NEW YORK UNIVERSITY ANNUAL SURVEY OF AMERICAN LAW

VOLUME 67
ISSUE 4

NEW YORK UNIVERSITY SCHOOL OF LAW
ARTHUR T. VANDERBILT HALL
Washington Square
New York City

New York University Annual Survey of American Law
is in its seventieth year of publication.

L.C. Cat. Card No.: 46-30523
ISSN 0066-4413

New York University Annual Survey of American Law is published quarterly at 110 West 3rd Street, New York, New York 10012. Subscription price: $30.00 per year (plus $4.00 for foreign mailing). Single issues are available at $16.00 per issue (plus $1.00 for foreign mailing). For regular subscriptions or single issues, contact the *Annual Survey* editorial office. Back issues may be ordered directly from William S. Hein & Co., Inc., by mail (1285 Main St., Buffalo, NY 14209-1987), phone (800-828-7571), fax (716-883-8100), or email (order@wshein.com). Back issues are also available in PDF format through HeinOnline (http://heinonline.org).

Manuscripts: The *Annual Survey* invites the submission of unsolicited manuscripts. Text and citations should conform to the 19th edition of *A Uniform System of Citation.* Please enclose an envelope with return postage if you would like your manuscript returned after consideration.

Editorial Office: 110 West 3rd Street, New York, N.Y. 10012
(212) 998-6540
(212) 995-4032 Fax
http: //www.law.nyu.edu/pubs/annualsurvey

For what avail the plough or sail
Or land or life, if freedom fail?
EMERSON

NEW YORK UNIVERSITY ANNUAL SURVEY OF AMERICAN LAW

2011-2012 BOARD OF EDITORS

SUMMARY OF CONTENTS

WHEN THE SHOP FLOOR IS IN THE LIVING ROOM: TOWARD A DOMESTIC EMPLOYMENT RELATIONSHIP THEORY

*ELIZABETH J. KENNEDY, J.D.**

INTRODUCTION

In the summer of 2010, nannies, housekeepers, and elder caregivers celebrated the passage of the New York Domestic Workers' Bill of Rights, a landmark victory in a national campaign to reverse the exclusion of domestic workers from the protections of state and federal labor and employment laws.[1] This spring, the International Labor Organization (ILO), of which the United States is a member, voted to adopt a set of international standards that

* Assistant Professor, Law & Social Responsibility Department, Loyola University Maryland.

1. Domestic Workers' Bill of Rights, ch. 481, 2010 N.Y. Sess. Laws 1315 (McKinney). On August 31, 2010, Governor David A. Paterson signed this bill into law. *Id.* The law went into effect on November 29, 2010. *Id.*

would extend to domestic workers the fundamental rights guaranteed to all workers by ILO nations.[2] While these are important steps for a nearly invisible, yet indispensable workforce subject to pervasive abuse that has struggled, at home and abroad, to have its labor recognized as "real work," questions remain about the relationship between domestic workers and their employers and how that relationship should be regulated.[3] Should the work "that makes all other work possible"[4] be treated identically to all other work? Or should the particular realities of domestic employment shape workplace standards, the enforcement of those standards, and the contours of collective bargaining? Other questions include whether or not domestic employers should be compelled to reinstate a nanny following retaliatory discharge, the standard remedy in other industries. The differences between domestic work and all other work might make it inappropriate for a court to require reinstatement of an aggrieved caretaker or housecleaner into a private home. Likewise, given the logistical (and potential Fourth Amendment) obstacles to performing worksite inspections in private homes, new strategies must be developed to ensure compliance with, and vigorous enforcement of, workplace regulations. Ultimately, a comprehensive assessment of the differences and similarities between domestic work and all other work is necessary in order to improve

2. Akito Yoshikane, *New ILO Convention Gives Domestic Workers Historic Labor Rights*, IN THESE TIMES (June 22 2011, 10:55AM), http://inthesetimes.com/working/entry/11549/new_ilo_convention_gives_domestic_workers_historic_labor_rights/.

3. DOMESTIC WORKERS UNITED & DATACENTER, HOME IS WHERE THE WORK IS: INSIDE NEW YORK'S DOMESTIC WORK INDUSTRY 31 (2006) [hereinafter HOME IS WHERE THE WORK IS], *available at* http://www.datacenter.org/reports/homeiswheretheworkis.pdf; *see also* PIERRETTE HONDAGNEU-SOTELO, DOMÉSTICA: IMMIGRANT WORKERS CLEANING AND CARING IN THE SHADOWS OF AFFLUENCE 9–12 (2d ed. 2007) (describing the unique features of caring work). For a detailed report on the pervasive abuse experienced by domestic workers in the United States and arguments about why permitting this abuse violates Articles 2, 3, 7, 8, 9, 12, 17, 19, 21, 22, and 26 of the International Covenant on Civil and Political Rights, see STEFANI BONATO ET AL., DOMESTIC WORKERS' RIGHTS IN THE UNITED STATES: A REPORT PREPARED FOR THE U.N. HUMAN RIGHTS COMMITTEE IN RESPONSE TO THE SECOND AND THIRD PERIODIC REPORT OF THE UNITED STATES (2006), *available at* http://www.law.unc.edu/documents/clinicalprograms/domesticworkersreport.pdf.

4. "Respect the work that makes all other work possible," is one of the central organizing themes of Domestic Workers United, a central domestic worker advocacy organization. CLAIRE HOBDEN, WINNING FAIR LABOUR STANDARDS FOR DOMESTIC WORKERS: LESSONS LEARNED FROM THE CAMPAIGN FOR A DOMESTIC WORKER BILL OF RIGHTS IN NEW YORK STATE 20 (2010), *available at* http://www.ilo.org/wcmsp5/groups/public/—ed_dialogue/—actrav/documents/publication/wcms_149488.pdf.

working conditions, transform the relationship between domestic workers and employers, and challenge industry-wide structural inequality.

Various scholarly models of the employment relationship[5]—which could help answer these questions—do not address many of the issues unique to domestic work, such as the isolation of domestic workers, the intimate nature of the home as workplace, and the complicated emotional bonds that develop between domestic workers and their employers (or their employers' children or elderly parents). Nor do they explain the unprecedented movement by some domestic *employers* calling for stronger regulation of the domestic industry. United by a vision that values the dignity of domestic work, these employers do not fit neatly into standard industrial relations paradigms.[6]

This article seeks to reconcile employment relationship theory with the unique characteristics of the domestic work industry. Just as traditional employment relationship models have shaped the regulation of other industries, a new model for domestic work could help regulators, workers, advocacy groups, and employers create and enforce standards and transform the employment relationships upon which most domestic work is based.

Part I of this Article provides an overview of the domestic work industry, including recent legislative and organizing campaigns led by domestic workers and employers; Part II identifies the limitations of using the predominant employment relationship theories to understand the domestic work industry; Part III proposes a new theoretical framework with which to analyze, regulate, and improve the domestic employment relationship; and Part V considers the New York Domestic Workers' Bill of Rights in light of the proposed "Domestic Employment" theoretical model.

5. *See generally* THEORETICAL PERSPECTIVES ON WORK AND THE EMPLOYMENT RELATIONSHIP (Bruce Kaufman ed., 2004) (collecting essays on employment relationship models).

6. *See generally* AI-JEN POO, DOMESTIC WORKERS UNITED, ORGANIZING WITH LOVE: LESSONS FROM THE NEW YORK DOMESTIC WORKERS BILL OF RIGHTS CAMPAIGN (2011), *available at* http://www.domesticworkersunited.org/media/files/287/OrganizingWithLoveFinal.pdf (charting the work of various non-profits that led the passage of the New York legislation).

I.
THE DOMESTIC WORK INDUSTRY

A. Industry Profile

The domestic work industry is growing, fueled by changes in local and global economies and population demographics.[7] As all wage earners work more hours each week, and as women's participation in the paid workforce continues to expand, domestic workers, such as nannies, housekeepers, and elder caregivers, play a critical role in making that work possible.[8] As noted by socio-legal scholars, domestic workers support their employers' participation in the global economy while perpetuating the "illusion of a family in which two wage earners do[ ] it all."[9]

---

7. Angela Charlton, *Study: Informal Employment at Record Levels*, ASSOCIATED PRESS FINANCIAL WIRE, Apr. 8, 2009, *available at* LEXIS, News Library, APFINL File; *see also* ANNETTE BERNHARDT, SIOBHÁN MCGRATH, & JAMES DEFILIPPIS, BRENNAN CTR. FOR JUST., UNREGULATED WORK IN THE GLOBAL CITY: EMPLOYMENT AND LABOR LAW VIOLATIONS IN NEW YORK CITY 61–62 (2007), *available at* http://www.brennancenter.org/dynamic/subpages/download_file_49436.pdf; Donna E. Young, *Working Across Borders: Global Restructuring and Women's Work*, 2001 UTAH L. REV. 1, 8 (2001).

8. *Domestic Working Circumstances and Conditions: Hearing Before the Assembly Standing Comm. on Labor, Children & Families and Social Services*, 2008 Leg., 231st Sess. (N.Y. 2008) [hereinafter *Domestic Working Hearing*] (testimony of Susan Wefald, Executive Vice President and Chief Operating Officer, Ms. Foundation for Women). In 2006, the following statistics described working mothers:

> Sixty-four percent of women with children under age 6, and 56 percent of women with infants (under age 1)— now work outside the home. A full seventy-seven percent of women with children age 6–17, and eighty-one percent of single women with children that age, are in the labor force. Seventy-six percent of employed mothers of children under eighteen work full time.

Id. (citing BUREAU OF LABOR STATISTICS, U.S. DEP'T OF LABOR, CURRENT POPULATION SURVEY, EMPLOYMENT CHARACTERISTICS OF FAMILIES IN 2006, tbls. 5 & 6, *available at* http://www.bls.gov/news.release/archives/famee_05092007.pdf). The increasing use of domestic workers may explain the overwhelming success of Kathryn Stockett's novel, THE HELP, as well as the *New York Time*s' placement of Barbara Ehrenreich's profile of Ai-jen Poo in the *Style Magazine*. Barbara Ehrenreich, *The Nannies' Norma Rae: Ai-jen Poo Fights for Domestic Workers' Rights*, THE N.Y. TIMES STYLE MAG. (Apr. 26, 2011, 9:00AM), http://tmagazine.blogs.nytimes.com/2011/04/26/the-nannies-norma-rae/; Marjorie Kehe, *With Book Sales Still Strong, 'The Help' Will Begin Filming*, CHRISTIAN SCIENCE MONITOR CHAPTER & VERSE BLOG (May 14, 2010), http://www.csmonitor.com/Books/chapter-and-verse/2010/0514/With-book-sales-still-strong-The-Help-will-begin-filming.

9. Teresa Carrillo, *The Best of Care: Latinas as Transnational Mothers and Caregivers*, *in* TECHNOFUTUROS: CRITICAL INTERVENTIONS IN LATINA/O STUDIES 191, 193 (Nancy Raquel Mirabal & Agustin Laó-Montes eds., 2007); *cf.* Kristi L. Graunke, *"Just Like One of the Family": Domestic Violence Paradigms and Combating On-the-Job Violence Against Household Workers in the United States*, 9 MICH. J. GENDER & L.

Although the exact number of domestic workers in the U.S. is unclear, the U.S. Census Bureau estimated that there were 1.3 million housekeepers and nannies nationwide in 2008,[10] a figure expected to increase at a rate of 11% between 2008 and 2018.[11] Further, as more older Americans (the first wave of "baby boomers" recently turned sixty-five) look to "aging in place" in their own homes,[12] many will require assistance with daily activities such as dressing, bathing, toileting, housekeeping, shopping, and visiting

131, 165 (2002) (explaining that domestic workers are "often treated as [a] more menial part of their employers" rather than as individuals). Though demographic information regarding domestic employers is scarce, most likely a majority or significant minority are dual income households with working mothers. NEW YORK STATE DEP'T OF LABOR, FEASIBILITY OF DOMESTIC WORKER COLLECTIVE BARGAINING 8 (2010) [hereinafter FEASIBILITY STUDY], *available at* http://www.labor.ny.gov/sites/legal/laws/pdf_word_docs/domestic-workers/domestic-workers-feasibility-study.pdf. Unpaid or deeply discounted domestic work has been the cushion on which the entire economy rests. *GRITtv with Laura Flanders: Domestic Worker Roundtable* (GRITtv television broadcast Nov. 1, 2011) [hereinafter *Laura Flanders*], *available at* http://blip.tv/file/984863/?utm_source=blip&utm_medium=site_search&utm_content=blip&utm_campaign=s_ab (last visited July 12, 2011).

10. Nannies hired to care for children often perform housekeeping work as well, "two jobs for the price of one." HONDAGNEU-SOTELO, *supra* note 3, at 5.

11. BUREAU OF LABOR STATISTICS, U.S. DEP'T OF LABOR, OCCUPATIONAL OUTLOOK HANDBOOK, 2010–11 EDITION, CHILD CARE WORKERS 3 (2010), *available at* http://www.bls.gov/oco/pdf/ocos170.pdf. Precise information about the size of the domestic work industry (or industries) is unavailable, in part because many domestic employment relationships are not reported to government entities. In a survey of 800 domestic employers in lower Manhattan and Brooklyn, 77% of respondents reported paying their nannies "off the books," while 14% reported paying "part on/part off" the books, and 9% said they pay on the books. PARK SLOPE PARENTS, THE PARK SLOPES PARENTS NANNY COMPENSATION SURVEY 2010 15 (2010) [hereinafter 2010 NANNY SURVEY], *available at* www.parkslopeparents.com/docs/NannySurvey2010.FINAL.pdf; *see also* FEASIBILITY STUDY, *supra* note 9, at 7 (explaining that advocates' estimates of the number of domestic workers in New York vary anywhere from 200,000 to 450,000). Domestic Workers United estimates the number of U.S. domestic workers to be 2.5 million. *See* Yoshikane, *supra* note 2.

12. NICHOLAS FARBER ET AL., NATIONAL CONFERENCE OF STATE LEGISLATURES & AARP PUBLIC POLICY INSTITUTE, AGING IN PLACE: A STATE SURVEY OF LIVABILITY POLICIES AND PRACTICES 1 (2011), *available at* http://assets.aarp.org/rgcenter/ppi/livcom/aging-in-place-2011-full.pdf.

doctors.[13] For most, relying on family members, neighbors, or other informal strategies will be insufficient to meet those needs.[14]

Domestic work can take various forms. "Live-in" domestic workers reside in their employers' homes and often perform more than one set of domestic tasks, such as cooking, cleaning, and childcare.[15] The majority of domestic workers, however, are likely hired exclusively for full-time child or elder care, or for a combination of cleaning, childcare, and elder care.[16] As a distinct group, housecleaners are typically hired for cleaning on a daily or weekly basis, often piecing together several jobs each week.[17] Au pairs and victims of trafficking are two other subgroups of domestic workers, each occupying an extreme position on either end of a continuum of legal protections and support.[18]

13. *See* Catherine Ruckelshaus & Ai-jen Poo, *When Companionship Doesn't Pay*, THE HILL'S CONGRESS BLOG (July 13, 2011, 12:01PM), http://thehill.com/blogs/congress-blog/labor/171207-when-companionship-doesnt-pay (predicting that "[o]ver the next two decades, the U.S. population over the age of 65 will grow to more than 70 million" and that the higher life expectancy means that "the demand for caregiving is expected to grow significantly"). An estimated 27 million people will need direct care by the year 2050. *Id.*

14. Peggie R. Smith, *Aging and Caring in the Home: Regulating Paid Domesticity in the Twenty-First Century*, 92 IOWA L. REV. 1835, 1844–45 (2007). A recent survey found that 71 % of domestic employers agreed with the statement "[o]ur nanny is my employee whose work I greatly respect," and 43% felt that "[o]ur nanny is like part of our family." PARK SLOPE PARENTS, THE PARK SLOPE PARENTS NANNY COMPENSATION SURVEY 2008 48, 50 (2008) [hereinafter 2008 NANNY SURVEY], *available at* http://test.parkslopeparents.com/images/Nanny%20Survey%20FINAL.pdf. Fifty-one percent reported that "[t]here are many times that my nanny has more patience than I have with my children," and 39% agreed that "[their] nanny has taught [them] ways to be a better parent." *Id.* at 49–50. Thirty-seven percent believed that their family was stronger because of their nanny. *Id.* at 50.

15. BERNHARDT, MCGRATH & DEFILIPPIS, *supra* note 7, at 61. In a survey of 1,100 domestic employers (parents employing nannies to care for their child(ren)) in the New York City area, 97% employed "live-out" nannies, while 3% hired "live-in" nannies. 2008 NANNY SURVEY, *supra* 14, at 7.

16. *See* BERNHARDT, MCGRATH, & DEFILIPPIS, *supra* note 7, at 61 (explaining that while accurate numbers for each segment of the domestic work population are unavailable, the demand for "live-out" domestic workers is growing). In addition to performing basic child-care duties, 82% of full-time nannies performed "light housekeeping," 64% did the child(ren)'s laundry, 39% shopped for kid-related supplies (e.g., milk, diapers), and 10% performed "heavy housecleaning." 2008 NANNY SURVEY, *supra* note 14, at 28. The percentage in each category was slightly less for part-time nannies. *Id.* Domestic workers also fill the unsung roles of "nurses, art teachers, counselors, tutors, assistants, and nutritionists." POO, *supra* note 6, at 2.

17. HONDAGNEU-SOTELO, *supra* note 3, at 70–71.

18. *See* BERNHARDT, MCGRATH, & DEFILIPPIS, *supra* note 7, at 61 (explaining that while au pairs are "generally brought into the country with visas and are hired

Domestic workers are disproportionately low-income women of color, whose earnings place them near the bottom of the economic ladder.[19] Domestic work's association with voluntary caring roles and "women's work" has had tangible economic consequences: it is likely that half of all domestic workers' earnings are close to or below the poverty line.[20] The demographic profile of domestic work-

exclusively for child care," victims of trafficking are "brought into the country . . . by professional traffickers . . . and live under conditions of servitude and imprisonment"). Only recently has the U.S. Department of Labor expressed a willingness to use its authority to help victims of trafficking and egregious wage and hour violations, including domestic workers, apply for special "U Visas" that would provide a pathway to permanent residency or citizenship. News Release, U.S. Dep't of Labor, US Labor Department to Exercise Authority to Certify Applications for U Visas (Mar. 15, 2010), *available at* http://www.dol.gov/opa/media/press/opa/opa20100312.htm. Au pairs, though not entirely protected from the workplace abuses experienced by many other domestic workers, do have a number of structural supports, and the congressionally sponsored program largely recruits middle-class women from Europe for "educational and cultural exchange" on J-1 visas. Joy M. Zarembka, *Migrant Maids and Modern Day Slavery*, *in* GLOBAL WOMAN: NANNIES, MAIDS AND SEX WORKERS IN THE GLOBAL ECONOMY 142, 149 (B. Ehrenreich & A. Hochschild eds., 2004). "Au pair" means "an equal" in French. *Id.* Formal check-ins and counseling are a regular part of the au pair employment relationship. *Id.* The domestic employment relationship model detailed herein would apply to both groups of workers (trafficked workers and au pairs). I do not specifically analyze how such an application would uniquely apply for each of those subgroups. Certainly the persistent abuse of domestic workers and their lack of formal legal protections enables the continued, often hidden, trafficking of domestic workers. Likewise, the establishment of more formal protections in the au pair industry indicates that such protections are possible in the context of domestic work, but the analysis contained in this paper does not explicitly address either of those implications.

19. HOME IS WHERE THE WORK IS, *supra* note 3, at 7, 16; *see also* Adam J. Hiller & Leah E. Saxtein, *Falling Through the Cracks: The Plight of Domestic Workers and Their Continued Search for Legislative Protection*, 27 HOFSTRA LAB. & EMPLOY. L.J. 233, 233–56 (2009). In 2008, domestic workers earned less per hour than workers employed as locker room and coatroom attendants. A survey of New York City domestic employers revealed that the average pay for live-in nannies is $500-$550 for a fifty-hour workweek. 2008 NANNY SURVEY, *supra* note 14, at 7. Full time "live-out" nannies caring for one child earned, on average $12.75 an hour when paid weekly (employers paid $2-$3 more per hour for each additional child). *Id.* at 7, 31.

