersox, 235 F.3d 1124, 1131 (8th Cir. 2001); United States v. Anderson, No. 95-3048, 1996 WL 135720 (D.C. Cir. Feb. 16, 1996) (per curiam)).

171. *See supra* note 65, and accompanying text.

172. *Shatzer*, 130 S. Ct. at 1222–24 (finding that the fourteen-day rule would conserve judicial resources by making the determination of voluntariness "easy" if a suspect has been out of custody for two weeks or longer).

173. *See id.* at 1222–23.

174. The Court has often expressed its preference for bright-line rules over totality of the circumstances approaches. *See, e.g.*, Davis v. United States, 512 U.S. 452, 461 (1994) ("[I]f we were to require questioning to cease if a suspect makes a statement that *might* be a request for an attorney, this clarity and ease of application [set forth in *Edwards*] would be lost."); Minnick v. Mississippi, 498 U.S. 146, 151 (1990) ("The merit of the *Edwards* decision lies in the clarity of its command and the certainty of its application.").

on ensuring ease of application for law enforcement, and abandoned its dedication to prophylactic measures established in *Miranda*.

After the Court barely kept the core of *Miranda*'s holding alive in *Dickerson*, it once again turned its back on one of *Miranda*'s main principles: that law enforcement's investigative powers, though valuable to society, are limited by the rights of the accused guaranteed in the Sixth Amendment.[175] Individual rights should not and do not have to be compromised to establish an effective system of law enforcement.[176] The *Miranda* Court acknowledged the importance of police investigations and interrogations but refused to abridge constitutional rights to make the prosecutor's job a little easier.[177] With *Shatzer*, the Court tipped the scales in the opposite direction based on a flawed focus on efficiency.

Contrary to the Court's assertion, the fourteen-day rule is not necessary to ensure the fair and effective administration of justice. The original *Edwards* rule did not prevent all confessions. If a suspect wishes to make a voluntary confession, he may do so even after invoking his rights, as *Edwards* allows questioning to resume if a suspect initiates the discussion.[178] And because only twenty percent of suspects invoke their rights,[179] *Edwards* ultimately has no effect on a vast majority of cases, and the number of confessions that may be suppressed is slim. Moreover, the Court's fixation with law enforcement is unfounded and contrary to a fair criminal justice system. In a just system, law enforcement should not "have to fear that if an accused is permitted to consult with a lawyer, he will become aware of, and exercise, these rights."[180] The mere fact that law enforcement fears a suspect's exercise of his fundamental rights is unsettling.

In fact, those fears may not be legitimate. Despite the Court's intentions in *Miranda*, false confessions are still prevalent,[181] showing that the numerous exceptions to *Miranda* have provided law enforcement with sufficient loopholes to continue to practice coercive tactics during interrogations. Although it is now well established that physical abuse is an illegal tactic to extract confessions, the line between acceptable psychological techniques and psychological coercion that is a violation of the Constitution remains blurred.[182] The fourteen-day rule, along with the

175. Holland, *supra* note 8, at 390.

176. Strauss, *supra* note 65, at 773 (quoting Escobedo v. Illinois, 378 U.S. 478, 490 (1964) ("If the exercise of constitutional rights will thwart the effectiveness of a system of law enforcement, then there is something very wrong with that system.")).

177. Miranda v. Arizona, 384 U.S. 436, 479–81 (1966).

178. Edwards v. Arizona, 451 U.S. 477, 484–85 (1981).

179. Strauss, *supra* note 65, at 774.

180. Escobedo v. Illinois, 378 U.S. 478, 490 (1964).

181. Brandon L. Garrett, *The Substance of False Confessions*, 62 STAN. L. REV. 1051, 1060 (2010).

182. *See* Laura Hoffman Roppe, *True Blue? Whether Police Should Be Allowed to Use Trickery and Deception to Extract Confessions*, 31 SAN DIEGO L. REV. 729, 732 (1994).

Court's progeny of exceptions to *Miranda*, has informed law enforcement officers of exactly how much questionable behavior they can legally employ.

Specifically, *Shatzer*'s rule will actually aid law enforcement in extracting confessions from suspects in custody because those suspects will be susceptible to incessant re-questioning every two weeks, even if they properly assert their right to counsel. Those people unable to make bail for whatever reason will therefore be more detrimentally impacted by the fourteen-day rule than will the rest of society. Hence, the rule will be arbitrarily more harmful to certain individuals with no justification.