20. *See* Graunke, *supra* note 9, at 155 n.136 (finding domestic workers' wages to be "often below or near the minimum wage"); Peggy R. Smith, *Laboring For Child Care: A Consideration of New Approaches to Represent Low-Income Service Workers*, 8 U. PA. J. EMP. & LAB. L., 583, 591 (2006) (linking low pay for a job to the perception of the job as "women's work"); HOME IS WHERE THE WORK IS, *supra* note 3, at 16. A survey of hundreds of workers in Maryland found that 51% of those surveyed reported earning less than Maryland's minimum wage. GEORGE WASHINGTON UNIV. STUDENT RESEARCH TEAM (SPONSORED BY MONTGOMERY CNTY. COUNCIL COMM. ON HEALTH AND HUMAN SERVS.), WORKING CONDITIONS OF DOMESTIC WORKERS IN MONTGOMERY COUNTY, MARYLAND 8, 13 (2006), *available at* http://www.montgom-

ers reflects structural issues of racial discrimination and immigration policies.[21] A complex racial and ethnic hierarchy exists throughout the domestic work industry, which reflects macro issues of migration and opportunity. For example, early in the twentieth century, U.S. domestic workers were often young and single Irish or Scandinavian immigrant women, who performed "live in" domestic work.[22] As opportunities expanded for white immigrant or working class women to obtain factory or white-collar work, by the middle of last century, domestic workers became more likely to be married minority women performing "live out" day work.[23] Then, as affirmative action policies created opportunities for more women of color, and as more women overall entered the paid workforce, the demographics of domestic work changed again.[24]

Described as a "feminization of migration," recent decades have seen a rising "push/pull" dynamic, as women from poor countries migrate to the U.S. and other developed nations in response to an increasing demand from those countries for paid domestic or care work.[25] Ninety-five percent of domestic workers in New York today are people of color, 93% in New York are women, and 17% across the nation lack the legal authorization to work in the U.S.[26]

---

erycountymd.gov/content/council/pdf/REPORTS/domestic_workers.pdf. Hila Shamir, in her work on the distributive effects of employment law in markets of care, notes that the exclusion of domestic workers from employment law distributes the cost of care to domestic workers themselves, who subsidize the cost of their work to primary market workers. Hila Shamir, *Between Home and Work: Assessing the Distributive Effects of Employment Law in Markets of Care*, 30 BERKELEY J. EMP. & LABOR L. 404, 453–454 (2009). Only 2% of domestic employers surveyed reported providing full medical benefits to their nannies, and another 4% reported helping with doctor bills. 2008 NANNY SURVEY, *supra* note 14, at 33. The most ubiquitous "benefit" offered was "the ability to eat what she wants out of our kitchen," which 89% of surveyed employers reported providing. *Id.*

21. BERNHARDT, MCGRATH, & DEFILIPPIS, *supra* note 7, at 61–62.

22. Christine E. Bose, *The Interconnections of Paid and Unpaid Domestic Work*, THE SCHOLAR & FEMINIST ONLINE (2009), http://barnard.edu/sfonline/work/bose_01.htm.

23. *Id.*

24. *Id.*

25. *Id.*

26. HOME IS WHERE THE WORK IS, *supra* note 3, at 10; *see also* Anna Gorman, *Day Labor Centers See Some New Faces: Immigrant Women*, L.A. TIMES (June 11, 2007), http://articles.latimes.com/2007/jun/11/local/me-mujeres11 (estimating that undocumented immigrants comprise 17% of the cleaning industry). In addition to the children they are hired to care for, domestic workers often have to raise their own children, and many are supporting families back in their own countries through remittances, creating an effective "triple charge" for domestic work. *Laura Flanders*, *supra* note 9 (interview with Marisa Franco, Domestic Workers United).

Domestic workers who report mistreatment by employers cite their race, immigration status, and language skills as significant factors contributing to the abuse.[27]

Domestic employers include "high-income families who hire live-in housekeepers and nannies," "middle-class professionals who hire live-out domestic workers, either full-time or part-time," and "immigrant employers, including diplomats, who hire domestic workers from their home country or region."[28] But not all domestic employers are upper or middle-class homeowners; they also include single mothers, elderly people living on fixed incomes, college students, and apartment dwellers.[29]

Domestic work creates, and at times requires, exceptionally intimate relationships between employers and employees, which directly shape working conditions.[30] Domestic workers face a constellation of workplace challenges such as isolation, communication barriers, informal and inconsistent terms of employment, and complicated emotional attachments to the families for whom they work.[31]

The objection of some employers to more formal regulation of the industry is that such regulations are unnecessary because most employers are "generous," and treat their employees "like one of the family."[32] The testimony of domestic workers and surveys of do-

27. BERNHARDT, MCGRATH, & DEFILIPPIS, *supra* note 7, at 61. Of the workers who reported mistreatment, one-third (33%) felt that immigration status was a factor in their employer's actions, one-third (32%) felt race was a factor, and 18% felt language played a role. HOME IS WHERE THE WORK IS, *supra* note 3, at 2.

28. BERNHARDT, MCGRATH, & DEFILIPPIS, *supra* note 7, at 63. Third party intermediaries also play a role in the industry by placing workers with domestic employers. These include storefront employment agencies, domestic work temp agencies, professional traffickers, and day labor corners. *Id.* at 63; *see also* Elizabeth J. Kennedy, *The Invisible Corner: Expanding Workplace Rights for Female Day Laborers*, 31 BERKELEY J. EMP. & LAB. L. 126, 129, 137 (2010) (noting that female day laborers often seek work on open-air corners and through referral networks, websites like www.craigslist.org, and flyers).

29. HONDAGNEU-SOTELO, *supra* note 3, at 9.

30. As one domestic worker testified, "When a person goes to work in someone's house, she doesn't know what she'll find." *Domestic Working Hearing*, supra note 8 (statements of "Elizabeth," a domestic worker in Manhattan). As detailed in the testimony of several nannies and housekeepers, "what she finds" includes husbands that sexually harass and intimidate, employers who ignore or disregard their concerns, and expectations that they will work additional hours or perform additional duties without additional compensation or negotiation. *See generally id.*

31. *See* Graunke, *supra* note 9, at 150–72 (documenting the various problems facing workers). *See generally* Hiller & Saxtein, *supra* note 19, at 233–56.

32. HONDAGNEU-SOTELO, *supra* note 3, at 10. Some employers may believe that the provision of "in kind benefits" offsets the low wages and lack of benefits

mestic employers contradict this putative familial status.[33] Pretending that employees are "part of the family" obscures the true power dynamics inherent in an employment relationship. A nanny who cares for a child over several years, and who develops "familial" type feelings for that child, has no legal rights to continue seeing that child should the employer choose to terminate that relationship. This is the paradox: domestic workers are denied basic employment rights because they are "like family," yet denied basic family law rights because they are workers.

Even those employers who do understand and respect the obligations that come with being an employer of a domestic worker are often frustrated with the lack of publicly recognized domestic work standards. Many have resorted to "swapping notes" regarding pay, policies, and benefits, anxious that by failing to pay market rates or offer competitive benefits, they will be unable to retain qualified caregivers for their children or parents.[34] Annette Bernhardt, Policy Co-Director of the National Employment Law Project, described a recent online discussion among domestic employers regarding whether "financial belt tightening would (or should) impact the compensation of their domestic workers."[35] Bernhardt explains the fact that these "novices . . . some well-intentioned, others not, without any legal background or information on what makes for a sufficiency wage in the city" were "making up standards on the spot"

that characterize the industry. *See* 2008 NANNY SURVEY, *supra* note 14, at 10 (including a question about the donation of furniture to the domestic employee, presumably as evidence of other "in-kind" benefits provided to domestic employees). However, since the survey targeted only domestic employers, and not employees, it is unclear whether the recipient employees regarded such donations as benefits. *Id.*

33. One worker described being "made to wear a white uniform like Florence Nightingale," push the employer's dog in a stroller around Manhattan, and sleep next to the dog at night so that she could get up every four hours to put drops in the dog's eyes. *Domestic Worker Hearing*, *supra* note 8 (statement of Jocelyn Gill-Campbell, Organizer, Domestic Workers United); *cf.* 2008 NANNY SURVEY, *supra* note 14, at 10 (listing benefits such as aiding nannies with school applications and giving their nannies furniture); 2010 NANNY SURVEY, supra note 11, at 11 (explaining the various benefits afforded to nannies by employers, including "open kitchen" policies, food allowances, and early dismissals with pay).

34. Speaking about the positive impact that a Domestic Worker Bill of Rights would have for domestic employers, Dara Silverman explained that the law would "give clarity to the relationship . . . . There have been informal attempts by moms in Park Slope Brooklyn to compare notes—how much do you pay? What's your policy on vacation and sick days?" *Laura Flanders*, *supra* note 9 (interview with Dara Silverman).

35. *Domestic Working Hearing*, *supra* note 8 (statement of Annette Bernhardt, Ph.D., Policy Co-Director, National Employment Law Project).

demonstrates the very real consequences of the lack of industry regulation in the context of domestic work.[36]

The private home setting in which domestic work is performed also has a significant impact on working conditions.[37] Live-in workers are frequently isolated from the kinds of social networks that might otherwise help protect them from continued abuse or help them to escape dangerous employment.[38] Their workplaces are outside of public view and governmental inspection. Geographic isolation, as well as a lack of co-workers or a common employer, makes collective bargaining in the traditional sense hard to conceptualize (though not, as discussed below, impossible).[39] Moreover, the endemic inequality in bargaining power between domestic workers and employers and a dearth of industry standards undercuts one-on-one negotiations.[40] As Ai-jen Poo, a domestic worker organizer, explains, "When individual workers try to bargain with their employers, termination is the standard result since employers can simply hire another worker."[41]

B. The Limitations of Employment Law in Regulating the Domestic Work Industry

Traditional theories about the nature of the employment relationship "presuppose[ ] a world in which workers leave the confines of their private homes and travel to public workspaces."[42] What, then, of the domestic worker who labors in her employer's private home? When the workplace is a private home, the traditional work-

36. *Id.*

37. *See* Graunke, *supra* note 9, at 132–33 (arguing that the home setting makes domestic workers more vulnerable to abuse); Hiller & Saxtein, *supra* note 19, at 244.

38. Chelsy Castro, *Dying to Work: OSHA's Exclusion of Health and Safety Standards for Domestic Workers*, 4 MOD. AM. 3, 4 (2008); *see also* NELP IMMIGRANT & NONSTANDARD WORK PROJECT, NAT'L EMP'T LAW PROJECT, HOLDING THE WAGE FLOOR: ENFORCEMENT OF WAGE AND HOUR STANDARDS FOR LOW-WAGE WORKERS IN AN ERA OF GOVERNMENT INACTION AND EMPLOYER UNACCOUNTABILITY 9 (2006), http://nelp.3cdn.net/95b39fc0a12a8d8a34_iwm6bhbv2.pdf (showing the poor conditions faced by workers); Graunke, *supra* note 9, at 132–33 (explaining that workplace harassment becomes "domesticized," since it occurs in "the privacy of the home"); Swarna v. Al-Awadi, 622 F.3d 123, 129 (2d Cir. 2010) (listing the allegations by plaintiff against her former employer, a Saudi Arabian diplomat, including trafficking, fraud, seizure of her passport, forced labor, and sexual abuse).

39. FEASIBILITY STUDY, *supra* note 9, at 15.

40. POO, *supra* note 6, at 4; *see also* Smith, *supra* note 13, at 1841.

41. POO, *supra* note 6, at 8.

42. Smith, *supra* note 14, at 1841.

place law regimes must adapt to the changing need for, and implementation of, workplace regulation.

Although domestic employers largely view themselves as benevolent, the industry thrives by maintaining a vulnerable workforce and perpetuating a characterization of domestic work as unregulated and informal.[43] The history of employment laws, legal decisions, legislative history, and supporting documents evinces a sociolegal construction of domestic work viewed as separate from nearly all other kinds of work.[44] In 1905, one newspaper writer observed that "it is unusual to think of any question of law as between the housewife and the lady who condescends to do her cooking and general work."[45] The idea that the relationship between domestic employer and employee should be subject to the same legal parameters as all other employment was, and in many ways remains, "unusual."

However, as those (predominately of northern European descent) "ladies who condescend" found jobs in the industrial sector during the first half of the 20th century, domestic service became synonymous with black women, who came to comprise roughly 50% of the workforce.[46] State-sanctioned racial discrimination effectively precluded black women from attaining most of the work that was newly available to white women, and legislative efforts to regulate the domestic employment relationship received limited public support.[47] In blocking the extension of federal New Deal labor legislation to domestic and agricultural workers, Southern Democrats acted to preserve racial subjugation and cheap supplies of labor.[48]

43. BERNHARDT, MCGRATH, & DEFILIPPIS, *supra* note 7, at 61. In a recent report on Domestic Work, the International Labour Office notes that in drafting its proposed convention concerning decent work for domestic workers, it "modified the text so as to avoid any discriminatory characterization of women as inherently 'vulnerable.'" INTERNATIONAL LABOUR OFFICE, DECENT WORK FOR DOMESTIC WORKERS 3 (2011), *available at* http://www.ilo.org/wcmsp5/groups/public/@ed_norm/@relconf/documents/meetingdocument/wcms_143337.pdf.

44. *See* Smith, *supra* note 14, at 1855.

45. *Id.* at 1851.

46. *Id.* at 1857. As one student writer notes, "The forces of racism and patriarchy have shaped the legal landscape surrounding domestic workers. In the southeastern United States, the work has historical roots in slavery and in the southwest, in colonization and genocide." Terri Nilliasca, Note, *Some Women's Work: Domestic Work, Class, Race, Heteropatriarchy, and the Limits of Legal Reform*, 16 MICH. J. RACE & L. 377, 381 (2011). Today, domestic work is less associated with black women than it is with immigrant women of color. *Id.* at 385.

47. Smith, *supra* note 14, at 1857.

48. *Domestic Working Hearing*, *supra* note 8 (statement of Premilla Nadasen, Associate Professor of History, Queens College, City University of New York)

The legacy of racism is preserved in several key pieces of federal workplace legislation. The National Labor Relations Act (NLRA), which guarantees employees the right to organize a union and bargain collectively with their employers, specifically excluded domestic workers from its definition of "employee,"[49] an exclusion that persists today. Until 1974, the Fair Labor Standards Act (FLSA) completely excluded domestic workers from coverage for minimum wage rates, maximum hours, and overtime compensation.[50] Today, "casual" employees such as "babysitt[ers]" and "companions" for the sick or elderly, categories which include many domestic workers, remain excluded from FLSA coverage.[51] Furthermore, in the first reported case to address the scope of the companionship-services exemption, *McCune v. Oregon Senior Services Division*, the Ninth Circuit held that the exemption applied to a group of home-care workers who "live[d] with their clients at a near poverty level providing around-the-clock care."[52] The workers were paid by various public entities, including the state of Oregon, to care for elderly and infirm clients.[53] The court found that the workers were "domestic service" employees, but it nevertheless excluded them from FLSA coverage because their work qualified as "companionship services."[54] While some states do include home care workers within

(describing New Deal-era concessions demanded by Southern congressmen). According to Rachel McCullough, organizer with Jews for Racial and Economic Justice, the response of some legislators to the demands by domestic workers for industry standards has been, "Why are we going to give special protections to this workforce?" To which the campaign's response has been, "If you thought we were 'special' enough to exclude all these years . . ." Telephone Interview with Rachel McCullough, Organizer, Jews for Racial and Economic Justice (Apr. 14, 2011).

49. 29 U.S.C. § 152(3) (2006). Even some legislators are unaware of this exemption, as evidenced by the position some New York State legislators took in response to legislation proposing paid time off and other benefits for domestic workers that they "form a union and collectively bargain like other workers." POO, *supra* note 6, at 8.

50. Smith, *supra* note 14, at 1860.

51. 29 U.S.C. § 213(a)(15) (2006); *see also* HOME IS WHERE THE WORK IS, *supra* note 3, at 4. Furthermore, live-in domestic workers, unlike most other employees in the U.S., cannot get overtime under the FLSA. 29 U.S.C. § 213(b)(21). Under New York state law, while domestic workers who do not live in their employer's home are entitled to overtime at a rate of one-and-a-half times their regular rate after forty hours of work in a week, live-in domestic workers are only entitled to overtime at a rate of one-and-a-half times the minimum wage, and then only after forty-four hours of work in a week. N.Y. COMP. CODES R. & REGS, tit. 12, § 142-2.2 (2010).

52. McCune v. Or. Senior Servs. Div., 894 F.2d 1107, 1110 (9th Cir. 1990).

53. *Id.* at 1108.

54. *Id.* at 1108, 1110.

their wage and hour regulations, most states follow the Ninth Circuit's broad application of the companionship exemption for home care workers.[55]

Not only are domestic workers denied protections under the FLSA, but regulations promulgated under the Occupational Safety and Health Act explicitly exclude domestic workers "[a]s a matter of policy."[56] Despite the fact that domestic workers directly support the continued participation of their employers in the paid workforce, they are not considered under the law to be "employed in a business of [their employer] which affects commerce."[57] Since domestic employers are typically homeowners with minimal or nonexistent safety or health training, or supplies like gloves or nontoxic cleaning products, this exclusion leaves domestic workers vulnerable to exposure to toxic chemicals, unsafe appliances, unsanitary conditions, and other risks related to poorly maintained homes, including poor lighting and decrepit stairs.[58] Sixty-three percent of domestic workers surveyed in California "considered their jobs hazardous, citing concentrated exposure to toxic cleaning chemicals and human contagions, risk of injury from cleaning high or difficult-to-reach places, and heavy lifting."[59] Seventy-five percent had not received from their employers "any protective gear such as facemasks or gloves to prevent workplace injuries," and 86% had not received "training in job safety or workplace injury preven-

55. PAUL K. SONN, CATHERINE RUCKELSHAUS & SARAH LEBERSTEIN, NATIONAL EMPLOYMENT LAW PROJECT, FAIR PAY FOR HOME CARE WORKERS: REFORMING THE U.S. DEPARTMENT OF LABOR'S COMPANIONSHIP REGULATIONS UNDER THE FAIR LABOR STANDARDS ACT 2 (2011), *available at* http://www.nelp.org/page/-/Justice/2011/FairPayforHomeCareWorkers.pdf?nocdn=1 (explaining that twenty-one states have given labor law protections to domestic workers).

56. 29 C.F.R. § 1975.6 (2009).

57. *Id.*

58. *See* RICK NEITZEL & NOAH SEIXAS, DEP'T OF ENVTL. & OCCUPATIONAL HEALTH SCIS., UNIV. OF WASH., DAY LABOR SAFETY AND HEALTH INITIATIVE REPORT 2 (2005) (describing generally poor working conditions, high injury rate, and lack of employer responsibility); *see also* HOME IS WHERE THE WORK IS, *supra* note 3, at 23 ("[N]ot only are new immigrants less likely to complain about job hazards, but they also tend to return to work quickly despite potentially serious job-related injuries and illnesses." (quoting AFL-CIO, IMMIGRANT WORKERS AT RISK: THE URGENT NEED FOR IMPROVED WORKPLACE SAFETY AND HEALTH POLICIES AND PROGRAMS 10 (2005), *available at* http://www.aflcio.org/aboutus/laborday/upload/immigrant_risk.pdf)); Megan Tady, *Unprotected by Laws, Domestic Workers Face Exploitation*, THE NEW STANDARD (Mar. 14, 2007), http://www.indybay.org/newsitems/2007/03/20/18380387.php; Smith, *supra* note 14, at 1873–74.

59. MUJERES UNIDAS Y ACTIVAS ET AL., BEHIND CLOSED DOORS: WORKING CONDITIONS OF CALIFORNIA HOUSEHOLD WORKERS 6 (2007), *available at* http://www.datacenter.org/reports/behindcloseddoors.pdf.

tion."[60] Nearly one-third had suffered "an injury or illness requiring medical attention" in the year prior to being surveyed, yet two-thirds of those injured did not receive medical attention, since they were neither able to afford treatment nor received employer-provided health care.[61] Testifying before the New York State Assembly Committee on Labor, the daughter of a domestic worker described a "particularly mischievous child" that tripped her mother on the stairs, "seriously injuring her already fragile knees, and breaking her front teeth." Eight years later, her daughter is still struggling to pay back her mother's medical loans.[62]

Likewise, although domestic workers report widespread discrimination on the basis of sex and race, including sexual harassment, Title VII of the Civil Rights Act applies only to employers with 15 or more employees, which excludes most domestic workers.[63] For the same reason, most domestic workers cannot seek relief under the Americans with Disabilities Act,[64] or the Age Discrimination in Employment Act,[65] each of which prohibits workplace discrimination.

In addition to these exclusions at the federal level, most states do not extend to domestic workers rights to overtime, sick time, vacation, health care, or workers' compensation.[66] Moreover, domestic workers often encounter significant barriers to asserting the

60. *Id.*

61. *Id.*

62. *Domestic Working Hearing*, *supra* note 8 (statement of Priscilla Gonzalez, Organizer, Domestic Workers United).

63. 42 U.S.C. § 2000e(b) (2006). Title VII prohibits discrimination on the basis of race, color, religion, sex, or national origin. *See, e.g.*, *Domestic Worker Hearings*, *supra* note 8 (statements of various domestic workers) (presenting accounts of abusive workplace behavior, all of which occurred in environments with fewer than fifteen employees). One domestic worker recounts, "How can I forget the numerous times when I resigned from a fulltime housekeeping position to avoid malicious sexual harassment of male employers?" *Id.* (statement of Monica Ledesma, Member, DAMAYAN Migrant Workers Association).

64. 29 U.S.C. § 630(b) (2006) (restricting the definition of "employer" to those who employ twenty or more employees).

65. 42 U.S.C. §§ 12111(5)(A) (2006) (restricting the definition of "employer" to those who employ fifteen or more employees).

66. Smith, *supra* note 14 at 1852. Of the twenty-six states with minimum wage laws in 1940, only Wisconsin had enacted a specific minimum-wage order applicable to domestic service. *Id.* Montgomery County, in Maryland, passed a Domestic Workers' Law that provides, among other things, a requirement that certain employers of domestic workers living in the county enter into a written employment contract that specifies the terms and conditions of employment. MONTGOMERY COUNTY, MD., CODE ch 11, § 11-4B (2010), *available at* http://www.montgomerycountymd.gov/content/ocp/domestic/pdfs/Law.pdf.

limited rights they do possess, including a lack of knowledge about existing employment and labor laws,[67] insufficient information regarding where to report violations,[68] collateral immigration consequences,[69] and, for domestic workers who lack the legal authorization to work in the U.S., the fear of being deported and separated from their families.[70]

Domestic workers are particularly vulnerable to misclassification as "independent contractors,"[71] which can further strip them of legal protections otherwise guaranteed by minimum wage standards, workers' and unemployment compensation laws, and Social

67. *See* NELP IMMIGRANT & NONSTANDARD WORK PROJECT, *supra* note 30, at 10 (finding that 61% of day laborers do not know their rights).

68. Laura Dresser, *Cleaning and Caring in the Home: Shared Problems? Shared Possiblities?, in* ANNETTE BERNHARDT ET AL., THE GLOVES-OFF ECONOMY 119 (2008); *cf.* NELP IMMIGRANT & NONSTANDARD WORK PROJECT, *supra* note 38, at 10 (finding that 80% of day laborers do not know where to report workplace abuses). As one housekeeper recounted, "My employer . . . did not allow us to sit down or talk to other people. During lunchtime, we were not allowed to use their utensils. We were supposed to use disposable plates, spoons, forks and cups. After using them, we were supposed to put them in the dishwasher and use them again. She yelled for no reason. She insisted on scrubbing the carpet on my knees. Every time she came into the room, I was supposed to stand. When she would pass by, I'd have to stand aside and not look at her. She always made me feel stupid." HOME IS WHERE THE WORK IS, *supra* note 3, at 21–22.

69. BONATO, *supra* note 3, at 9 ("Because domestic workers' visas are tied to their employment, if an employer fires his worker for reporting a violation, the worker could face deportation or may economically be unable to stay in the United States if she can not obtain work authorization.").

70. *See, e.g.,* Rivera v. NIBCO, 364 F.3d 1057, 1064–65 (9th Cir. 2004) (agreeing that "most undocumented workers are reluctant to report abusive or discriminatory employment practices" because they fear criminal prosecution and deportation), *cert. denied*, 544 U.S. 905 (2005); *cf.* NELP IMMIGRANT & NONSTANDARD WORK PROJECT, *supra* note 38, at 9 (finding that 67% of day laborers fear being reported to the INS). This article uses the term "unauthorized" workers to describe immigrant workers who do not have the legal authorization to work under U.S. laws, and who are distinguishable from "undocumented" workers, or a subset of the immigrant population that is unauthorized to work.

71. Domestic workers are considered household employees regardless of whether they are paid on an hourly, daily or weekly basis, or by the job, or whether they are hired through an agency, so long as the employer can control how the work is done. DEP'T OF TREAS., I.R.S. PUBL'N 926, HOUSEHOLD EMPLOYER'S TAX GUIDE 2–3 (2011), *available at* http://www.irs.gov/pub/irs-pdf/p926.pdf. Self-employed independent contractors generally supply their own cleaning equipment and materials; if provided by the employer, the domestic worker is more likely to be an employee, not an independent contractor. *Id.*

Security and disability benefits.[72] While they often work "independently," without direct supervision, domestic workers' testimonials about their workdays paint a very different picture of independence and control. Parents and homeowners may sometimes dictate with extreme specificity the manner in which the house is to be cleaned, how child or parent is to be cared for, and what supplies will be used.[73] Many domestic workers occupy a legal no man's land: neither protected by workplace laws, nor possessing self-employed independent contractors' control over their work. While other countries, such as Canada, provide for such workers using a "dependent contractor" classification, U.S. employment and labor law lacks such a classification.[74]

Domestic workers have spent decades trying to upend the attitude that because domestic work is unlike all other work it is therefore unsuitable for formal regulation. It is important, however, to recognize the ways in which domestic work can legitimately be distinguished from other occupations, and the ways in which the relationship between domestic employers and employees is unique, in order to more effectively regulate the industry.

C. Distinguishing Domestic Work

What is it about domestic work that relegates domestic workers to a second-tier status? As discussed in the previous sections, the gendered association of the work, patterns of global migration and immigration policies, and race-based discrimination, combined with the charged, private setting of the home and intimate nature of the work, all serve to set domestic work apart from most other types of paid employment.[75] In addition to the unique qualities of

72. Katherine V. W. Stone, *Legal Protections for Atypical Employees: Employment Law for Workers Without Workplaces and Employees Without Employers*, 27 BERKELEY J. EMP. & LAB. L. 254, 279–80 (2006).

73. *See, e.g.*, Tady, *supra* note 58 (explaining how one domestic recounts of an employer who "made her wash her hands with ammonia before starting work," a practice that "burned [her] hands"). Most employers in one survey in the Brooklyn area reported some kind of restrictions on the discretion of their nannies. *See* 2008 NANNY SURVEY, *supra* note 14, at 29. Only half of surveyed parents agreed with the statement, "Our nanny has her own style and we accept that she won't do things exactly like we would." *Id.* at 49.

74. See Elizabeth J. Kennedy, Comment, *Freedom from Independence: Collective Bargaining Rights for "Dependent Contractors"*, 26 BERKELEY J. EMP. & LAB. L. 143, 148–49, 153 (2005).

75. POO, *supra* note 6, at 4; DOMESTIC WORKERS UNITED ET AL., DOMESTIC WORKERS AND COLLECTIVE BARGAINING: A PROPOSAL FOR IMMEDIATE INCLUSION OF DOMESTIC WORKERS IN THE NEW YORK STATE LABOR RELATIONS ACT 7 (2010) [hereinafter DWU REPORT], *available at* http://www.urbanjustice.org/pdf/publications/

domestic work and those who perform it, domestic *employers* are in many ways "fundamentally different" from their counterparts in other industries.[76] They hire nannies and housekeepers not as part of their primary business, but in addition to (and so that they may maintain) their own jobs.[77] Employers themselves (many of whom are also women) are experiencing the economic strain of the Great Recession, as well as the elimination of a social safety net that could otherwise help manage the costs of child and elder care.[78] These men and women are, in large part, reluctantly thrust into employer roles. They often regard themselves as consumers of domestic services, not as employers of domestic workers, and do not view their private homes as workplaces.[79] Rather than provide their employees with direct instruction, domestic employers often shy away from engaging in the kind of negotiation and communication required. In the context of housekeeping, many even prefer to be out of the house when the work is performed.[80] However, treating domestic employees like contractors or one-time service providers has consequences for both parties. An employment *relationship* is, under most

Domestic_Workers_and_Collective_Bargaining.pdf ("As compared to other workplaces, the home is a charged, sensitive space often lacking the emotional distance necessary for negotiation to take place."); Jennifer Steinhauer, *City Nannies Say They, Too, Can Be Mother Lions*, N.Y. TIMES (July 16, 2005), http://www.nytimes.com/2005/07/16/nyregion/16nanny.html?scp=1&sq=%22brunilda+tirado&st=nyt (explaining that "[m]utual suspicions and resentment" arise, since a baby sitter may "know every inch of her employer's home," while "many parents never set foot in their nanny's neighborhood, let alone her home").

76. FEASIBILITY STUDY, *supra* note 9, at 2.

77. 2008 NANNY SURVEY, *supra* note 14, at 8.

78. Based on survey data collected by the Park Slope Parents organization, the economy has caused four in ten, or 39% of, domestic employers to make changes to the employment relationship they have with their nanny, including cutting back the number of hours worked and increasing "nanny sharing" with another family. 2010 NANNY SURVEY, *supra* note 11, at 15. Survey data also revealed a pattern of decreased pay rates, and well as fewer and smaller raises than were reported by the same organization in 2008. *Id.* As described by a domestic employer organizer, domestic employers are experiencing "an assault on their way of life, the elimination of the social safety net, an assault on working mothers, an assault on the ability of these women to be professional and be a mother [sic] at the same time." Interview with Rachel McCullough, *supra* note 48.

79. While speaking at a Human Rights Tribunal organized by DWU, one domestic employer relayed the following:

> The first time I heard Debbie—our son's caregiver—refer to me as her boss, I was taken aback. The word seemed too formal. I had hopes for the kind of intimacy I'd known other parents and nannies to experience, and I wanted Debbie to relate to me as someone other than her employer.

POO, *supra* note 6, at 14.

80. HONDAGNEU-SOTELO, *supra* note 3, at 47.

of the existing models, just that: a relationship. And one that, like any other, requires investments of resources and a commitment to communication in order to ensure that the relationship is mutual, sustainable, and successful. Hiring a contractor to replace your hot water heater requires no such investment by either party, and yet many domestic employers regard their employees in that light. Moreover, skilled craftsmen and contractors who perform work in private homes typically do not suffer the kind of abuse endemic to domestic work and are protected by state and federal laws.

Even when domestic employers recognize the additional responsibilities that attach to employers, as opposed to consumers, of domestic work, few are able (while struggling to balance the needs of their own families and their work responsibilities) to take on additional administrative and human resource roles.[81] Unlike most other employers, they "cannot pass on the costs of domestic worker salaries or benefits to consumers."[82] In addition, because of a complicated set of emotions related to the need for paid care work,[83] domestic employers often feel awkward discussing terms and conditions of that work.[84] An informal hiring process, in which workers

81. *See* 2008 NANNY SURVEY, *supra* note 14, at 52–53. Only 17% of employers entered into a "Nanny contract" with their nanny, and only 35% of employers reported having conducted a "review" with the nanny in which they had discussed the nanny's work (almost the same percentage that had "shown up unexpectedly in order to verify that a nanny is doing a good job"). *Id.* at 4, 52–53. Seventy-two percent reported at least one instance when they had not let their nanny know that something was bothering them, 35% had came home later than expected without notice, and 21% had failed to pay their nanny on time, a cognizable violation of wage and hour law. *Id.* at 53. One domestic employer, speaking in support of the Domestic Workers Bill of Rights, confessed the following:

> [M]y resistance to seeing myself as an employer meant that it took too long for Debbie [her son's nanny] to be treated like an employee. Rather than signing a contract and agreeing to the terms of work on day one, we talked about benefits casually, after she'd already started work. I would not have tolerated such lack of professionalism in my own job.

POO, *supra* note 6 at 14.

82. FEASIBILITY STUDY, *supra* note 9, at 16. Some domestic employers report providing raises that are tied to the raises that they or their spouse received; raises tended to be $1 an hour, per year. 2008 NANNY SURVEY, *supra* note 14, at 38.

83. Describing many domestic employers, an organizer with Jews for Racial and Economic Justice explained, "I've learned so much from the women we organize about the shame and guilt they feel from not being able to spend enough time with their own children. Their own lives [are] more precarious, they worry about layoffs, budget cuts." Telephone Interview with Rachel McCullough, *supra* note 47.

84. FEASIBILITY STUDY, *supra* note 9, at 15. In her testimony before the New York State Assembly Committee on Labor, Annette Bernhardt, Policy Co-Director of the National Employment Law Project, described a recent online discussion among domestic employers regarding whether "financial belt tightening would (or

are hired largely through word-of-mouth referrals and few enter into formal employment agreements, is perpetuated by a lack of publicly available data on salaries, benefits, and policies in the industry.[85] Operating in such a vacuum, 78% of employers felt they paid their employee an average or above average salary.[86]

Another reason for the persistent lack of domestic workplace regulation is the "home as castle" doctrine, which structures the relationship between governments and individuals within many different contexts, including search and seizure prohibitions,[87] the use of self-defense,[88] and tort law.[89] The labor performed by domestic workers inside private homes is also performed by workers who en-

should) impact the compensation of their domestic workers" and noted that the fact that these "novices . . . some well-intentioned, others not, without any legal background or information on what makes for a sufficiency wage in the city" were "making up standards on the spot" demonstrates the very real consequences of the lack of industry regulation in the context of domestic work. *Berhardt Statement*, *supra* note 35, at *3. When asked, "What do you do," most domestic employers probably would not include their role as an employer of a nanny, housekeeper, or elder caregiver. *See supra* note 32 and accompanying text. The work of sociologist Pierrette Hondagneu-Sotelo explores in depth domestic employers' reluctance to view themselves as employers and their homes as work sites, citing ambivalence, embarrassment, uncomfortable associations with feudalism and slavery, and guilt. HONDAGNEU-SOTELO, *supra* note 3, at 10–11, 139. Moreover, since domestic worker employers, in contrast to many service sector employers, are "the ultimate consumers of their employees' services," a conversation about such services is "an even more delicate issue." FEASIBILITY STUDY, *supra* note 9, at 15.

85. 2008 NANNY SURVEY, *supra* note 14, at 4, 45. "Through a friend" is the number one way respondents found their Nannies. *Id.* at 4. Only 17% of employers enter into a "Nanny contract" with their employees. *Id.* Seventy-eight percent of employers of full-time nannies believe that they pay their nannies either about the same or more than most people. *Id.* at 45. Only 13% believe that they paid less than other people. *Id.*

86. *Id.* at 45.

87. *See* Thomas Y. Davies, *Recovering the Original Fourth Amendment*, 98 MICH. L. REV. 547, 642–43 (1999) (noting that at common law, the home was accorded sacrosanct legal protection under the castle doctrine; for example, intrusions into the home required a judicial warrant). While contemporary law extends the Fourth Amendment to all privately held property, it does still recognize that one's home—one's castle—is entitled to special protection. *See, e.g.*, Kyllo v. United States, 533 U.S. 27, 34 (2001) (holding that the use of thermo-imaging technology to detect the presence of narcotics in a home constituted a "search" for Fourth Amendment purposes).

88. *See* Weiand v. State, 732 So. 2d 1044, 1049 (Fla. 1999) (explaining that one need not retreat "to the wall" in one's home from an intruder to claim self-defense).

89. Linda C. McClain, *Inviolability and Privacy: The Castle, the Sanctuary, and the Body*, 7 YALE J.L. & HUMAN. 195, 202–203 (1995) ("The home as castle appears in defenses of privacy rights sounding in tort law.").

joy very different regulatory protections and negotiate wages in very different economic markets. For example, daycare center employees, nursing home aides, hotel housekeepers, and restaurant line cooks are all entitled by statute to safe working conditions, to the right to form unions, and to workplaces free from sexual harassment. While employers enjoy some modicum of privacy, many employment laws are enforced, at least in part, through a system of workplace inspections.[90] Such a system would be exceedingly difficult to reproduce in the domestic work industry, given the practical limitations (including the lack of a centralized database needed to locate domestic workers within thousands of individual workplaces), as well as Fourth Amendment limitations and public opposition to the regulation of private, home-based activity.[91]

Domestic employers, and their children or parents, often develop emotional bonds that distinguish domestic work from most other employment relationships. As an illustration, *The New York Times* recently reported on the "bewilderment" of rescue workers upon discovering that the child for whom a woman risked her life to save in the wake of a collapsing building (screaming, "My baby! My baby!") was not, in fact, the child's mother.[92] She was the infant's nanny. The *Times* also reported the "surprise[ ] by the surprise," felt by fellow domestic workers. One explained, "These children are your babies because you are their parents all day," and another remarked, "We didn't make them . . . . [b]ut they're ours and we love them dearly, dearly."[93] These emotional attachments cut both ways—they can cause a domestic worker to tolerate conditions she might not in another workplace, but they can also make it less likely that she would be arbitrarily terminated, since she is much less fungible than she might be in another workplace.

However, as the New York State Department of Labor notes, in distinguishing domestic work from other industrial or service indus-

90. The Fourth Amendment protects individuals against unreasonable governmental intrusion into their homes, and, on a more limited basis, protects employers against such intrusions into workplace in the public sphere. For an interesting examination of Fourth Amendment issues in the context of ICE enforcement strategies, see Marisa Antos-Fallon, Note, *The Fourth Amendment and Immigration Enforcement in the Home: Can ICE Target the Utmost Sphere of Privacy?*, 35 FORDHAM URB. L.J. 999 (2008).

91. *See supra* note 87–90 and accompanying text; *see also* Smith, *supra* note 14, at 1857–58 (explaining that early objections to regulating domestic work stemmed from the perception that home life should not be bound by the rigid market transactions of the public sphere).

92. Steinhauer, *supra* note 75.

93. *Id.*

tries, these relationships often have a "presumptive end point: in the case of childcare . . . when the children enter elementary school; and in the case of elder care . . . when an elderly person goes into a nursing home or dies."[94] And yet, any presumed duration of employment is not normally a consideration for whether or not labor and employment law protections should apply.[95] For domestic workers, however, collective bargaining and other labor and employment rights remain elusive.

D. Building Power Through Alliances: Domestic Workers United

Over the past decade, domestic workers have organized dozens of affinity groups and developed legislative campaigns. In New York City, Domestic Workers United (DWU) has organized thousands of predominately Caribbean, Latina, and African nannies and elder caregivers.[96] In addition to improving working conditions and building power and respect for domestic workers, DWU's mission is to build a broader movement for social change.[97] To do this, DWU has developed a coalition of allies that includes domestic employers, the children of employers, communities of faith, and the larger labor movement.[98] Those alliances proved critical to advancing a statewide legislative campaign and creating pathways for domestic employers to voluntarily improve conditions for domestic workers in their homes.[99]

---

94. FEASIBILITY STUDY, *supra* note 9, at 15.

95. *See* Cmty. for Creative Non-Violence v. Reid, 490 U.S. 730, 751 (1989) (stating that duration plays a role as part of an eleven part test in determining employment status but "no one of these factors is determinative").

96. DWU REPORT, *supra* note 75, at 1; *see also* DOMESTIC WORKERS UNITED, http://www.domesticworkersunited.org.

97. PAM WHITEFIELD, SALLY ALVAREZ, & YASMIN EMRANI, CORNELL UNIV. ILR SCH., IS THERE A WOMAN'S WAY OF ORGNIZING? GENDER, UNIONS, AND EFFECTIVE ORGANIZING 12 (2009), *available at* http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=1025&context=reports. With the help of the Urban Justice, City University of New York Immigrant and Redugee Rights Clnic, and other legal partners, DWU has helped domestic workers recover over $450,000 in stolen wages. POO, *supra* note 6, at 5. In 2002, DWU successfully campaigned to compel domestic worker employment placement agencies to educate workers and employers about basic labor rights. *Id.* Additionally, DWU offers ESL training, GED classes, and a Nanny Training Program in association with Cornell Labor Studies. The Nanny Training Program teaches domestic workers basic pediatrics and negotiation skills. WHITEFIELD, ALVAREZ, & EMRANI, *supra* note 94, at 21.

98. POO, *supra* note 6, at 6. Thousands of New Yorkers took action in support of domestic workers, demonstrating the relevance of domestic work relationships to the lives of so many different people, including the children of domestic workers and the children cared for by domestic workers. *Id.* at 7.

99. *Id.* at 2.

DWU was able to create these unlikely alliances between workers and employers, in part by focusing their campaign on common issues of structural inequality that affect domestic workers and many employers. These include the devaluation of women's work, the widening income inequality, and a shrinking social safety net that places stress on domestic workers and employers alike.[100] Thus, without glossing over the stark racial and economic inequities between domestic workers and their employers, DWU was able to develop strategic alliances with domestic employers who wanted to be fair employers—working mothers supporting other working mothers.[101]

Likewise, the emotional bonds generated between nannies and children were recast by DWU as further justification for equal rights. In the summer of 2009, DWU organized a march down Broadway led by the children of domestic workers and children of domestic employers, demanding the passage of the New York Domestic Workers' Bill of Rights.[102] Following the march, children spoke about the role that domestic workers had played in their own lives. As one child related, "[My nanny] raised me—which is not the same thing as being paid to do a job. She taught me, she accepted me, and if I had not known her, if she had not supplied those things, I don't know what I'd be now, or who."[103]

In building a broad base of allies, DWU sought to change individual relationships between domestic workers and their employers, as well as make structural changes within the industry. The campaign brought together labor, religious, and community groups united by the idea that "every one of us has needed care, provided care, or relied upon someone else for care at some point in our

100. *Id.* at 11. Recognizing that many domestic workers are mothers, as are many domestic employers, Moms Rising, an advocacy organization based in New York was enlisted as another ally in the campaign. *Laura Flanders*, *supra* note 9 (interview with Deirdre Schiefling, Center for Working Families).

101. Poo, *supra* note 6, at 13. This sentiment is shared by the organizers of high-road domestic employers as well. Explains Rachel McCullough, "We want to build a consciousness in which women can be in the same boat, without blurring real issues of privilege, racism, class, what it feels like to employ an immigrant women of color, why don't we want to talk about it? . . . Our organizing model needs to analyze these relations of power in a new way. Those who fall under the label of employers have a stake in justice for all workers, building up the left and building up the percentage of workers who have access to collective bargaining." Telephone Interview with Rachel McCullough, *supra* note 48.

102. *Pictures from June 14th's Children and Families March*, Domestic Workers United (June 24, 2009, 10:21PM), http://domesticworkersunited.blogspot.com/2009/06/pictures-from-june-14th-children-and.html.

103. Poo, *supra* note 6, at 12.

lives."[104] DWU bet, successfully, that if they framed the campaign around values, "people will choose fairness and love even when it cuts against their immediate self-interest."[105]

II.
EMPLOYMENT RELATIONSHIP THEORY AND DOMESTIC WORK

Industrial relations and legal scholars conceptualize the employment relationship in various ways, using contract theory, neoliberal economic paradigms, human resources and pluralistic models, and models based on social norms. Central to almost every paradigm is the assumption that the relationship contains conflicts of interest between employers and their employees.[106] Each of the predominant theories and how they might frame the domestic employment relationship will be addressed in turn.