Ultimately, by continuing to institute exceptions and limitations to the application of *Miranda* rights, the Court is instituting a dangerous pattern that actually helps police engage in trickery and coercion. Nonetheless, the *Shatzer* decision continued the Court's precedent of chipping away at *Miranda* rights to satisfy concerns about effective law enforcement. Moreover, as law enforcement officers become more clever and confident in their techniques, suspects struggle to assert and protect their fundamental rights, which are particularly critical in today's criminal prosecutions.

C. Why Fifth and Sixth Amendment Rights are More Important to Suspects Today

The *Shatzer* fourteen-day rule limited a suspect's rights at a time when a suspect's pretrial rights are becoming increasingly critical to the outcome of his case. Approximately ninety percent or more of today's criminal trials are resolved by negotiated disposition rather than trial, meaning defendants "rarely face their accusers during traditional courtroom proceedings that pit skilled trial lawyers against each other."[183] This is a recent development in criminal law that differs from the reality the Court faced at the time of *Miranda*. Specifically, between 1980 and 2002, the rate of federal criminal cases concluded by a bench and jury trial fell from 23 percent to 4.8 percent.[184] So today, pretrial contexts, such as interrogation settings, are the stage for judgment, where damage can be minimized, bargains can be struck, and cases can be won or lost.[185] In fact, only in rare cases does the "compulsion" sought to be protected by the Fifth Amendment occur at trial.[186]

This modern reality makes a suspect's right against self-incrimination incredibly valuable, as there may not be a trial to argue the

183. Holland, *supra* note 8, at 382.

184. Frank O. Bowman, III, Response, *American Buffalo: Vanishing Acquittals and the Gradual Extinction of the Federal Criminal Trial Lawyer*, 156 U. PA. L. REV. PENNUMBRA 226, 226 (2007), http://www.pennumbra.com/responses/11-2007/Bowman.pdf.

185. Holland, *supra* note 8, at 382–83.

186. Manheimer, *supra* note 12, at 1265.

voluntariness of a statement. Even more troubling for suspects, the rate of acquittal has declined alongside the falling rate of criminal trials.[187] The declining rate of acquittals has been attributed to the enactment of the United States Sentencing Guidelines, which provided prosecutors more bargaining leverage.[188] Thus, suspects are at a disadvantage from the initiation of the investigation because the interrogation context is increasingly more influential to the result of their case, *and* prosecutors have increased bargaining power. And at this time when pretrial contexts are especially valuable to suspects, the Court is continuing to curtail the pretrial rights of the accused.

The trend of modern criminal prosecution also clouds the line separating Fifth and Sixth Amendment protections, which further increases a suspect's need for prophylactic protections that transcend the bare text of the Constitution. Unlike the Fifth Amendment, the Sixth Amendment textually guarantees a suspect's right to counsel in a "criminal prosecution."[189] This right need not be invoked, but automatically takes effect when prosecution commences.[190] However, this right does not attach until the "critical stage" of the proceedings, which can include post-charge interrogations and lineups.[191] Today, as initial and pre-charge interrogations grow increasingly influential in criminal prosecutions, the definition of this "critical stage" is changing. While the increasingly blurred line between the critical and non-critical stage of criminal prosecution would support stronger rights earlier in the process, the Court has done the opposite. Ignoring the realities of modern criminal prosecution, the Court has made the right to counsel harder to invoke[192] and more difficult to maintain.[193]

D. How the Court's Focus on Edwards *Makes it Easier for the Court to Curtail Suspect Rights*

The *Shatzer* Court confidently flexed its muscles by declaring its prerogative to alter its own "judicially prescribed prophylaxis."[194] The *Shatzer* Court justified its drastic limitation on individual rights by pro-

187. Bowman, *supra* note 184, at 227.

188. Bowman, *supra* note 184, at 226–27 (citing Ronald F. Wright, *Trial Distortion and the End of Innocence in Federal Criminal Justice*, 154 U. PA. L. REV. 79, 101–06 (2005) (presenting data to support the assertion that the post-1987 federal sentencing system consisting of the United States Sentencing Guidelines provided prosecutors more bargaining leverage, directly resulting in the declining number of acquittals)).

189. U.S. CONST. amend. VI.

190. Holland, *supra* note 8, at 390.

191. *Id.* (citing Moran v. Burbine, 475 U.S. 412, 428 (1986)).

192. Davis v. United States, 512 U.S. 452, 459 (1994) (holding that a suspect's invocation of the right to counsel must be unambiguous to halt the interrogation).

193. Maryland v. Shatzer, 130 S. Ct. 1213, 1223 (2010) (holding that a suspect's assertion of his right to counsel forbids police from interrogating the suspect again for fourteen days if he has not obtained counsel).