Contract theorists use a "norms versus contracts" framework to determine whether employers and employees have structured their relationship in ways that maximize their economic self-interests.[107] Viewed through that lens, employers and employees rely heavily on a combination of social norms and legal contracts, which either are "self-enforcing," or require "judicial enforcement."[108] Contract law

104. *Id.* at 13.

105. *Id.*

106. *See* Christopher T. Wonnell, *The Influential Myth of a Generalized Conflict of Interests Between Labor and Management*, 81 GEO. L.J. 39, 39 (1992).

107. Michael L. Wachter, *Theories of the Employment Relations: Choosing Between Norms and Contracts* 1, 6 (Scholarship at Penn Law, Paper 70, 2005), *available at* http://lsr.nellco.org/cgi/viewcontent.cgi?article=1074&context=penn_wps.

108. *See id.* at 28 (acknowledging that in the context of employment regulations, certain protections are left to "judicial enforcement," while others are left to a firm's "hierarchical governance mechanism outside the purview of judicial review"). Informal norm governance is most effective in "repeat-play" situations, in which an aggrieved party can leverage the frequency of the employee-employer interaction to sanction and deter bad faith action. *Id.* The degree to which norms are relied upon more strongly than contract terms, or the extent to which such contracts require third party enforcement, depends on the nature of the employment, the employee, and the employer. *Id.* at 2. For example, in a unionized workplace, it is assumed that the collective bargaining agreement has been negotiated in order to maximize the economic self-interests of employees and employers. In the nonunion firm, employees and employers more heavily rely on norms to guide their behavior and resolve disputes. Enforcing contracts through litigation is "expensive and wastes resources"; therefore, even contracts that require judicial enforcement are considered most effective when they deter breach, rather than remedy breaches through judicial enforcement. *Id.* at 24. Contracts protect the interest of each party only to the extent that the contract contains terms that accomplish the goals of each party, and only to the extent that both parties expect

is considered to be an efficient tool for shaping the contours of an employment relationship only when workers voluntarily enter into their relationship and negotiate an agreement that has few, if any, mandatory terms.[109] This model treats issues of fairness and unequal bargaining power as issues of "opportunistic behavior and asymmetric information," to be corrected by market forces.[110] The labor contracting literature is helpful when examining how domestic employees and employers resolve problems arising from the employment relationship, and how they integrate the four central elements of industrial organization theory: (1) match-specific assets, (2) asymmetric information, (3) risk-aversion, and (4) transaction costs.[111] In analyzing a particular employment relationship and determining whether norms or contract law should govern, contract theorists ask how the terms of the employment relationship work to protect the parties' agreement with respect to these four elements. Do the terms of the contract maximize mutual interests? Are the mechanisms used to enforce the contract adequate?[112]

Under contract theory, employers are deterred from acting opportunistically because once the labor market discovers the poor conduct, they suffer reputational losses that outweigh any potential gains. Such an employer would face higher labor costs in the long run, since it would eventually be forced to pay higher wages to induce new workers to join the firm.[113] In the case of domestic workers, however, workers have to negotiate the terms of their employment individually, day-by-day, house-by-house, in situations where they lack any real bargaining power.[114] Very few domestic workers negotiate formal, written agreements regarding their employment.[115] Instead, domestic working conditions are often unilat-

that the contract can and will be enforced. *Id.* Statutory protections, such as the NLRA, are justified by labor contract scholars based on the perceived vulnerability of most workers in their dealings with employers. When parties omit terms to a contract, "the courts fill the gaps with the default terms of contract law." *Id.* at 22. Moreover, "there is widespread agreement that when a contract is inadvertently incomplete, the court should and, in fact, does fill the gap by adopting the term that the parties themselves would have written had they appreciated the contingency." *Id.*

109. *Id.* at 31.

110. *Id.* at 3–7.

111. *Id.* at 6. The term "match-specific assets" refers to job-specific productivity an individual brings to his or her work, based on experience or training. *Id.* at 7.

112. *Id.* at 6–7.

113. *Id.* at 21.

114. Poo, *supra* note 6, at 13.

115. 2008 Nanny Survey, *supra* note 14, at 4.

erally based on informal, non-standard, gendered, and race-based social norms and cultural values.[116]

Simply grafting contract principles onto the domestic employment relationship disregards the range of relational interests that are at the heart of that relationship.[117] For example, caring work, like most domestic work, requires attachment, affiliation, intimate knowledge, patience, and listening.[118] However, most models of employment relationship are based on concepts derived from the manufacturing sector, which focus on inputs, outputs, efficiency, and productivity. While not wholly inapplicable, the utility of contract theory, or any of the following prevailing theoretical models, in guiding regulation of the domestic workplace is inadequate.[119]

Neoclassical "egoist" economic models emphasize the role of the market in mediating those conflicts and balancing the inherently oppositional interests between employer and employee.[120] This paradigm makes the assumption that employers and employees act according to their own self-interest, within perfectly competitive labor markets.[121] The economic treatment of labor as a commodity considers domestic work to be "nonproductive," since it is, in economic parlance, "immediately consumed" by the family.[122]

116. *Cf.* Bruce E. Kaufman, *The Future of Employment Relations, Insights in Theory* 5 (W.J. Usery Workplace Research Group Paper Series, Working Paper No. 2010-1-3, 2010), *available at* http://aysps.gsu.edu/usery/files/UWRG_10-1-3_Kaufman.pdf ("[T]he employment relationship creates an asymmetrical control and power relationship in the organization, moderated to various degrees by market forces, employment laws, social norms, and cultural values.").

117. Graunke, *supra* note 9, at 158.

118. HONDAGNEU-SOTELO, *supra* note 3, at 10.

119. *Cf.* POO, *supra* note 6, at 9 (explaining that the DWU campaign is built on an organizing model that rejects a narrow definition of "material self-interest"—"the historical assumption on which a great deal of organizing models are based"—and envisions the possibility of more transformative relations between domestic employer and worker—"a campaign that mobilized many different communities of people based on an expanded sense of self-interest that acknowledged our relationships and our interdependencies"—than the more narrow conflict-driven approach of all of the existing theoretical models).

120. George R. Boyer & Robert S. Smith, *The Development of the Neoclassical Tradition in Labor Economics*, 54 INDUS. & LAB. REL. REV. 199, 210 (2001) (noting the determination of the neoclassical labor economist "to find maximizing behavior and equilibrium outcomes throughout the labor market").

121. John W. Budd, *Fairness at Work, and Maybe Efficiency but Not Voice*, 29 COMP. LAB. L. & POL'Y J. 477, 479–480.

122. Smith, *supra* note 14, at 1855; *see also* POO, *supra* note 6, at 18 ("[B]ecause women's work in the home has never been factored into national labor statistics, it is difficult to quantify the economic contributions of the domestic workplace.").

Because it isn't considered to "build capital," domestic work has been persistently undervalued by mainstream economic theory.[123]

Another influential theoretical model is *Human Resource Management* (HRM), which takes a "*unitarist*"[124] view of the employment relationship as a long-term "partnership" between employers and employees who share certain interests.[125] The HRM model assumes that these interests may be harmonized simply by implementing the appropriate workplace policies and procedures.[126] To be sure, in the rhetoric over regulation, the shared interests of domestic workers and their employers (for example, safe workplaces and the well-being of a child or parent) are frequently overlooked, discounted, or ignored. However, the focus on the more formal approach of human resource policies is incongruous in the context of domestic work, since domestic employers often lack the time, resources, or experience necessary to develop a series of policies and practices that would apply, in most cases, to one employee only.

Alternatively, the pluralist model frames the employment relationship as a negotiated exchange between stakeholders in a democratic society.[127] Under this paradigm, the economic interests of employers and employees are both shared (such as the continued success of business) and competing (such as individual wages), and workers are entitled to human rights.[128] The pluralist model comes closest to capturing the particular dynamics of the domestic employment relationship. Strains of pluralism resound in the rationale for regulating domestic work, which includes the need to establish a level playing field so that domestic workers can assert basic

123. *See* Peggie R. Smith, *Regulating Paid Household Work: Class, Gender, Race, and Agendas of Reform*, 48 AM. U. L. REV. 851, 899 (1999) ("Situated within the family sphere and outside the purview of capital, paid household labor, similar to unpaid household labor, was understood to involve the creation of simple use-values, i.e., those values that the employing family consumed immediately and thus were thought never to enrich capital.")

124. So called for its emphasis on the unity of interests among employees and their employers.

125. *See* David R. Hannah & Roderick D. Iverson, *Employment Relationships in Context: Implications for Policy and Practice*, *in* THE EMPLOYMENT RELATIONSHIP: EXAMINING PSYCHOLOGICAL AND CONTEXTUAL PERSPECTIVES 332, 337–343 (Jacqueline Coyle-Shapiro et al. eds., 2004) (explaining the relationship of inducements and contributions to the employer-employee dynamic).

126. Budd, *supra* note 121, at 480.

127. *Id.*

128. *Id. See generally* Bruce E. Kaufman, *The Social Welfare Objectives and Ethical Principles of Industrial Relations*, *in* THE ETHICS OF HUMAN RESOURCES AND INDUSTRIAL RELATIONS 23-59 (John W. Budd & James G. Scoville eds., 2005).

human rights of safety, security, and dignity.[129] However, the pluralist model does not adequately capture the unique conflicts and common ground that lie at the heart of the domestic employment relationship. While employees in a traditional firm may share an interest with their employer in the ongoing operation of the business, the failure of that business has economic, rather than emotional, ramifications. Categorizing domestic employers and employees as "stakeholders" in the pluralist tradition, presumes that each has a different stake in the same game. Yet, in the context of domestic employment, the privacy concerns of employers and the emotional bonds of employees distinguish the relationship from those found in most other industries, and the pluralist model is insufficient in explaining those differences.

Lastly, *Critical Industrial Relations* theory views the employment relationship as fundamentally unequal, and inseparable from systemic inequalities throughout the socio-political-economic system.[130] Under the critical model, the interests of employers and employees are inherently, interminably, in conflict.[131] This model best captures the perspective of some in the domestic work movement regarding the nature of labor relations more broadly, but the vision of Domestic Workers United and other labor advocacy organizations diverges on the question of whether common ground is reachable by domestic workers and those for whom they care.[132]

As set forth in Table 1, each of the models (other than contract theory) considers the need for regulation of the employment relationship differently. Mainstream economic ("egoist") theory, which assumes perfectly competitive markets and self-interested workers and employers, considers workplace regulation necessary only in exceptional circumstances.[133] In the context of domestic work, this approach ignores factors that impel domestic workers and employ-

129. Domestic Workers' Bill of Rights, ch. 481, §1, 2010 N.Y. Sess. Laws 1315 (McKinney).

130. *See* Samuel Bowles & Herbert Gintis, *Contested Exchange: New Microfoundations for the Political Economy of Capitalism*, 18 POL. & SOC. 165, 177–82 (1990); *see also* Paul Thompson & Kirsty Newsome, *Labor Process Theory, Work, and the Employment Relation*ship, *in* THEORETICAL PERSPECTIVES ON WORK AND THE EMPLOYMENT RELATIONSHIP 133, 135–36, 147 (Bruce E. Kaufman ed., 2004).

131. Budd, *supra* note 121, at 480; *see also* Kaufman, *supra* note 113, at 12 ("The essence of the employer-employee relationship under capitalism, from a Marxist perspective, is domination, control and exploitation of labor in order to provide profit so firms can further accumulate capital.").

132. *See* Telephone Interview with Rachel McCullough, *supra* note 48 (explaining that while DWU educated the domestic workers themselves, JFREJ's focus has been on outreach to employers).

133. Budd, *supra* note 121, at 480.

Models of the Employment Relationship and Government Regulation				
Model	**View of Labor**	**View of Labor Markets**	**View of Employee-Employer Objectives**	**Resulting View of Government Regulation**
Egoist	A commodity; a rational self-interested economic agent.	Perfectly competitive.	Emphasis is on self-interest; exchanges occur when self-interest align.	Minimal. Fix market failures only when regulation does not do more harm than good.
Unitarist	A psychological being.	Imperfectly competitive.	Emphasis is on shared employer-employee interests; alignment occurs with effective human resource policies.	Low. Promote cooperation and prevent destructive competition.
Pluralist	An economic and psychological being; a democratic citizen with rights.	Imperfectly competitive.	Emphasis is on a mixture of shared and conflicting interests.	Essential. Establish safety nets and equalize bargaining power to balance efficiency, equity, and voice.
Critical	An economic and psychological being; a citizen with democratic rights.	Imperfectly competitive; part of broader, unequal institutional structure.	Emphasis is on inherent conflicts of interest; power differentials lead to exploitation.	Mixed. Important for protecting employees. Inadequate because of systemic imbalances inherent in capitalism.

Table 1

ers to act contrary to their self-interest (such as immigration status, emotional bonds, language barriers, and lack of alternative care options). The unitarist model emphasizes workplace-specific human resource policies over more expansive regulation.[134] Under that

134. *Id.*

model, regulation is needed only to thwart "destructive competition" (for example, immigration reform, minimum wage, and safety standards), and to cover those few employers that do not understand the value of cooperation.[135] The pluralist approach considers regulation essential to protecting basic democratic rights, by establishing the conditions necessary for workers to negotiate fair terms and conditions of their employment.[136] This model would support more detailed, expansive, and holistic regulation of domestic work, and is designed to correct the imbalance of power experienced by domestic workers in one-on-one negotiations. Lastly, while critical models take a view similar to that taken by the pluralists that government regulation is necessary, critical scholars do not believe that regulation itself is capable of fully mitigating the pervasive socio-political-economic inequalities in the domestic work industry.[137]

III.
DOMESTIC EMPLOYMENT MODEL

A "Domestic Employment" model that acknowledges the differences inherent in domestic work, without allowing those differences to perpetuate a second-tier status of rights, could help shape regulation of the industry. For example, the close bonds formed between a nanny or elder caregiver and the family members for whom she cares should be seen as a core *asset* of the employment relationship, not as a justification for its continued exclusion from workplace safety laws. Likewise, the privacy concerns of domestic employers should not be wholly discounted when considering ways of enforcing domestic workplace rights within the confines of an individual's home.

Integrating this Domestic Employment model into the existing table of theoretical frameworks (see Table 2, *infra*), the Domestic Employment model's view of labor and labor markets is substantially similar to that of the pluralist and critical paradigms. Domestic workers possess fundamental human and democratic rights, and the labor markets in which they work are not perfectly competitive. Broader, unequal institutional structures continue to serve as barri-

135. *Id.*

136. *Id.* at 482; *see also* Bruce E. Kaufman, *The Social Welfare Objectives and Ethical Principles of Industrial Relations*, *in* THE ETHICS OF HUMAN RESOURCES AND INDUSTRIAL RELATIONS 23, 43–56 (John W. Budd & James G. Scoville eds., 2005).

137. Budd, *supra* note 121, at 482; Kaufman, *supra* note 116, at 12 ("The essence of the employer-employee relationship under capitalism, from a Marxist perspective, is domination, control and exploitation of labor in order to provide profit so firms can further accumulate capital.").

ers to domestic workforce in attaining and asserting equal rights—including, but not limited to, immigration laws, racism, unequal access to education, a dismantling social safety net, and gender-based discrimination.

The objectives of domestic employers and employees are a mix of shared and conflicting interests. Domestic employers, most of whom are paying domestic workers out of the wages they earn in their primary occupations, have an interest in keeping the cost of care down. But they also have an interest in ensuring the quality of care provided by the domestic worker to their child, their parent, or their home. That interest frequently overlaps with the interest of domestic workers, many of whom have developed emotional connections with the children or parents of their employers in providing quality care. But the domestic worker, who often has her own childcare, elder care, and other economic and non-economic needs, certainly has an interest in maximizing her compensation. The intersection of these interests provides opportunities for creating a more transformative employment relationship, and one that can be regulated in two ways: with industry regulations that establish a "floor" of workplace standards, and with regulations that incentivize employers and employees to negotiate terms and conditions that exceed those standards.

Domestic Employment Relationship and Government Regulation				
Model	**View of Labor**	**View of Labor Markets**	**View of Employee-Employer Objectives**	**Resulting View of Government Regulation**
Domestic Employment	An economic and psychological being; a citizen with democratic rights.	Imperfectly competitive; part of broader, unequal institutional structure.	Mixture of shared and conflicting interests; power differentials lead to exploitation; opportunities for transformative relationship through shared interests.	Essential. Establish core industry standards, collective bargaining rights, co-regulatory regimes. Larger issues of immigration, institutional discrimination.

Table 2

A. The Domestic Employment Relationship

In framing the domestic employment relationship, it is necessary to recognize the differences in the nature of domestic work (for example, privacy concerns, emotional bonds, and isolation) without allowing those distinctions to perpetuate the industry's "second-tier" status and regulatory void. This new model recognizes opportunities for transformative organizing around shared interests (including the well-being of a child or elderly person, inequities experienced by working women generally, and the shrinking social safety net). In some ways, the lack of an existing regulatory regime in the domestic work industry (the result of enduring racism, sexism, and anti-immigrant animus) offers workers and high-road employers an opportunity to restructure their relationship.

The campaign to pass the Domestic Workers Bill of Rights revealed opportunities that had previously been viewed only as limitations. DWU led a campaign that mobilized many different communities of people, based on a more expansive and inclusive sense of "self-interest," which acknowledged the interdependence created by the domestic employment relationship. As Ai-jen Poo, writing on behalf of Domestic Workers United explains, "We learned that the historical assumption on which a great deal of organizing models are based—that we need to build our campaigns based on people's material self-interest—is not the whole story."[138]

One natural intersection of interests lies in the well being of the child or elderly person cared for by both the domestic worker and employer. As one domestic employer explains, "A lot of [working parents] want to make sure that their children are well taken care of and that the person caring for their child is happy, feels valued, feels well paid. That has a direct impact on how she cares for my daughter."[139] While "happy employees" are considered by HRM and unitarist theories as engines of productive workplaces, in the case of domestic work, the employer's self-interest in ensuring that the domestic workers he or she employs is difficult, if not impossible, to quantify using standard economic metrics.

Another intersection of interests in regulating the home as workplace is safety. While the type of home-based inspections traditionally under workplace health and safety laws such as OSHA have traditionally been dismissed because of the complicated privacy issues that arise when the workplace is in a private home, the Domes-

138. Poo, *supra* note 6 at 9.

139. *Laura Flanders*, *supra* note 9 (Interview with Dara Silverman, Jews for Racial and Economic Justice).

tic Employment model takes into account benefits incurred by employers as a result of such regulation. For example, standards regarding the toxicity of cleaning products or lifting requirements benefit the employee (by reducing the employee's exposure to toxic chemicals, and reducing the chance of injury from falls or strain), as well as employer (by reducing the employer's exposure to toxic chemicals, and reducing the chance of injury to possessions from being dropped). When the home is the workplace, safety standards benefit everyone in the home, including domestic employers and their families.[140]

As partners in the effort to transform the domestic work industry, high-road domestic employers have had to "explain to themselves, their neighbors, and their community why peace, justice, dignity, and respect in one's home [is] in their interest."[141] Likewise, a broader definition of "self-interest" helped to motivate and embolden domestic workers to organize the campaign for the New York Domestic Workers' Bill of Rights. Explains DWU organizer Aijen Poo, "They didn't only talk about bigger paychecks or days off for themselves. [T]hey talked about their mothers and their grandmothers who had done this work before, and they talked about their children for whom they wanted the opportunity to choose different futures."[142]

In regulating the domestic employment relationship, these broader, more inclusive self- and mutual-interests must be a part of a framework that undergirds a structure of core workplace standards, dignity, and respect. This approach provides a voice for domestic workers, both individually in the workplace, and collectively within the industry. Most importantly, it includes methods of enforcing standards that are responsive to the particular realities of domestic work.

B. Regulating the Domestic Employment Relationship

The Domestic Employment model relies upon a tripartite regulatory framework: employee voice, core industry standards, and pragmatic enforcement. Each leg is interdependent. The direct participation of domestic workers (their "voice") is critical to the development of meaningful, relevant, and enforceable industry

140. *See* Poo, *supra* note 6, at 11 (explaining that "[s]tandards benefit everyone" was one of the DWU campaign slogans in support of the Domestic Workers' Bill of Rights to highlight the "interconnectedness" between all people); *see also* Smith, *supra* note 14, at 1843.

141. Telephone Interview with Rachel McCullough, *supra* note 48.

142. Poo, *supra* note 6, at 10.

standards. These regulatory priorities are not unique to the Domestic Employment model, and are central to other pluralistic and critical models as well. However, the unique nature of domestic work, and the relationships that such work engenders, require a new approach to achieving each of these regulatory goals. Pluralist models bring us closer to an understanding of how the domestic work industry should be regulated than do neoliberal economic models that rely entirely on free market assumptions, or do contract theories that ignore fundamental inequalities of bargaining power between domestic employers and employees. However, such models do not go far enough in balancing the rights of domestic workers with the practical realities and privacy issues that arise when the workplace is a private home, nor do they recognize the complicated set of emotions between workers and employers, and workers and employers' loved ones. The Domestic Employment model further refines the pluralist stakeholder approach and, in its assessment of employee voice, core industry standards, and pragmatic enforcement, it incorporates many of the structural considerations embedded in the critical models.

1. Employee Voice and Collective Representation

Promoting and protecting the right of domestic workers to negotiate for better working conditions is a necessary and feasible means of raising standards across the domestic work industry.[143] While collective bargaining is hard to imagine in the domestic work industry in which isolated workers negotiate individually with employers, it is not unprecedented. Workers in a range of similarly isolated industries have surmounted these obstacles, including residential building superintendents, musicians, writers, and home healthcare workers.[144] The recent successes of Domestic Workers United dispel the myth that isolation and intimidation are insurmountable obstacles to organizing domestic workers.[145] Describing

143. DWU REPORT, *supra* note 75, at 10–11 (outlining a four-step plan for pursuing collective bargaining rights on behalf of domestic workers).

144. *Id.* at 13.

145. For example, unions across the country have used innovative approaches to overcome obstacles to organizing child-care workers. AFSCME, SEIU, the United Auto Workers, and The Association of Community Organization for Reform Now (ACORN) have successfully organized low-wage child-care providers in New York, Illinois, Massachusetts, Seattle, Pennsylvania, and Los Angeles. Paul F. Brooks, *New Turf for Organizing: Family Child Care Providers*, 29 LAB. STUD. J. 45, 45–48 (2005). For a detailed account of the historical exclusion of domestic workers from mainstream labor unions, see DOROTHY SUE COBBLE, THE OTHER WOMEN'S MOVEMENT: WORKPLACE JUSTICE AND SOCIAL RIGHTS IN MODERN AMERICA

these obstacles in the context of the New York campaign, one organizer explained, "There were potentially many employer-families who could be moved. And there were certainly many workers who could be organized."[146] Through these campaigns, domestic workers, employers, and their allies have developed an innovative, transformative organizing model aimed at "improving domestic workers' whole lives as workers, parents, caregivers, and community members."[147] Any system of collective representation for domestic workers would require cross-workplace employee organizations and, for purposes of collective bargaining, multi-employer units.[148] Indeed, DWU has recommended the development of collaborative, community-based mechanisms for enforcing standards within the industry.[149]

Contemplating a system of collective representation in the absence of NLRA sanctioned bargaining is not unique to the domestic work industry. Professor Cynthia Estlund, the Catherine A. Rein Professor of Law at New York University School of Law, has developed a model of "co-regulation" that seeks to improve and enforce

(2004) and Smith, *supra* note 14, at 1853–1855 (noting the exclusion of domestic service from developing labor standards). Smith further observes that in the context of home health care, workers often care for clients who rely on public funds for health care, giving rise to two models of third party arrangements, traditional agency-based care or consumer-directed care:

> Under the agency-based care model, a home-care agency hires, trains supervises, and assigns workers to provide publicly funded services to eligible clients. In such a scenario, the agency qualifies as the worker's employer. The second approach to delivering publicly funded home care is based on a consumer-directed care model that delegates to the client some responsibility for recruiting, hiring, training, and supervising the worker. Because clients often exercise considerable control over workers under this model, the clients may qualify as employers. Under the consumer-directed care model, however, third-parties—including public agencies that may administer publicly funded home-care services—may qualify as "joint employer" such that the law regards both the individual client and the agency as the workers' employers.

Smith, *supra* note 14, at 1862–1863.

146. Poo, *supra* note 6, at 13.

147. Whitefield, Alvarez & Emrani, *supra* note 97, at 28. The formation of the "Excluded Worker Congress" at the 2010 U.S. Social Forum reflects a growing movement to organize all "excluded workers," those who by law or policy are denied the right to organize and other fundamental labor rights, including (but not limited to) farm workers, guest workers, day laborers, domestic workers, and workers in Right to Work states. *See* Excluded Workers Congress, Unity of Dignity: Expanding the Right to Organize to Win Human Rights at Work 3 (2010), http://www.excludedworkerscongress.org/images/stories/documents/EWC_rpt_final4.pdf.

148. *See generally* Whitefield, Alvarez & Emrani, *supra* note 97, at 37–47.

149. *Id.* at 39.

workplace standards even in the absence of formal union representation.[150] Under a system of co-regulation, employee voice in workplace decisions is supported by internal employee committees and protected through independent, external monitors. The Domestic Employment model builds upon this framework by acknowledging that high road domestic employers can, and should, play a slightly different role in "co-regulation," because the domestic employment relationship is unlike most other employment relationships. Estlund's framework contemplates "internal employee committees" that would balance presumed employer authority structures, including management hierarchies and human resource departments. It also envisions a workplace that is larger than that of most domestic workers, for whom employee voice must potentially be fostered and protected at two levels: at the level of the individual relationship between the domestic employee and employer, and at a collective level in multi-employer standard setting and negotiations. Moreover, it is difficult to determine where high-road employers might fit in the co-regulatory model. They are neither employee insiders, nor truly independent outside monitors. The Domestic Employment model reconciles the co-regulatory model with the unique qualities of domestic work.

A critical step toward developing an effective system of co-regulation in the domestic work industry is the extension of collective bargaining rights to domestic workers. Recent state-based attacks on the collective bargaining rights of public sector workers suggest that an extension of those rights to domestic workers, in the current political climate, would require a tough battle. However, even in the absence of formal collective bargaining rights, domestic workers and "high road" employers—those who strive to meet or exceed industry standards—can pursue a system of workplace governance. Multi-employer domestic worker committees, perhaps organized geographically, can work together with domestic employers to hash out collective standards. By convening meetings of domestic employers and workers in particular neighborhoods, both groups have the opportunity to engage in conversations about wages and benefits that can, on an individual basis, feel awkward or imbalanced. Workers can talk more freely about the need for paid sick

150. CYNTHIA ESTLUND, REGOVERNING THE WORKPLACE: FROM SELF-REGULATION TO CO-REGULATION 21–23 (2010). Prof. Estlund's model emerges in light of the "bleak prospect" for "reviving and dramatically extending" the New Deal models of collective bargaining and standard setting in the U.S., and recognizes that the employer-driven move toward self-regulation has coincided with the drastic decline in unionization. *Id.* at 21–22.

leave or higher wages, and employers can share concerns about privacy and the quality of child or elder care, in an environment less fraught with fear or trepidation. Through these discussions, domestic workers and employers can establish standards—5 days paid sick leave, for example—based on a commitment to mutual respect, rather than based on word-of-mouth speculation. Enforcing those agreed upon standards is a separate, but equally important, component of the Domestic Employment model.

Until recently, the voices of domestic workers in setting standards and holding employers accountable to those standards has been largely piece-meal. Domestic workers are best positioned to determine how the work should be done, and how standards could be enforced. And yet, for all the reasons discussed above, domestic workers have found it difficult to assert their voices, either individually in negotiation with employers, or collectively to establish industry-wide standards. *Shalom Bayit*, a project of the New York City-based Jews for Racial and Economic Justice (JFREJ), has developed campaigns aimed at radically transforming that dynamic.[151] The campaign aims to bring together progressive domestic employers and the broader Jewish community to support domestic workers' rights, drawing on traditional Jewish values and progressive unionism. JFREJ and DWU have proposed neighborhood-based standard-setting negotiations among organizations of domestic workers and high road employers. The Domestic Employment model acknowledges the role that high road employers, like those organized by JFREJ, can play in helping to transform this industry.

Because the domestic employment relationship has been marked, historically, by a stark imbalance of power, domestic working conditions would improve with the building of power by domestic workers. Nonetheless, the work of JFREJ and DWU suggest that, at least in the case of high road employers, better working conditions may also be attained through more effective, open communication between domestic employees and employers about certain workplace issues. By legitimizing employer concerns surrounding privacy, autonomy (especially as relates to discipline and termination), and emotional attachments, the Domestic Employment model is capable of bringing together domestic employers and employees not as purely adversarial parties or co-equal "stakeholders," but as two groups of people whose unique concerns set them apart

151. *About the Shalom Bayit: Justice for Domestic Workers Campaign*, JEWS FOR ECONOMIC & SOCIAL JUSTICE, http://www.jfrej.org/shalom-bayit-campaign (last visited Feb. 7, 2012). "Shalom Bayit" means "peace in the home," in Hebrew. REUBEN ALCALAY, THE COMPLETE ENGLISH-HEBREW DICTIONARY 1776, 2674 (1959).

from most other employer-employee pairings. The Domestic Employment model contemplates the inclusion of high road employers in collaborative, community-based conversations, negotiations, and public awareness campaigns, in order to strengthen the voices of domestic workers individually and collectively.

2. Core Industry Standards

Given the persistent exclusion of domestic work from local, state, and federal labor and employment laws, any regulation of the industry must include a statutory floor of rights and protections, such as minimum wages and overtime, safety standards, paid sick leave, and protections against discrimination and harassment. Professor Estlund uses the term "responsive regulation," to describe a regulatory approach that incentivizes employers to comply with or even exceed standards, while sanctioning abusive employers.[152] While much of the litigation and press surrounding domestic work has focused on so-called "low-road employers" who take advantage of the isolation and vulnerability of their domestic employees, less attention has been paid to the role of high road domestic employers.[153] Indeed, several organizations of domestic employers were among the most vocal supporters of the New York Domestic Workers' Bill of Rights, and can help establish core industry standards that protect the rights of domestic workers, while taking into account the unique concerns (such as privacy) of domestic employers.[154] The Domestic Employment model builds upon Estlund's co-regulatory framework, and sets even higher expectations for high road employers to not only comply with or exceed standards, but indeed, to help establish those standards in the first place.

Through organizing a network of domestic employers, JFREJ is helping to accomplish this leg of the Domestic Employment model; its members have established best practices across the industry and

152. ESTLUND, *supra* note 150, at 217.

153. Moreover, formal trade associations of domestic employers do not exist to protect their rights; in other industries, these types of organized associations can help educate their members about applicable laws. NEW YORK STATE DEP'T OF LABOR, REPORT ON OUTREACH EFFORTS FOR DOMESTIC WORKERS LEGISLATION 1 (2010), *available at* http://www.labor.ny.gov/legal/laws/pdf/domestic-workers/report-to-governor-outreach.pdf. For an example of a call for a high road approach in the care sector industries like child care, see generally Nancy Folbre, *Demanding Quality: Worker/Consumer Coalitions and "High Road" Strategies in the Care Sector*, 34 POL. & SOC'Y 11 (2006).

154. POO, *supra* note 6, at 13–16.

serve as role models within their broader communities.[155] JFREJ has developed an outreach and education program through which domestic employers encourage their peers to make concrete improvements in their employment practices by first taking stock of the standards they have set, and then taking "one step up" (for example, by providing five paid holidays instead of three, or by entering into a written contract rather than verbal, ad hoc, agreements).[156] On the national level, "Hand in Hand" is an employers' association that is helping domestic employers understand the New York Domestic Workers' Bill of Rights, and that advocates for the passage of similar legislation in other states.[157] Through employer questionnaires, surveys, handbooks, and personal conversations, Hand in Hand hopes to formalize and "professionalize" the domestic employment relationship.[158]

Members of JFREJ's "Employers for Justice" network are also expected to speak out publicly when egregious cases of violence and abuse are revealed in the industry. In that way, JFREJ helps to set and enforce these core industry standards, by speaking out to say, "There is another way. I am an employer, and I am an employer for justice."[159] Certainly employers across a broad spectrum of in-

155. *About the Shalom Bayit: Justice for Domestic Workers Campaign*, *supra* note 151.

156. *See* Jews for Economic & Social Justice & Hand in Hand, *Take One Step Up with Hand in Hand*, HAND IN HAND: THE DOMESTIC EMPLOYERS ASSOCIATION, *available at* http://domesticemployers.org/wp-content/uploads/2011/04/2010_One_Step_Up_Form.pdf (last visited July 27, 2011).

157. *See Who We are*, HAND IN HAND: THE DOMESTIC EMPLOYERS ASSOCIATION, http://domesticemployers.org/who-we-are/ (last visited Jan. 31, 2012).

158. A questionnaire developed by Hand in Hand includes questions for domestic employers like, "Have you clearly defined your employee's responsibilities in the form of a contract or written agreement, so that both parties understand their obligations and responsibilities?" Jews for Economic & Social Justice & Hand in Hand, *supra* note 156. Additional "one step up" questions follow, such as, "Do you have a defined schedule for evaluation that includes a review of employer and employee expectations and experience and a plan of action to respond to any concerns?" The questions themselves appear to set market standards, with questions covering "fair wages" (e.g., severance and the living wage of $15–$18), benefits (e.g., insurance coverage and two weeks' paid vacation), and advocacy (e.g., conversations about workers' rights that have led to change in employment practices and communicating with legislators about these rights). *See* Jews For Economic & Social Justice & Hand in Hand, *Take One Step Up with Hand in Hand*, HAND IN HAND: THE DOMESTIC EMPLOYERS ASSOCIATION, *available at* http://domesticemployers.org/wp-content/uploads/2011/04/2010_One_Step_Up_Form.pdf (last visited July 27, 2011).

159. *See generally Justice for Domestic Workers*, JEWS FOR ECONOMIC & SOCIAL JUSTICE, *available at* http://www.jfrej.org/shalom-bayit-campaign (last visited Feb. 11, 2012).

dustries believe in fair working conditions, access to justice, and other tenets of corporate social responsibility. What is unique about the Domestic Employment model, however, is that it requires domestic workers be leaders and active participants in the process of establishing core industry standards, setting those standards through a combination of statutory minimums and collective negotiations with domestic employers.

3. Pragmatic Enforcement

Enforcing core standards in the domestic work industry will require the same level of engaged, direct participation by domestic workers, as well as a reflexive, creative, and vigorous monitoring system. The nature of the workplace as the private home, discussed in greater detail above, makes the traditional means of enforcing statutory standards and labor agreements (pickets, strikes,[160] and unannounced inspections[161]) difficult, if not impossible.

At minimum, any regulation of the domestic work relationship must include independent domestic worker organizations as part of the enforcement framework, as described in both the "voice" and "collective standards" sections, *supra*. Just as with standard setting, such organizations would ideally be permitted by law to engage in collective bargaining. Yet even in the absence of formal collective bargaining rights, such organizations can play a vital enforcement role.[162] By educating domestic workers and employers about indus-

160. In *State v. Cooper*, the Minnesota Supreme Court heard the case of a domestic worker, arrested and charged with disorderly conduct for picketing in front of his former employer's home to protest his recent termination. 285 N.W. 903, 904 (Minn. 1939). In that case, the court held that the chauffer was not protected by state labor law protecting workplace picketing, even though the home was his place of employment, because "the home is a sacred place for people to go and be quiet and at rest and not bothered with the turmoil of industry[.]" *Id.* at 905 (internal quotation marks and citation omitted). It is "a sanctuary of the individual and should not be interfered with by industrial disputes." *Id.* at 904–905 (internal quotation marks and citation omitted). For a further discussion of *State v. Cooper*, see Peggie R. Smith, *Organizing the Unorganizable: Private Paid Household Workers and Approaches to Employee Representation*, 79 N.C. L. REV. 45, 103–109 (2000).

161. Student writer Terri Nilliasca aptly points out that the protections afforded to the "private home" are not race-neutral. Nilliasca, *supra* note 44, at 381. In fact, the private homes (whether owned or rented) of poor people of color are frequently, and disproportionately, subject to public inspection, either through police searches or in connection with welfare benefits. *Id.* at 392.

162. *See* ESTLUND, *supra* note 150, at 148–49 (arguing for a new form of employee representation beyond formal union structures). Prof. Estlund warns that without assurance that courts and regulators are capable (or willing) to distinguish real from cosmetic compliance, co-regulation can devolve into a "thinly disguised form of deregulation." *Id.* at 17.

try standards and legal protections, helping workers to report violations of such standards, and uniting an inherently decentralized and dispersed group of workers, independent domestic worker organizations can help domestic workers build the collective strength and individual confidence necessary to enforce and advance domestic workplace standards. Rooted in communities of low-wage, immigrant workers, many domestic worker cooperatives, collectives, and worker centers have already helped to enforce evolving industry standards through formal and informal mediation and litigation, and by providing domestic workers with leadership training and organizing opportunities.[163]

For example, members of the San Francisco-based *La Colectiva* ("Organized Labor for an Organized Home") receive $70 for a minimum of three hours of labor, and $15 for each additional hour.[164] The women meet weekly to develop strategies for improving workplace safety (such as through the use of non-toxic cleaning products and ergonomics), and they develop advocacy, communication, and leadership skills through a set of classes on English, computer training, and workplace rights.[165] In order to establish and enforce mutual expectations and industry standards, *La Colectiva* developed an employer's handbook.[166] Designed using an aesthetic likely to appeal to the progressive, Bay Area professional milieu, the "conscientious cleaning service on a mission" helps guide conversations between domestic workers and employers about their relationship.[167] Artsy illustrations of appliances, furniture, and cleaning supplies, in both English and Spanish, are to be used in setting cleaning priorities.[168] Explicit pricing structures and policies on non-toxic cleaners are also included.[169]

Similar "eco-friendly" housecleaning co-ops have emerged in other cities,[170] including New York, where the Brooklyn-based

163. *See generally* Janice Fine, *Worker Centers: Organizing Communities at the Edge of the Dream*, 50 N.Y.L. SCH. L. REV. 417 (2005) (documenting the efforts of various worker centers).

164. LA COLECTIVA, EMPLOYER HANDBOOK 6, *available at* http://lacolectivasf.org/brochure.pdf.

165. Preeti Shekar, *Day Labor Program Unites Politics and Services Urban Habitat*, 14 RACE, POVERTY & THE ENV'T 42, 42–43 (2007), *available at* http://www.urbanhabitat.org/node/1183.

166. LA COLECTIVA, *supra* note 164.

167. *Id.* at 10.

168. *Id.* at 3–7.

169. *Id.* at 6–7.

170. Other such co-ops include UNITY Housecleaners Cooperative in Long Island, N.Y., and Las Senoras of St. Mary in Staten Island, N.Y. Vanessa Bransburg, *The Center for Family Life: Tackling Poverty and Social Isolation in Brooklyn with Worker*

Center for Family Life supports the *Si Se Puede!* (Yes We Can!) Cleaning Cooperative, as well as child-care, interior painting, and elder care co-ops.[171] Co-op members operate as individual businesses, and have the flexibility to work as much or as little as they would like.[172] By organizing together, however, the women have been able to increase their wages from an average of $7 to $8 an hour, to an average of $20 an hour, as a result of several interwoven factors.[173] First, by marketing themselves as a co-op, the members were able to tap into a market of middle-class, progressive clients that are sympathetic to the goals of the immigrant-run co-op.[174] By developing a standard contract, members are able to confidently assert a set of rates and standards (subject, in some cases, to negotiation).[175] Third, the women are able to enforce those standards, often informally, by refusing to work for clients who refuse to comply; and the success of the co-op's marketing tactics gives the members greater autonomy in turning down work, given the more consistent stream of new business.[176] In creating the *Si Se Puede!* Co-op, organizers looked to Women's Action to Gain Economic Security (WAGES), another San Francisco Bay Area organization, which "incubat[es]" worker owned "green cleaning co-ops."[177] For over five years, WAGES has partnered with Seventh Generation, a manufacturer of "eco-friendly cleaning products," which has garnered resources and expertise for co-op members, and public acclaim for Seventh Generation's commitment to social equity.[178]

Cooperatives, 2 GRASSROOTS ECON. ORGANIZING 13 (2011), *available at* http://geo.coop/node/636. The *Si Se Puede*! Co-op has been in existence for five years, and provides cleaning services to over 1,000 households and offices; members of the coop have more than doubled their hourly wage, from $7-$8/hour at their previous jobs, to $20/hour for work they secure through co-op referrals. *Id.* at 14.

171. Members of the Beyond Care Childcare Cooperative use a contract adapted from a model developed by Domestic Workers United. *Id.* at 15. The contract clearly sets out work duties, vacation days, sick days, overtime, cancellation fees, and an acknowledgment that the "[c]lient understands that labor laws, regardless of race, gender, immigration status, age, sexual orientation or religion, protect members." *Id.*

172. Telephone Interview with Vanessa Bransburg, Cooperative Coordinator, Center for Family Life (July 19, 2011).

173. *Id.*

174. *Id.*

175. *See id.* ("[The contract] is standardized, and from that the member has the confidence to say 'I'm not going to accept work that doesn't meet these standards.'").

176. *Id.*

177. Hilary Abell, *WAGES Model and the Value of Partnerships*, 2 GRASSROOTS ECON. ORGANIZING 27 (2011), *available at* http://geo.coop/node/635.

178. *Id.*

Under the Domestic Employment model, co-ops and collectives can function as an important, alternative vehicle for enforcing workplace standards in the largely unregulated and undervalued domestic work industry. Such organizations help domestic workers reduce the effects of isolation, inconsistent schedules, and one-on-one negotiations. By pooling resources, domestic workers can develop marketing strategies and strategic partnerships (for example, Seventh Generation, UC Berkeley Center on Occupational Health), needed to sustain the co-op ability to enforce such standards over the long haul. However, developing a co-op, collective, or worker center requires a tremendous amount of up-front resources. And, even when successful, the scale is limited, at least initially. Participation in a co-op or collective is often transformational for the individual domestic worker-members, but a transformation of the domestic work industry requires solutions that can be applied industry-wide. The Domestic Employment model, therefore, views high road employers and local government as key partners in developing effective enforcement strategies that do not necessarily require the resources and organizational prowess required for a successful co-op or collective.