194. *Id.* at 1220.

claiming that the *Edwards* rule is not a constitutional mandate.[195] But only a decade ago in *Dickerson*, the Court avoided overruling *Miranda* by declaring unconstitutional an act of Congress that purported to reduce *Miranda* warnings to a mere factor for consideration in determining the voluntariness of a statement.[196] Critical to the *Dickerson* Court's reasoning was that *Miranda* was a "constitutional decision" of the Court, which may not be overruled by an act of Congress.[197] Although the *Dickerson* Court qualified its decision by explaining that constitutional rulings are not immutable, but are subject to judicial modification,[198] the Court explicitly classified *Miranda* as a "constitutional decision."[199] Only ten years after this controversial classification, the *Shatzer* Court declared that *Edwards*, a direct offspring of *Miranda*,[200] is not a constitutional mandate.[201]

Because *Edwards* and *Miranda* are so intimately related, this shift is not fully justified. The *Edwards* opinion simply reconfirmed[202] the *Miranda* mandate that an accused has a constitutional right to have counsel present during custodial interrogation.[203] The *Edwards* Court aimed to provide "substance" to *Miranda* and its progeny, and emphasized that it is inconsistent with *Miranda* for police to reinterrogate a suspect in custody after he has clearly asserted his right to counsel.[204] Therefore, the *Edwards* rule was fashioned completely on *Miranda*'s heels, and if one rule is a constitutional mandate, the other should be as well.[205]

Characterizing *Edwards* as merely a judicially created prophylaxis increases and assists the Court's ability to further curtail suspect rights. By switching its focus from *Miranda* to *Edwards*, the *Shatzer* Court has found an easier way to limit rights of the accused. Specifically, limiting the *Edwards* rule is easier than limiting the *Miranda* rules because by classifying the *Edwards* rule as a judicial prophylaxis instead of a constitutional rule, the Court need not defend its limitations on that prophylaxis, as it did with *Miranda* in *Dickerson*.[206]

---

195. *Id.*
196. Dickerson v. United States, 530 U.S. 428, 432 (2000).
197. *Id.*
198. *Id.* at 441.
199. *Id.* at 432.
200. *Shatzer*, 130 S. Ct. at 1219–20 (describing the advent of the *Edwards* rule as an expansion of *Miranda* rights).
201. *Id.* at 1220.
202. Edwards v. Arizona, 451 U.S. 477, 485 (1981).
203. *Id.* at 482.
204. *Id.* at 485.
205. *See Shatzer*, 130 S. Ct. at 1228 (Stevens, J., concurring) ("The source of the holdings in the long line of cases that includes both *Edwards* and *Miranda*, however, is the Fifth Amendment's protection against compelled self-incrimination applied to the 'compulsion inherent in custodial' interrogation . . . .").
206. Dickerson v. United States, 530 U.S. 428, 439–44 (2000).

CONCLUSION

The United States Supreme Court began running with the idea of additional safeguards for the accused in *Miranda* and *Edwards*, but tripped over those safeguards with *Davis* and eventually fell backwards with *Shatzer*. In *Shatzer*, the Court correctly asserted its prerogative to clarify and constrain its own prophylactic creations, but lost sight of its concurrent responsibility to protect individual rights, albeit those the Court itself has created. By focusing on the *Edwards* rule instead of *Miranda* rights generally, the Court was able to create the fourteen-day rule with minimal constitutional challenge, although the changing process of criminal justice makes *Edwards*' prophylactic measures more closely related to explicit constitutional rights.

Consequently, the Court ignored the realities of the modern criminal prosecution process and drastically limited *Miranda* rights at a time when suspects need them the most. The Court, and Justice Thomas in his concurrence, ask their audience to find solace in *Zerbst* protections still available to defendants.[207] But because *Zerbst* predates *Miranda* by twenty-eight years, it can be argued that *Miranda* replaced the need for *Zerbst*, meaning that *Shatzer* was the Court's last opportunity to salvage *Miranda* rights. Given this opportunity, the Court not only constricted the accused's right to counsel, but also jeopardized his right to remain silent.

Given the Court's flawed reasoning, detrimental impact, and inconsistency with the realities of modern criminal prosecutions, *Shatzer* was wrongly decided and will hinder the fair administration of criminal justice in America.

Hannah Misner[*]

207. *Shatzer*, 130 S. Ct. at 1227 n.1 (Thomas, J., concurring).

* J.D. Candidate, 2012. I would like to thank the members of the Denver University Law Review Board and editorial staff for their significant and untiring roles in refining this Comment. I would also like to thank Karen Steinhauser for her time and invaluable contributions. Lastly, I thank my family and friends for their continuing support throughout all of my incredible journeys.