IV.
THE DOMESTIC WORKER BILL OF RIGHTS: A RELATIONSHIP BASED ANALYSIS

As a general premise, the stated objectives of any particular piece of workplace legislation should be aligned with the drafters' conception of the particular employment relationship being regulated (for example, a pluralist's vision for workplace regulation should, ideally, be reflected in legislation that is inclusive of various stakeholder perspectives). Such legislation should, in turn, be implemented in such a way that the regulation achieves those stated goals.[179] A sponsor's draft of the Domestic Workers' Bill of Rights paints a pluralistic picture of a vulnerable workforce that needs legislation to protect basic human rights, explaining that "many domestic workers fall through the cracks of U.S. government," and

179. Budd, *supra* note 121, at 478. Advocating for greater transparency by public policymakers about the assumptions upon which workplace laws are based, Stephen F. Befort and John W. Budd have proposed a framework to analyze individual pieces of workplace legislation. *See generally id.* This framework requires three separate inquiries: (1) an analytical undertaking that asks whether the regulation is coherent and consistent with the objectives of that legislation; (2) a normative assessment of the goals of the legislation; and (3) a pragmatic appraisal that asks whether the legislation will work as intended. *See generally id.*

must thus shoulder the burden of ensuring that employers comply with basic workplace standards and human rights.[180] And because many domestic workers are "isolated, exploited and psychologically abused by their employers," legislation is necessary to protect employees working in homes.[181] Using the pluralist model of employment relationship, the New York Domestic Workers' Bill of Rights helps improve working conditions for domestic workers, while considering different stakeholders.[182] In order to determine whether this pioneering legislation is likely to be effective given the conception of the employment relationship articulated by the Domestic Employment model, we must evaluate whether the law (i) fosters employee voice, (ii) establishes core industry standards, and (iii) contains pragmatic means of enforcement.

Does the regulation foster genuine employee voice? As described above, the Domestic Employment model builds upon the pluralistic "co-regulatory" approach of Estlund's model, which requires that domestic workers have the "ability to provide meaningful input[,] individually and collectively," into workplace decisions, industry standards, and enforcement.[183] Although the voices of domestic workers (and high road employers) were critical to the passage of the New York Domestic Workers' Bill of Rights,[184] the version of the law that was finally enacted did not include the right to collectively bargain with employers.[185] This omission denied to domestic workers one of the most direct means of raising their voices, shaping

180. S. 2311E, 2009 Leg., 232nd Sess. (N.Y. 2009), *available at* http://open.nysenate.gov/legislation/bill/S2311E-2009.

181. *Id.*

182. Domestic Workers' Bill of Rights, ch. 481, §1, 2010 N.Y. Sess. Laws 1315 (McKinney).

183. Budd, *supra* note 121, at 484.

184. Kicking off the campaign, DWU held a "Having Your Say" Convention, which brought together hundreds of domestic workers, from over a dozen countries. Poo, *supra* note 6, at 5. Despite the lack of a common language, the domestic workers "found a common voice as they shared their experiences of laboring without respect or basic labor standards." *Id.* The convention's emcee, a domestic worker from the Caribbean, opened the program by stating, "Ladies, we are making history here today. You have a voice, and together we are going places." *Id.*

185. In passing the law without that provision, the New York State Assembly directed the New York State Department of Labor to prepare a report on the feasibility of collective bargaining in the domestic work industry. Domestic Workers' Bill of Rights § 10. The New York Labor Law itself declares it as the state's public policy "to encourage the practice and procedure of collective bargaining, and to protect employees in the exercise of full freedom of association, self-organization and designation of representatives of their own choosing for the purposes of collective bargain, or other mutual aid and protection, free from the interference, restraint or coercion of their employers." N.Y. Lab. Law § 700 (McKinney 2011).

their working conditions, and transforming their employment relationship.[186] The completed feasibility study commissioned by the state legislature did find that collective bargaining was a feasible option for domestic workers and cited several industries in which unionization is not hampered by barriers such as multiple workplaces and employers, or isolated workers.[187] To date, however, the legislature has not taken the next step in amending the state labor laws to permit domestic workers to form unions.

Formal union representation is not, however, the only means of fostering genuine employee voice and collective participation. Co-regulation, for example, envisions a combination of internal employee committees and external independent monitors.[188] In its report "Feasibility of Domestic Worker Collective Bargaining," the New York State Department of Labor proposed several alternatives for collective representation, including voluntary recognition, single person collective bargaining units, hiring halls and cooperatives, mandatory employment contracts, and continued outreach to workers and employers.[189] However, none of these proposals explicitly includes the voices of high road employers, a critical component of the Domestic Employment model.

While the New York State Domestic Workers' Bill of Rights did not go far enough in fostering employee voice or collective participation, the various frameworks proposed in the state's Feasibility Study do have the potential to do so. DWU and JFREJ have proposed a series of voluntary, collective negotiations among domestic workers and neighborhood-based employer associations.[190] This in-

186. Section 8 of the proposed legislation would have eliminated the domestic worker carve-out to the New York State Labor Law. S. 2311D, 2009 Leg, 232d Sess. (N.Y. 2009*), available at* http://open.nysenate.gov/legislation/bill/S2311D-2009. Closing this gap would require amending Section 701(3) of the State Labor Relations Act (SLRA) to eliminate the following language from the definition of "employees":

> [B]ut shall not include any individual employed by his parent or spouse or in the domestic service of and directly employed by his parent or spouse in the domestic service of and directly employed, controlled and paid by any person in his home, any individual whose primary responsibility is the care of a minor child or children and/or someone who lives in the home of a person for the purpose of service as a companion to a sick, convalescing or elderly person or any individual employed only for the duration of a labor dispute[.]

N.Y. LAB. LAW § 701(3) (McKinney 2011).

187. FEASIBILITY STUDY, *supra* note 9, at 3.

188. *See* ESTLUND, *supra* note 150, at 162 (arguing that both employee voice and the support of regulators and enforcers are critical).

189. FEASIBILITY STUDY, *supra* note 9, at 17–28.

190. Telephone Interview with Rachel McCullough, *supra* note 48.

itiative falls squarely within the proposals offered by the Feasibility Study, which acknowledges that "although it is unclear which approach to organizing will emerge as the most effective strategy, [extending collective bargaining rights to domestic workers] provides the opportunity for domestic workers and employers to begin the process of exploring these various approaches in an effort to ultimately achieve more harmonious labor relations. *Both domestic workers and their employers must determine their best form of organization*."[191]

Does the regulation establish core industry standards? Reversing a legacy of underpayment, either legally by virtue of a lack of minimum standards for domestic work, or illegally through rampant "wage-theft" in the domestic work industry, the New York Domestic Workers' Bill of Rights is one step toward a more efficient, productive, and mutually beneficial domestic work industry. The core gains achieved by the Bill of Rights were the expansion of minimum wage coverage to "companions" employed solely by households;[192] a higher rate of overtime pay for live-in domestic workers and live-out companions;[193] mandatory time off of work and paid vacation (one unpaid day of rest every calendar week, and three paid days of vacation after one year of employment);[194] protection from sexual harassment and other forms of workplace discrimination;[195] and temporary disability benefits for part-time and full-time domestic workers.[196]

191. FEASIBILITY STUDY, *supra* note 7, at 29 (emphasis added).

192. Domestic Workers' Bill of Rights, ch. 481, §2, 2010 N.Y. Sess. Laws 1315 (McKinney). One criticism of the New York Domestic Workers' Bill of Rights, however, has been that the legislation preserved the "companionship services" exclusion in the newly created definition of "domestic worker," further perpetuating the exclusion of certain domestic workers from protection under this, and any future, domestic work regulation. *See, e.g.*, Nilliasca, *supra* note 46, at 399–400.

193. Prior to the passage of the New York Domestic Workers' Bill of Rights, "live-in" domestic workers were entitled to overtime pay, but only at one and a half times the state minimum wage (not their negotiated hourly wage, which may have been higher), and only after they worked forty-four hours in a week. DWU REPORT, *supra* note 75, at 2. Today, those workers are entitled to one and a half times their regular rate of pay. *Id.*

194. Domestic Workers' Bill of Rights §7.

195. The New York Domestic Workers' Bill of Rights amends the New York Human Rights Law's definition of employer to expand coverage to domestic workers, and prohibit sexual harassment and harassment based on race, religion, or national origin aimed specifically at "domestic workers." Domestic Workers' Bill of Rights §§ 2–3.

196. A person is considered a domestic worker under the law if they work in another person's home to care for a child; serve as a companion for a sick, convalescing, or elderly person; do housekeeping; or perform any other domestic service purpose. N.Y. LABOR LAW § 2.16 (McKinney 2011); N.Y. EXEC. LAW § 296-b (Mc-

The final version of the bill did not include other proposed benefits, such as paid sick days, paid personal days, advance notice of termination, and severance pay.[197] While it is true that other private-sector workers are not guaranteed those benefits by statute, most other workers have important means of obtaining them. Workers who enjoy those benefits often do so as the result of collective bargaining—either as direct beneficiaries of union contracts, or indirectly from market pressures that such agreements place on non-union competitors. However, as discussed in greater detail above, the New York State Legislature chose not to extend collective bargaining rights to domestic workers, opting instead to defer the decision pending a "feasibility study."[198]

Overall the New York State Domestic Workers' Bill of Rights did create a set of minimum standards and legal protections for domestic workers. Including elder caregivers in the state's minimum wage framework will result in tangible economic gains, especially because the New York state minimum wage has historically been higher than the floor set at the federal level. Higher overtime rates, mandatory days of rest (paid and unpaid), and protection from harassment are all very basic standards enjoyed by most other workers, yet constitute an unprecedented legislative victory for domestic workers. What is lacking under the Domestic Employment model, however, is an opportunity for domestic employers to engage in structured dialogue with domestic workers and their advocates about the issues of privacy, autonomy, and emotional attachment in order to find common ground on those issues.

Does the regulation include pragmatic strategies for enforcement? The New York Domestic Workers' Bill of Rights requires the Commissioner of Labor to report to the Governor and Speaker of the Assembly about how to best provide easily accessible educational and informational material on workplace laws covering domestic workers.[199] Soon after the bill's passage, the Commissioner of Labor met with advocacy groups, community-based organizations, employee and employer representatives, and state agencies to discuss with

Kinney 2011). However, a person performing companionship services as an employee of an agency without additional domestic work such as cleaning services is not subject to the overtime and day of rest rules. NEW YORK STATE DEP'T OF LABOR, FACT SHEET: LABOR RIGHTS AND PROTECTIONS FOR DOMESTIC WORKERS, *available at* http://www.labor.ny.gov/legal/laws/pdf/domestic-workers/facts-for-domestic-workers.pdf .

197. DWU REPORT, *supra* note 75, at 3.

198. *Id.*

199. Domestic Workers' Bill of Rights §10.

stakeholders the optimal means for reaching out to this isolated group of workers.[200]

Because neither domestic workers nor their employers are likely to have much experience with wage, employment, tax, or other laws, the state has developed materials in "plain language" that "avoid jargon or overly technical language."[201] Posting notice of applicable laws in "visible location[s] in the workplace" might seem awkward in a personal home, so the state created a "small card that will not intrude on a home setting yet will include required information."[202] Included in the planned efforts for reaching domestic employers are "[p]arenting blogs and list-serves, such as *Park Slope Parents*, *Urban Baby*," and through the Chamber of Commerce (with the assumption that many of its members are also employers of domestic workers).[203]

Enforcing workplace standards in private homes will remain a challenge for the New York State Department of Labor, but continued collaboration with domestic workers and their advocates is critical to ensuring such efforts are successful. Through such collaboration, the Labor Department may develop pragmatic enforcement methods that could be replicated in other jurisdictions for the enforcement of workplace protections within many different industries of lower-wage workers.

CONCLUSION

"We are enacting a culture shift in which progressive employers in a given community will collectively take action and say, 'No, in our homes we employ dignity; this isn't a factory, this isn't going to be like other employment.'"[204]

As we've seen, the domestic workplace and the relationship between domestic workers and their employers is unique. But what sets domestic work apart—physical isolation, private home workplaces, emotional attachment, and a legacy of race and gender discrimination—should not perpetuate the disparate treatment of domestic work by regulators, employers, and other industry stakeholders. Instead, regulation of the domestic industry should be developed in partnership with domestic workers, and should not only

200. NEW YORK STATE DEP'T OF LABOR, *supra* note 153, at 1.

201. *Id.*

202. *Id.* at 3. Domestic Workers' Rights Poster, P713-Spanish-English, N.Y. STATE DEP'T OF LABOR, *available at* http://www.labor.ny.gov/formsdocs/wp/P713-english-spanish.pdf.

203. NEW YORK STATE DEP'T OF LABOR, *supra* note 153, at 4.

204. Telephone Interview with Rachel McCullough, *supra* note 46.

take into account those differences that make regulation and enforcement difficult, but should also acknowledge the unique features of the domestic employment relationship. The Domestic Employment model should be used to assess legislative and policy proposals to regulate domestic work, in order to ensure that such regulation includes core industry standards and effective systems of enforcement (including collective bargaining, domestic worker coops, neighborhood-based bargaining, external public monitoring, and enforcement).

The need for care—childcare, elder care, house cleaning—cuts across race, class, and gender lines. The strong—if complicated—bonds that form between workers, employers, and those receiving care, and the issues of structural inequality, gender-discrimination, and a disappearing social safety net that affect both domestic workers and their employers, reflect opportunities for improving, if not transforming, the domestic employment relationship.

NAVIGATING IMPLIED WAIVER OF THE ATTORNEY-CLIENT PRIVILEGE AFTER ADOPTION OF FEDERAL RULE 502 OF THE FEDERAL RULES OF EVIDENCE†

JAMES P. MCLOUGHLIN, JR.
NEIL T. BLOOMFIELD
RENEE D.K. MILLER
*TONYA L. MERCER**

* Mr. McLoughlin and Mr. Bloomfield are Members, Ms. Miller is an Associate and Ms. Mercer is Counsel in Moore & Van Allen, PLLC's Litigation and Government Investigation practice groups in Charlotte, North Carolina.

INTRODUCTION

> If we intend to serve the interests of justice by encouraging consultation with counsel free from the apprehension of disclosure, then courts must work to delineate the scope of the privilege in ways that are predictable and certain. "An uncertain privilege—or one which purports to be certain, but rests in widely varying applications by the courts—is little better than no privilege."[1]

Assume that a publicly held New York corporation with nationwide business retains counsel for its latest filing with the United States Securities and Exchange Commission. As part of this representation, counsel advises the client on what information should and should not be included in the filing about the company's retention of independent counsel to investigate suspicions that a Virginia subsidiary bribed public officials. Based on the disclosures made in that filing, the corporation is sued in district courts in the Second and Fourth Circuits. The court in the Second Circuit, as expected, finds that the communications between the corporation and its counsel related to the filing are privileged. However, under Fourth Circuit precedent, none of the communications between the corporation and its counsel are privileged. The court in the Fourth Circuit orders production of all of the information and advice relating to the disclosure about the investigation, including information counsel advised the company not to disclose. This result shocks the client and even the attorneys, who practice in jurisdictions with more generous interpretations of the attorney-client privilege. The impact of Rule 502 of the Federal Rules of Evidence on this dichotomy—and the dichotomy's effect on Rule 502—is the subject of this Article.

The circumstances under which the attorney-client privilege is deemed waived by voluntary action, whether by express waiver or the less-understood "implied waiver," are volatile and uncertain. Indeed, as recently as 2003, the First Circuit could say that "[l]ike most courts, this court has yet to develop a jurisprudence clarifying the scope of such implied waivers."[2] Many attorneys and their clients do not appreciate the disparate approaches that various federal courts take to implied waiver, or the uncertainty over privilege that results.

1. *In re* Lott, 424 F.3d 446, 450 (6th Cir. 2005) (quoting Rhone-Poulenc Rorer Inc. v. Home Indem. Co., 32 F.3d 851, 863 (3d Cir. 1994)).

2. *In re* Keeper of the Records, 348 F.3d 16, 23 (1st Cir. 2003).

Federal Rule of Evidence 502, which became effective September 19, 2008, was enacted to resolve what the drafters characterized as "long-standing disputes in the courts" about the effect of disclosures of privileged information.[3] The Advisory Committee reported: "The rule seeks to provide a predictable, uniform set of standards under which parties can determine the consequences of a disclosure of a communication or information covered by the attorney-client privilege or work product protection."[4]

Our discussion focuses on Rule 502(a), which addresses "implied waiver." Implied waiver may occur in two ways. First, implied waiver may occur when the privilege-holder partially discloses confidential communications with counsel. Alternatively, implied waiver may occur when the privilege-holder puts an otherwise privileged communication into controversy—for example, by asserting the defense of attorney-client privilege. In these situations, the law implies waiver of the privilege as to other, undisclosed, communications.

The circumstances in which waiver is implied, and the scope of that waiver, vary by circuit. Generally speaking, circuits have adopted one of two views of waiver of the attorney-client privilege, referred to here as the Classic View and the Modern View. These views diverge because they are grounded upon different evaluations of the benefits and costs of the attorney-client privilege to the judicial system and society. Although it is an oversimplification, the Classic View begrudgingly acknowledges the necessity of the attorney-client privilege to allow the attorney-client relationship to function, but emphasizes the countervailing loss of transparency and the obstruction of the truth-finding process. As a result, under the Classic View, the attorney-client privilege is a necessary evil, and courts holding to the Classic View are rigorous in the application of the requirements of the privilege. The failure of a client to comply with the strictures of the classic requirements of the privilege results in a complete or broad loss of the privilege. In contrast, the Modern View emphasizes the attorney-client privilege as a useful and beneficial component of the judicial system and broader society, and so takes a more generous view of the privilege and a much more limited view of the circumstances under which the privilege is lost and the scope of that loss.

3. ADVISORY COMM. ON EVIDENCE RULES, REPORT OF THE ADVISORY COMMITTEE ON EVIDENCE RULES (MAY 15, 2007), *available at* http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/Reports/2007-05-Committee_Report-Evidence.pdf.

4. FED. R. EVID. 502 advisory committee's note.

Rule 502(a) provides that disclosure of privileged information or communications in a federal proceeding, or to a federal office or agency, waives the attorney-client privilege or work product protection as to the information disclosed. However, the waiver extends to undisclosed communications or information only if: "(1) the waiver is intentional; (2) the disclosed and undisclosed communications or information concern the same subject matter; and (3) they ought in fairness to be considered together."[5] Rule 502's use of the malleable concept of "fairness" adopts the Modern View, but leaves courts with tremendous leeway to define the context and scope of the waiver without resolving conflicts between the Modern and Classic Views.

By leaving room for both the Classic and Modern Views, Rule 502 thus provides less certainty than first appears. This Article examines the mechanics and consequences of that uncertainty. We will examine the conflict between the Classic View and the Modern View with respect to implied waiver, including the scope of implied waiver and the contexts in which implied waivers occur, particularly in investigations. We will also examine situations when a portion of privileged information is published with the intent to keep a separate portion of the information privileged. Along the way we will cover the essentials of the attorney-client privilege under the Classic and Modern Views and summarize the development of Rule 502.

Our analysis suggests that though Rule 502 may seem to constrain the Classic View's expansive waiver doctrine, Classic View courts may nonetheless stand their ground based on the conclusion that the balancing of "fairness" considerations required by Rule 502 can support a broad waiver.[6] Further, because the rule does not address privilege generally, or waiver by acts other than disclosure, Classic View courts may find other ways to preserve their broader sense of waiver.[7] For example, before a communication can be sub-

5. FED. R. EVID. 502(a).

6. As recently as 2005, the Federal Circuit, sitting *en banc*, declared that broad, subject matter waiver, which is an essential feature of the Classic View, is a widely applied standard founded on fairness principles. *In re* Seagate Tech., LLC, 497 F.3d 1360, 1372 (Fed. Cir. 2007) (en banc) ("'The widely applied standard for determining the scope of a waiver . . . is that the waiver applies to all other communications relating to the same subject matter.' . . . This broad scope is grounded in principles of fairness and serves to prevent a party from simultaneously using the privilege as both a sword and a shield; that is, it prevents the inequitable result of a party disclosing favorable communications while asserting the privilege as to less favorable ones." (quoting Fort James Corp. v. Solo Cup Corp., 412 F.3d 1340, 1349 (Fed. Cir. 2005))).

7. S. REP. NO. 110-264, at 3 (2008).

ject to implied waiver, it must be privileged in the first place. When analyzing implied waiver, especially in the context of investigations, this threshold issue is often overlooked. A Classic View court may resolve what first appears to be a question of implied waiver by deciding the communication is not privileged in the first instance because the client anticipated disclosure of the communication not made in confidence.

This inter-circuit conflict is likely to be exacerbated by Rule 502's grant to private parties of the power to negotiate agreements regarding waiver and preservation of privilege. The Advisory Committee on Evidence Rules ("Advisory Committee") expressed the view that "[p]arties should be able to contract around common-law waiver rules by entering into confidentiality agreements . . . ."[8] While such contracts cannot bind third parties absent a court order, Rules 502(e) and (d) have the potential to fundamentally alter the way that privilege operates. If Modern View courts are willing to approve the panoply of possible agreements while Classic View courts balk, the conflict between Classic View and Modern View courts may produce even more confusion among the circuits as to privilege than we see today. This confusion may not be resolved for some time, as rulings on attorney-client privilege are rarely appealed and the standard for appellate review is typically deferential.

Rule 502's inability to fully resolve the tension between the Classic and Modern Views is increasingly important as corporations face heightened regulatory scrutiny not experienced since the New Deal gave birth to the era of modern regulation. Part I of this Article provides an overview of the Classic and Modern Views of privilege and implied waiver, as well as of the choice of law principles that dictate when and how the Classic View or the Modern View controls. Part II explores Rule 502 and its influence upon the dichotomy between the Classic and Modern Views. Part III describes the Classic and Modern Views of the application of the attorney-client privilege to information communicated with the intent that the information will be published. Finally, Part IV examines implied waiver of attorney-client privilege, and Part V describes the topical and temporal scope of waiver.

8. ADVISORY COMM. ON EVIDENCE RULES, REPORT OF THE ADVISORY COMMITTEE ON EVIDENCE RULES (May 16, 2006), *available at* http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/Reports/EV05-2006.pdf. *See* Dowd v. Calabrese, 101 F.R.D. 427, 439–40 (D.D.C. 1984) (finding no waiver where the parties stipulated in advance that certain testimony at a deposition "would not be deemed to constitute any waiver of the attorney-client or work product privileges").

I.
THE BATTLE BETWEEN THE MODERN AND CLASSIC VIEWS OF PRIVILEGE

A. The Classic View of the Privilege and Implied Waiver

It cannot be said too often that the attorney-client privilege protects communications, not facts.[9] To be protected from disclosure, the communications must meet the requirements of privilege.[10] The policy imperative underpinning the Classic View is that a client must be free to disclose damaging confidences to his lawyer in order to get advice to solve the problem, but that the privilege frustrates the "search for truth,"[11] so any assertion of the attorney-client privilege must be construed narrowly and compliance with its requirements must be judged strictly.[12] This policy imperative is not mere lip service to transparency; it defines the Classic View and drives the outcomes as to virtually all facets of privilege. The United States Court of Appeals for the Fourth Circuit has been an unflagging advocate of the Classic View.

The Fourth Circuit's decision in *United States v. Jones* describes the elements of the attorney-client privilege under the Classic View:

> (1) [T]he asserted holder of the privilege is or sought to become a client;
> (2) the person to whom the communication was made must be a lawyer or his subordinate, and acting as or on behalf of a lawyer;
> (3) the communication must relate to a fact of which the attorney was informed by his client in confidence and without the presence of "strangers;"
> (4) for the primary purpose of securing legal advice or services;
> (5) not for the purpose of committing a crime or tort; and
> (6) the privilege has not been waived by the client.[13]

9. *See, e.g.*, Upjohn Co. v. United States, 449 U.S. 383, 395 (1981); P. & B. Marina, Ltd. v. Logrande, 136 F.R.D. 50, 56 (E.D.N.Y. 1991) ("[It is] the well established rule that only the communication, not the underlying facts, are privileged." (quoting Sneider v. Kimberly-Clark Corp., 91 F.R.D. 1, 4 (N.D. Ill. 1980))).

10. *See, e.g.*, United States v. Jones, 696 F.2d 1069, 1072 (4th Cir. 1982).

11. Burroughs Wellcome Co. v. Barr Labs., Inc., 143 F.R.D. 611, 615 (E.D.N.C. 1992).

12. *See, e.g.*, Weil v. Investment/Indicators, Research & Mgmt., Inc., 647 F.2d 18, 24 (9th Cir. 1981) ("Because it impedes full and free discovery of the truth, the attorney-client privilege is strictly construed."); Permian Corp. v. United States, 665 F.2d 1214, 1219 (D.C. Cir. 1981).

13. *See, e.g.*, *Jones*, 696 F.2d at 1072.

Under this classic definition, the function of the privilege is limited to the protection of the confidentiality of a client's confidences shared with counsel. Thus, a communication between the client and counsel that reflects counsel's advice is privileged only if it would reveal a confidential communication from the client to counsel made for the purpose of obtaining legal advice.[14] For instance, a memorandum from counsel that gives legal advice but does not include a client confidence is not privileged because it fails the third and fourth prongs of the classic test.[15]

The Classic View takes a sweeping approach to the question of waiver.[16] In virtually any context, either intentionally disclosing privileged information, or intentionally putting a privileged communication into controversy, results in a complete waiver of the privilege of all communications regarding that subject matter.[17]

14. *Jones*, 696 F.2d at 1072–73 (questioning whether communications were privileged where the attorneys were retained for the primarily commercial purpose of preparing tax opinions for a private offering memorandum); North Carolina Elec. Membership Corp. v. Carolina Power & Light Co., 110 F.R.D. 511, 514 (M.D.N.C. 1986) ("In order for the privilege to apply, the attorney receiving a communication must be acting as an attorney and not simply as a business advisor."); CHARLES ALAN WRIGHT & KENNETH W. GRAHAM, JR., FEDERAL PRACTICE AND PROCEDURE: EVIDENCE, § 5491 (1986).

15. *See, e.g.*, *In re* Grand Jury Proceedings, 727 F.2d 1352, 1358 (4th Cir. 1984) (finding that communications with legal counsel for the purpose of preparing a prospectus for publication were not privileged); United States v. Kelsey-Hayes Wheel Co., 15 F.R.D. 461, 465 (E.D. Mich. 1954) (holding privilege lost as to inadvertently produced documents when in-house counsel did not maintain a segregated file for privileged documents, thereby failing to take reasonable steps to ensure and maintain the confidentiality of privileged communications).

16. *E.g.*, *In re* Sealed Case, 877 F.2d 976, 980–81 (D.C. Cir. 1989) (inadvertent production of privileged documents in discovery resulted in subject matter waiver). The harshness of some precedents was used to support passage of Rule 502 in the Congress: "Outdated legal precedents from an earlier era continue to create uncertainty. There are precedents, for example, holding that an inadvertent disclosure of a single document or communication not only can waive the privilege as to that one item, but can result in a blanket waiver as to all information concerning the same subject. That can collapse a case." 154 CONG. REC. 18,016 (2008).

17. *E.g.*, Anderson v. Calderon, 232 F.3d 1053, 1099 (9th Cir. 2000) ("[T]he privilege is an evidentiary rule designed to prevent the forced disclosure in a judicial proceeding of certain confidential communications between a client and a lawyer. The rule does not apply where there is no forced disclosure of a confidential communication in a judicial proceeding. Thus, a court's authority to 'protect' the attorney-client privilege simply does not extend, at least absent some compelling circumstance, to non-compelled, voluntary, [disclosures], any more than it does to an after-dinner conversation. The attorney-client privilege is a rule of evidence.") (citations omitted).

B. The Modern View of the Privilege and Implied Waiver

A majority of federal circuits have come to a more lenient view of the privilege. These courts acknowledge that the privilege may frustrate the search for truth, but focus on the policy that for the client and counsel relationship to function, the client must be able to trust that all confidential communications are protected from disclosure.[18] This assessment of the privilege's impact on the judicial process is starkly different from that of the Classic View. Under the Modern View, "As the privilege serves the interests of justice, it is worthy of maximum legal protection."[19]

As described by the Second Circuit in *In re Grand Jury Subpoena Duces Tecum dated Sept. 15, 1983*, the generally accepted Modern View test is that the privilege attaches:

(1) Where legal advice of any kind is sought;
(2) from a professional legal advisor in his capacity as such;
(3) the communications relating to that purpose;
(4) made in confidence;
(5) by the client;
(6) are at his instance permanently protected;
(7) from disclosure by himself or by the legal advisor; and
(8) except the protection would be waived. . . .[20]

Under the Modern View, all confidential communications between the attorney and her client are protected by privilege. Modern View courts believe that the privilege cannot function unless the client can have confidence that the advice that she receives is as protected as the information that she provides in order to get that

18. *See* Swidler & Berlin v. United States, 524 U.S. 399, 409 (1998) (rejecting exception that would allow for disclosure of privileged communications in criminal cases after a client's death, because such an exception would introduce "substantial uncertainty into the privilege's application"); *see also In re* LTV Sec. Litig., 89 F.R.D. 595, 600 (N.D. Tex. 1981).

19. Rhone-Poulence Rorer v. Home Indem. Co., 32 F. 3d 851, 862 (3d Cir. 1994) (citing Haines v. Liggett Group Inc., 975 F.2d 81, 90 (3d Cir. 1992)).

20. *In re* Grand Jury Subpoena Duces Tecum Dated Sept. 15, 1983 (Marc Rich & Co. A.G. v. United States), 731 F.2d 1032 (2d Cir. 1984) (citations omitted).

advice.[21] The United States Court of Appeals for the Second Circuit has been a leader in developing the Modern View.[22]

One hallmark of the Modern View is that it takes a forgivingly pragmatic approach to waiver of the privilege,[23] abandoning ideological purity in favor of seeking "fairness" in the way that the privilege operates in litigation. The Modern View takes the approach that the implied waiver of the attorney-client privilege that results from intentional disclosure of privileged information should be no broader than fairness requires.[24]

Of course, the Classic View courts would argue that privilege, in fairness, cannot be both a sword and a shield, and so fairness requires complete disclosure. The results of these two differing approaches are sometimes termed implied "subject matter waiver" and implied "limited waiver." The waivers are "implied" because they are not the result of publication of the information at issue; rather, the waiver results from a determination by the court that some other act compels the waiver. Under a Classic View of subject

21. *See* Walsh v. Northrop Grumman Corp., 165 F.R.D. 16, 18 (E.D.N.Y. 1996); Bank Brussels Lambert v. Credit Lyonnais (Suisse) S.A., 160 F.R.D. 437, 441–42 (S.D.N.Y. 1995); *see also* Sprague v. Thorn Ams., Inc., 129 F.3d 1355, 1370 (10th Cir. 1997) (summarizing the difference in approach among courts as to advice from counsel that does not include a client confidence and affirming the circuit's acceptance of the broader view of the privilege). Some courts take a hybrid approach. For example, the Ninth Circuit takes a Classic View with respect to proof of the privilege and waiver "'[b]ecause it impedes full and free discovery of the truth, the attorney-client privilege is strictly construed.' '[T]he privilege stands in derogation of the public's "right to every man's evidence" and as "an obstacle to the investigation of the truth," [and] thus, . . . "[i]t ought to be strictly confined within the narrowest possible limits consistent with the logic of its principle."'" United States v. Ruehle, 583 F.3d 600, 607 (9th Cir. 2009) (citations omitted). However, the Ninth Circuit protects attorney to client communications. *Id.* ("'The attorney-client privilege protects confidential disclosures made by a client to an attorney in order to obtain legal advice . . . as well as an attorney's advice in response to such disclosures.'") (citing United States v. Bauer, 132 F.3d 504, 507 (9th Cir. 1997)).

22. *See, e.g.*, *Marc Rich & Co. A.G.*, 731 F.2d at 1036.

23. *Compare* James Julian, Inc. v. Raytheon Co., 93 F.R.D. 138, 142 (D. Del. 1982) (accepting that it was essential for the efficient operation of corporation to keep privileged documents in the general file to reduce the administrative burden, given needed access to non-lawyer executives working on that subject matter), *with* United States v. Kelsey-Hayes Wheel Co., 15 F.R.D. 461, 465 (E.D. Mich. 1954) (holding privilege lost as to inadvertently produced documents when in-house counsel did not maintain a segregated file for privileged documents, thereby failing to take reasonable steps to ensure and maintain the confidentiality of privileged communications).

24. *See, e.g.*, *In re* Keeper of the Records, 348 F.3d 16, 24 (1st Cir. 2003) (noting that implied waiver is "almost invariably premised on fairness concerns").

matter waiver, if some communications on a subject are disclosed, all communications on the entire subject lose the privilege by implied waiver.[25] Under a limited waiver analysis, the unpublished material may remain privileged—subject to the imprecise qualification that no unfairness to an adversary results.[26]

We will see that the Modern View defines "fairness" narrowly, as the avoidance of material disadvantage to an adversary in an official, contested proceeding. With this definition, in contrast to the Classic View, the Modern View considers the context of the waiver as essential to the analysis of fairness. Thus, because waiver outside of an official contested proceeding is deemed not to be harmful to an opposing party, disclosure of privileged information outside of such a proceeding does not result in waiver of the privilege protecting any information or communications other than those actually disclosed.

C. Choice of Law

To appreciate the uncertainty that may infect the attorney-client privilege requires a discussion of when and how the Classic View or the Modern View controls. Under Rule 501 of the Federal Rules of Evidence, privilege in federal criminal proceedings and in civil cases involving only federal claims is governed by the federal common law of privilege.[27] When an action involves both state and federal claims and the disputed evidence is relevant to both, the overwhelming majority of courts apply the federal common law of privilege, especially if federal law allows the evidence and state law excludes it.[28] However, if an action involves both federal and state

25. *E.g.*, United States v. Woodall, 438 F.2d 1317, 1324 (5th Cir. 1970) (en banc) ("[A] client's offer of his own or his attorney's testimony as to a specific communication constitutes a waiver as to all other communications on the same matter [because] 'the privilege of secret communication is intended only as an incidental means of defense, and not as an independent means of attack, and to use it in the latter character is to abandon it in the former.'" (quoting WIGMORE ON EVIDENCE § 2327 (1961))).

26. *See, e.g.*, Pritchard v. County of Erie, 546 F.3d 222, 227–28 (2d Cir. 2008).

27. FED. R. EVID. 501; *see also* Lego v. Stratos Lightwave, Inc., 224 F.R.D. 576, 578 (S.D.N.Y. 2004).

28. *See* FED. R. EVID. 501 (providing that in civil cases "with respect to an element of a claim or defense as to which State law supplies the rule of decision, the privilege of a witness [or] person . . . shall be determined in accordance with State law"); Guzman v. Mem'l Hermann Hosp. Sys., No. H-07-3973, 2009 U.S. Dist. LEXIS 13336, at *12 (S.D. Tex. Feb. 20, 2009) (listing cases); *see also* Virmani v. Novant Health Inc., 259 F.3d 284, 286 n.3 (4th Cir. 2001); Pearson v. Miller, 211 F.3d 57, 66 (3d Cir. 2000); Mem'l Hosp. for McHenry County v. Shadur, 664 F.2d 1058, 1061 n.3 (7th Cir. 1981); Cont'l Cas. Co. v. Under Armour, Inc., 537 F. Supp.

law claims, but the evidence is relevant only to the state law claims, the majority of federal courts apply state privilege law.[29]

When federal law applies, choice of law amongst federal circuits is often a misnomer for privilege because federal courts overwhelmingly apply the law of their home circuit, and most do so without discussion.[30] While choice of law can be a factor in transfer cases,[31] the decisions hold uniformly that a federal choice of law analysis is inappropriate between or among circuits because the "[f]ederal courts comprise a single system applying a single body of law, and no litigant has a right to have the interpretation of one federal court rather than that of another determine his case."[32] Of course, uniformity is a fiction, as circuits have adopted varying interpretations of attorney-client privilege.

2d 761, 765 n.3 (D. Md. 2008); Weiss *ex rel.* Estate of Weiss v. County of Chester, 231 F.R.D. 202, 205 n.19 (E.D. Pa. 2004) (listing cases).

29. *E.g.*, *Lego*, 224 F.R.D. at 578; *Guzman*, 2009 U.S. Dist. LEXIS 13336, at * 22.

30. *See, e.g.*, *In re* von Bulow, 828 F.2d 94 (2d Cir. 1987); Kingsway Fin. Servs., Inc. v. PricewaterhouseCoopers, LLP, No. 03 Civ. 556ORMBHBP, 2007 WL 1837133, at *2 (S.D.N.Y. June 27, 2007).

31. When a federal question case is transferred for the convenience of the parties and witnesses pursuant to 28 U.S.C. § 1404, the law of the transferor federal forum applies. McMasters v. United States, 260 F.3d 814, 819 (7th Cir. 2001); Newton v. Thomason, 22 F.3d 1455, 1459 (9th Cir. 1994); Tel-Phonic Servs., Inc. v. TBS Int'l, Inc., 975 F.2d 1134, 1138 (5th Cir. 1992). When a case is transferred pursuant to 28 U.S.C. § 1406 because it was brought in an improper venue, that improper venue choice cannot establish a choice of law, so the proper venue's law applies. Nelson v. Int'l Paint Co., 716 F.2d 640, 643–44 (9th Cir. 1983); Ellis v. Great Sw. Corp., 646 F.2d 1099, 1109–11 (5th Cir. 1981). The civil precedent suggests that only one circuit's law should govern privilege for all cases in the multidistrict litigation context. Highland Tank & Mfg. Co. v. PS Int'l, Inc., 246 F.R.D. 239, 243–44 (W.D. Pa. 2007); New York v. Microsoft Corp., No. 98-1233, 2002 WL 649492, at *1 (D.D.C. Apr. 8, 2002). When two or more cases are considered under multi-district litigation rules, the interpretation of the transferee court will most often be applied. *See, e.g.*, Menowitz v. Brown, 991 F.2d 36, 40 (2d Cir. 1993). However, some courts have found that the law of the transferor forum should apply—at least to the extent that federal law would borrow from an applicable state statute of limitations. Eckstein v. Balcor Film Investors, 8 F.3d 1121, 1127 (7th Cir. 1993); *In re* UMW Emp. Benefit Plans Litig., 854 F. Supp. 914, 921 (D.D.C. 1994). The decisions do not address the law of privilege, so what position might control in the multi-district litigation context is an open question.

32. *E.g.*, *In re* Ramaekers, 33 F. Supp. 2d 312, 315 (S.D.N.Y. 1999) (per curiam) (quoting *Menowitz*, 991 F.2d at 40) (holding that the First Amendment reporter's privilege of the circuit in which a subpoena was issued, and in which the court sat, governed a motion to compel compliance with the subpoena, not the law of the circuit in which the litigation was pending).

When state claims are at issue and state law applies, the state law of attorney-client privilege controls.[33] When state law applies, the federal court must determine which state's choice-of-law analysis applies, and, applying the selected choice-of-law analysis, determine which jurisdiction's privilege rules apply.[34] The court hearing the discovery dispute must apply the choice of law rules of the state in which it sits.[35]

Federal courts applying state choice of law rules may or may not yield predictable results. Applying New York's choice of law principles, federal courts in the Second Circuit have held the governing law is that of the jurisdiction which, "because of its relationship or contact with the occurrence or the parties, has the greatest concern with the specific issue raised in the litigation."[36] Under this rule, courts generally apply the law of the jurisdiction in which the assertedly privileged communications were made, which in most of the cases is also the jurisdiction in which the party that made the communications resides.[37] Courts have articulated the rationale for this rule as: "[T]he parties who made the communications expected that those communications would remain confidential under the law of that jurisdiction, and the state has an interest in furthering the policies behind the privilege at issue."[38]

However, the result is not always predictable. The Eastern District of North Carolina applied North Carolina choice of law rules to determine whether New York or North Carolina law should govern privilege in a dispute over the financing of a construction project.[39] The district court concluded that North Carolina courts would apply "the law of the State having the most significant rela-

33. *See, e.g.*, Limbach Co. v. Zurich Am. Ins. Co., 396 F.3d 358, 361–62 (4th Cir. 2005).

34. Lego v. Stratos Lightwave, Inc., 224 F.R.D. 576, 578 n.7 (S.D.N.Y. 2004); *Limbach Co.*, 396 F.3d at 361–62; Metric Constructors, Inc. v. Bank of Tokyo-Mitsubishi, Ltd., No. 5:97-CV-369BR1, 1998 U.S. Dist. LEXIS 18428, at *4 (E.D.N.C. Sept. 28, 1998).

35. *Lego*, 224 F.R.D. at 578 n.7; *Limbach Co.*, 396 F.3d at 361–62.

36. *Lego*, 224 F.R.D. at 578 (quoting Tartaglia v. Paul Revere Life Ins. Co., 948 F. Supp. 325, 326 (S.D.N.Y. 1996)).

37. *Id.* at 579.

38. *Id.* at 579 & n.10 (listing cases); *cf.* Chin v. Rogoff & Co., No. 05 Civ. 8360(NRB), 2008 WL 2073934, at *5 (S.D.N.Y. May 8, 2008) ("[S]ince the evidence will be introduced in New York, the dispute involves an attorney-client relationship with a New York law firm, and both parties cite to New York law, we hold that New York law governs the assertion of the privilege.").

39. *Metric Constructors, Inc.*, 1998 U.S. Dist. LEXIS 18428, at *5.

tionship to the occurrence giving rise to the action."[40] Counsel and clients were from New York and the privileged communications occurred there. Nonetheless, the court concluded that North Carolina law applied because the dispute concerned construction of a waste-to-energy project located in North Carolina.[41]

Thus it can be difficult to predict what law will govern disputes concerning privileged communications. Whether the court takes the Classic View or the Modern View can have drastic consequences on the privilege and the nature of any waiver. Rule 502 appears to have been intended, at least in part, to address this issue. As we will see, it is not always successful. The commentary to Rule 502(a), which addresses implied waiver, states that "[t]his subdivision does not alter the substantive law regarding when a party's strategic use in litigation of otherwise privileged information obliges that party to waive the privilege regarding other information concerning the same subject matter, so that the information being used can be fairly considered in context."[42] How the Classic View courts and the Modern View courts interpret this instruction may decide the impact that Rule 502 has on the debate between the two views.

The scope of Rule 502, and the ways in which the rule impacts the Classic and Modern Views (and vice versa), may not be addressed by the appellate courts for some time; and until Congress or the Supreme Court speaks in a clear voice, circuit court rulings are unlikely to be definitive. Rulings over attorney-client privilege are rarely appealed, and the standards of appellate review are typically deferential. For example, in the Second Circuit, determinations about the scope of waiver are reviewed under the abuse of discretion standard.[43] The Fourth Circuit reviews attorney-client privilege determinations using a two-fold standard: if the district court's ruling rests on findings of fact, review is for clear error. If the district court's decision rests on legal principles, the standard is *de novo* review.[44] Exploring the tension between the Classic and

40. *Id.* at *6 (quoting Andrew Jackson Sales v. Bi-Lo Stores, 314 S.E.2d 797, 799 (N.C. 1984)).

41. *Id.*

42. 154 CONG. REC. 18,016 (2008).

43. United States v. Doe (*In re* Grand Jury Proceedings), 219 F.3d 175, 182 (2d Cir. 2000).

44. Hawkins v. Stables, 148 F.3d 379, 382 (4th Cir. 1998); Better Gov't Bureau, Inc. v. McGraw (*In re* Allen), 106 F.3d 582, 601 (4th Cir. 1997). *But see* United States v. Ruehle, 583 F.3d 600, 606 (9th Cir. 2009) ("The district court's conclusion that statements are protected by an individual attorney-client privilege is 'a mixed question of law and fact which this court reviews independently and without deference to the district court.' . . . That is, whether the party has met the requirements

Modern Views and Rule 502 is therefore necessary because navigating these waters may be an uncertain enterprise for a long time.

II.
FEDERAL RULE OF EVIDENCE 502

A. The Rule

Into the battle between the Classic View and the Modern View, Congress inserted Rule 502. In pertinent part, Rule 502 states:

> Rule 502 Attorney-Client Privilege and Work Product; Limitations on Waiver
>
> The following provisions apply, in the circumstances set out, to disclosure of a communication or information covered by the attorney-client privilege or work-product protection.
>
> **a. Disclosure Made in a Federal Proceeding or to a Federal Office or Agency; Scope of a Waiver—**When the disclosure is made in a Federal proceeding or to a Federal office or agency and waives the attorney-client privilege or work-product protection, the waiver extends to an undisclosed communication or information in a Federal or State proceeding only if:
>
> 1. the waiver is intentional;
> 2. the disclosed and undisclosed communications or information concern the same subject matter; and
> 3. they ought in fairness to be considered together.
>
> ***
>
> **d. Controlling Effect of a Court Order—**A Federal court may order that the privilege or protection is not waived by disclosure connected with the litigation pending before the court—in which event the disclosure is also not a waiver in any other Federal or State proceeding.
>
> **e. Controlling Effect of a Party Agreement—**An agreement on the effect of disclosure in a Federal proceeding is binding only on the parties to the agreement, unless it is incorporated into a court order.
>
> ***

to establish the existence of the attorney-client privilege is reviewed de novo. . . . We also review de novo the district court's rulings on the scope of the attorney-client privilege. . . . Factual findings are reviewed for clear error.") (citations omitted).

B. The History of Rule 502

Rule 502 was a long time coming. The necessity of an additional rule that would address waiver of attorney-client privilege has long been the subject of debate, and an effort to enact such a rule was shelved in 2002.[45] At that time, the Advisory Committee agreed to continue its privilege project with the intent of providing a survey of the existing federal common law of privilege. Professor Kenneth S. Broun of the University of North Carolina School of Law and Professor Daniel J. Capra of Fordham Law School led that effort and would ultimately draft the proposed rule.[46]

In May of 2004, the Advisory Committee gave notice that it was proposing a rule to address the issue of inadvertent waiver of the attorney-client privilege, especially in complex litigation.[47] The effort was motivated primarily by the escalating costs and delay resulting from steps taken by litigants to prevent the inadvertent disclosure of privileged information in the age of electronic discovery.[48] Though inadvertent, such disclosures are nonetheless "voluntary," and some courts had ruled that such a disclosure resulted in subject matter waiver.[49] The Advisory Committee and legal commentators were troubled by the fact that, in order to prevent waiver

45. ADVISORY COMM. ON EVIDENCE RULES, MINUTES OF THE MEETING OF APRIL 29TH AND 30TH, 2001, *available at* www.uscourts.gov/uscourts/RulesAndPolicies/rules/Minutes/EV04-2004.pdf.

46. Kenneth S. Broun & Daniel J. Capra, *Getting Control of Waiver of Privilege in the Federal Courts: A Proposal for a Federal Rule of Evidence 502* 1 n.1 (Bepress Legal Series, Working Paper No. 1485, 2006).

47. *See Federal Rule of Evidence 502—Timeline*, FEDERAL EVIDENCE REVIEW, http://federalevidence.com/node/288 (last visited Oct. 5, 2011) (reporting that the Advisory Committee's subcommittee on privileges' survey of the existing federal law of the attorney-client privilege is near completion and reporting that the Advisory Committee is considering whether to propose a statute to cover the problem of "unintentional waiver of privileged information," particularly in complex litigation).

48. ADVISORY COMM. ON EVIDENCE RULES, *supra* note 3, at 2 ("Enormous expenses are put into document production in order to protect against inadvertent disclosure of privileged information, because the producing party risks a ruling that even a mistaken disclosure can result in a subject matter waiver."). The Chairman of the Committee on the Judiciary of the House of Representatives, then Representative F. James Sensenbrenner, Jr., had requested that the issue be addressed. Memorandum from James N. Ishida to the Standing Comm. on Federal Rules of Practice, Procedure and Evidence (Dec. 18, 2006) (on file with the Administrative Offices of the U.S. Courts).

49. ADVISORY COMM. ON EVIDENCE RULES, *supra* note 4, at 7 (rejecting the holding of *In re Sealed Case*, 877 F.2d 976 (D.C. Cir. 1989), which the Committee states held inadvertent disclosure of privileged communications in discovery automatically resulted in subject matter waiver).

in this scenario, litigants were forced, at great expense, to review virtually everything produced.

In April of 2005, a draft statute was proposed. The draft began with a general statement of the fairness doctrine and followed with exceptions to waiver.[50] The Advisory Committee minutes focus on the inadvertent disclosure problem. At the request of Representative Sensenbrenner, Chairman of the Committee on the Judiciary of the House of Representatives, the Advisory Committee went on to address waiver of attorney-client privilege when one is the subject of a government investigation. The Advisory Committee incorporated into the proposed Rule 502 the "selective waiver" doctrine, under which disclosure of privileged information to a government agency in the course of an investigation does not result in a waiver of the privilege as to the information disclosed or the subject matter against non-government persons or entities.[51] The selective waiver proposal, which had been requested by Representative Sensenbrenner, generated criticism in part on the ground that it would foster a coercive "culture of waiver."[52]

Advocates of a robust privilege feared that adoption of selective waiver would make it impossible to resist the "pernicious" damage to the attorney-client relationship because the government would push for disclosure in all cases using the argument that the target would suffer no harm.[53] Ultimately, the Advisory Committee de-

50. The draft read:

> (a) Waiver by disclosure in general.—A person waives an attorney-client privilege or work product protection if that person—or a predecessor while its holder—voluntarily discloses or consents to disclosure of any significant part of the privileged or protected information. The waiver extends to undisclosed information concerning the same subject matter if that undisclosed information ought in fairness to be considered with the disclosed information.
>
> (b) Exceptions in general.—A voluntary disclosure does not operate as a waiver if: (1) the disclosure is itself privileged or protected; (2) the disclosure is inadvertent and is made during discovery in federal or state litigation or administrative proceedings—and if the holder of the privilege or work product protection took reasonable precautions to prevent disclosure and took reasonably prompt measures, once the holder knew or should have known of the disclosure, to rectify the error, including (if applicable) following the procedures in Fed. R. Civ. P. 26(b)(5)(B); or (3) the disclosure is made to a federal, state, or local governmental agency during an investigation by that agency, and is limited to persons involved in the investigation.

Broun & Capra, *supra* note 47, at 50–51.

51. In short, it would have codified the decision in *Diversified Industries, Inc. v. Meredith*, 572 F.2d 596 (8th Cir. 1978). Broun & Capra, *supra* note 47, at 60.

52. ADVISORY COMM. ON EVIDENCE RULES, *supra* note 4, at 36–38.

53. ADVISORY COMM. ON EVIDENCE RULES, HEARING ON PROPOSED 502 (Apr. 24, 2006), *available at* http://federalevidence.com/pdf/2008/06-June/MiniConf_

cided not to include the selective waiver provision as part of Rule 502; instead, they included it as an option presented to the Standing Committee.[54] The Standing Committee adopted Rule 502 as proposed with the caveat that a "strong showing" in favor of selective waiver would have to be made during the public comment period, particularly by state attorneys general and other regulatory authorities, to incorporate the selective waiver provision. A proposed Rule 502 was approved by the Standing Committee in June 2006.[55] The rule, as proposed, did not incorporate selective waiver, leaving it as a separate report to Congress.[56] The debate over the selective waiver provision appears to have drowned out discussion of implied waiver, especially in the investigation context, perhaps leaving less of a record of the intent of the final rule than might otherwise have been available.

The Standing Committee sent the final proposed rule to Congress in June of 2007.[57] The Standing Committee's proposal took a decidedly Modern View approach to waiver.[58] Surveying the congressional materials is not as illuminating with regard to implied waiver as one might hope.[59] The vast majority of the advocacy with

April_2006.pdf (Comments of David Brodsky, member of the ABA Presidential Task Force on the Attorney Client Privilege); *see also* Liesa L. Richter, *Corporate Salvations Or Damnation? Proposed New Federal Legislation on Selective Waiver*, 76 FORDHAM L. REV. 129, 173 (2007).

54. *Minutes of June 22-23 Meeting*, COMM. ON RULES OF PRACTICE & PROCEDURE, http://www.uscourts.gov/uscourts/rulesandpolicies/rules/minutes/st06-2006-min.pdf (last visited Oct. 6, 2011).

55. *Id.* Professor Capra was the Reporter to the Conference Committee's Advisory Committee on Evidence Rules, and Judge Jerry E. Smith was its Chair.

56. The selective waiver provision—on which the Evidence Rules Committee had never voted affirmatively—was dropped from the Proposed Rule 502.

57. The public comment period ended in February of 2007.

58. "A subject matter waiver should be found only when privilege or work product has already been disclosed, and a further disclosure 'ought in fairness' to be required in order to protect against a misrepresentation that might arise from the previous disclosure." ADVISORY COMM. ON EVIDENCE RULES, REPORT OF THE ADVISORY COMMITTEE ON EVIDENCE RULES (MAY 16, 2006), *available at* http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/Reports/EV05-2006.pdf. The comments to the proposed Rule 502 can be interpreted to settle any dispute in favor of the Modern View: "[Rule 502] resolves some longstanding disputes in the courts about the effect of certain disclosures of communications or information protected by the attorney-client privilege or as work product—specifically those disputes involving inadvertent disclosure and subject matter waiver." *Id.*

59. Representative Jackson Lee of Texas, sponsor of the companion bill in the House, H.R. 6610, urged passage of the Senate Bill on the House floor. She assured the House that the rule would have "no negative impact on those lawyers representing defendants or those lawyers representing plaintiffs." 154 CONG. REC. 18,016 (2008).

respect to the proposed Rule 502 was addressed to the inadvertent disclosure provisions, about which there was a broad consensus.[60] For example, the Senate Report includes no discussion of any problem that the rule is intended to address except for inadvertent waiver, and the report's statement of purpose declares that the rule is intended to leave untouched the question of whether material is privileged in the first instance,[61] and "merely modifies the consequences of inadvertent disclosure once a privilege is found to exist."[62] Interpretation of Rule 502 to address inadvertent waiver is easily explained; the Senate Report and its disclaimer ignore the more far reaching provisions of the rule. Floor proceedings were a few short speeches in support that focused on inadvertent waiver.[63] Senate Bill 2450 became Rule 502 when it was passed by the House.

The limited Congressional commentary points to a preference for the Modern View, which it describes as the majority view.[64] Congress supplemented the Rules Committee's Explanatory Note with a Statement of Congressional Intent.[65] The Statement provides some guidance,[66] but it also appears to give courts leeway to continue to apply the Classic View, as the Statement disclaims any intent to affect the law of implied waiver—at least so far as concerns the con-

60. *Id.* at 18,017; *id.* (remarks of Rep. Jackson Lee) ("The new rule 502 reaffirms and reinforces the attorney-client privilege and work product protection by clarifying how they are affected by, and withstand, inadvertent disclosure in discovery."); *id.* (remarks of Rep. King) ("The bill only modifies the consequences of an inadvertent disclosure once a privilege exists.").

61. S. REP. NO. 110-264, at 1–3 (2008).

62. *Id.* at 3.

63. 154 CONG. REC. 18,016 (2008).

64. *Id.* at 18,016 (Rep. Jackson-Lee) ("The attorney-client privilege and work product protection are crucial to our legal system. They encourage businesses and individuals to obtain legal counsel when appropriate by protecting the confidentiality of communications between clients and their attorneys, and documents prepared by attorneys to assist their clients in litigation. In fact, this is the backbone, the infrastructure of civil and criminal litigation.").

65. *Id.* at 18,016–17.

66. "The standard set forth in this subdivision for determining whether a disclosure operates as a waiver of the privilege or protection is, as explained elsewhere in this Note, the majority rule in the federal courts. The majority rule has simply been distilled here into a standard designed to be predictable in its application. This distillation is not intended to foreclose notions of fairness from continuing to inform application of the standard in all aspects as appropriate in particular cases . . . ." *Id.* at 18,016. Rep. King of Iowa spoke in support of the bill, stating in part: "The content of the new rule includes the following provisions: If a waiver is found, it applies only to the information disclosed, unless a broader waiver is made necessary by the holder's intentional and misleading use of privileged or protected communications or information." *Id.* at 18,017.

cept that privilege cannot be used as both a sword and a shield: "This subdivision does not alter the substantive law regarding when a party's strategic use in litigation of otherwise privileged information obliges that party to waive the privilege regarding other information concerning the same subject matter, so that the information being used can be fairly considered in context."[67]

III.
INTENT, EXPECTATION, AND PRIVILEGE

Under both the Classic and the Modern Views, the attorney-client privilege never applies to information communicated with the intention that the information will be published.[68] For example, a public filing is never privileged, even if drafted by a lawyer. However, the application of this seemingly straightforward principle differs between the two views with dramatic consequences.[69] What we call the "publication dichotomy" becomes important for Rule 502 and implied waiver when a court considers the privileged status of information disclosed to the government. If the court rules that the intent to share some of the results with the government means that the communications were never privileged in the first instance, the court will not have the opportunity to explore the questions of implied waiver and the scope of such waiver under Rule 502. Rule 502's Commentary states it is not intended to change privilege law generally, meaning that there is a strong argument that this publication dichotomy is outside the scope of the rule.

The requirement that any communication from the client to counsel must be made with the intention that it will remain confidential is enforced strictly under the Classic View. Thus unless information is provided with the iron-clad intent that it will remain confidential, it is not privileged from the outset.

Under the Modern View, the approach is often a functional one. When the client has no expectation that the information will remain confidential, the result matches the Classic View. However, when the expectation is contingent—i.e., a client provides information with the expectation that what can be kept confidential will be, but what the attorney determines should be disclosed will be dis-

67. *Id.* at 18,016.

68. *See* United States v. Tellier, 255 F.2d 441, 447 (2d Cir. 1958); *see also In re* Grand Jury Proceedings, 727 F.2d 1352, 1355–56 (4th Cir. 1984).

69. United States v. Bergonzi, 216 F.R.D. 487, 493–94 (N.D. Cal. 2003) (investigation materials prepared with the anticipation some or all would be shared with the SEC are not privileged).

closed—the Modern View will protect the client's contingent expectation by preserving the privilege as to the information that is not disclosed.

A. The Classic View and the Evolution of the Modern View

A brief review of the relevant history is instructive. In *United States v. Tellier*,[70] decided in 1958, the Second Circuit took the Classic View. Mr. Tellier ran an underwriting firm engaged to raise funds through a public offering of debentures for an independent telephone and electric power system in Alaska. He employed counsel to handle SEC work necessary for the offering. When a third issuance of debentures was proposed by Tellier, the attorney warned that the debentures looked like a Ponzi scheme. Tellier and the attorney agreed that the attorney would put his warning in a letter in order to warn the co-venturers. Tellier stated that he had expected his attorney's warning letter to be forwarded "to the joint venturers for their enlightenment," but the attorney sent the letter only to Tellier, who did not publish the letter to anyone. When the company went bankrupt, Tellier and his co-venturers were indicted for securities fraud. Tellier's attorney testified at trial, over Tellier's objection, that the letter about the Ponzi scheme was a privileged communication. He relied in part on the fact that the warning letter was never published. The Second Circuit ruled that there was no privilege. The fact that the letter was never sent to the joint venturers (and that in all likelihood Tellier never truly intended that it be sent) was of no consequence because the letter was prepared with the professed plan of publication. The Second Circuit held:

> It is of the essence of the attorney-client privilege that it is limited to those communications which are intended to be confidential. 'The moment confidence ceases,' said Lord Eldon, 'privilege ceases.' . . . Thus it is well established that communications between an attorney and his client, though made privately, are not privileged if it was understood that the information communicated in the conversation was to be conveyed to others.[71]

Under the Classic View, drafts of any document to be published to third parties, and even drafts prepared with the intention that the final product *might* be published, are deemed to lack the intent to keep the communication confidential, so the drafts, in-

70. *Tellier*, 255 F.2d at 441.

71. *Id.* at 447 (citations omitted).

cluding notes on the drafts, are not privileged from the outset.[72] The future expectation of confidentiality is critical. Today the Classic View generates results that many corporate counsel find dangerously counter-intuitive. Drafts assumed protected from scrutiny are not. For example, in most instances privilege will not protect drafts of public securities filings or documents created in the course of preparing a report that will be turned over to government investigators, because none are prepared with an expectation that all of the information in the documents will be kept confidential.[73] The same is true for communications of information to prepare tax returns or patent applications.[74] Documents prepared for litigation are the exception.[75]

After *Tellier* in 1958, the Second Circuit cases exemplify the movement to the Modern View of future publication as the circuit started to move away from the convention that the attorney-client privilege does not protect communications of information intended to be published. In *Colton v. United States*, decided in 1962, the Second Circuit concluded that not all documents and communications exchanged between the client and attorney during preparation of tax returns lack an expectation of confidentiality.[76] The court recognized, and the government conceded, that privilege protects "confidential papers which were 'specifically prepared by the client for the purpose of consultation with his attorney' and any of the firm's memoranda and worksheets 'to the extent of any unpublished expression made by an attorney therein of confidences which had passed between him and his clients.'"[77]

72. *In re Grand Jury Proceedings*, 727 F.2d at 1356; *Tellier*, 255 F.2d at 447 (communication from counsel warning defendant not to make a public offering of debentures which was intended to be included in a letter forwarded to others renders the advice not privileged).

73. *Bergonzi*, 216 F.R.D. at 493–94; *see also* United States v. Sudikoff, 36 F. Supp. 2d 1196, 1204–05 (C.D. Cal. 1999).

74. *E.g.*, United States v. Cote, 456 F.2d 142, 144 (8th Cir. 1972).

75. United States v. Under Seal (*In re* Grand Jury Subpoena), 341 F.3d 331, 335–36 (4th Cir. 2003) (holding that waiver does not apply to litigation filings and similar documents).

76. Colton v. United States, 306 F.2d 633, 638–39 (2d Cir. 1962).

77. *Id.* at 639. In *United States v. McDonald*, 313 F.2d 832, 835 (2d Cir. 1963), the subpoena called for an attorney to produce copies of closing statements and sales contracts for all real estate closings in which the client was involved. Citing *Colton* and *Tellier*, the court held these materials were not privileged because "[h]is client necessarily contemplated divulging the information requested to other parties at the closing." *McDonald*, 313 F.2d at 835. However, the court did not extend disclosure to all underlying communications and documents that contained client confidences. *Id.*

While the Second Circuit was moving toward the Modern View as early as 1962, in 1984, the Fourth Circuit was still committed to the Classic View.[78] During that year, the Fourth Circuit issued the second of four influential Classic View decisions, *In re Grand Jury Proceedings*.[79] Doe was a securities attorney subpoenaed to testify about his work for the investigation's targets in their effort to create and market a limited partnership in coal mining equipment. Doe's clients had retained him to prepare a prospectus for marketing the limited partnership interests. He met once with the three targets to discuss "the preparation of a prospectus to be used in the enlistment of investors in the private placement," and he had conversations with one target about the information to be included.[80] About two weeks after he had been retained, Doe was instructed to stop. He did no further work. When he was subpoenaed, his former clients objected to his testifying about their discussions on grounds of the attorney-client privilege. The Fourth Circuit ruled, "we have no difficulty in concluding under the admitted facts of this case that *all information* given [Doe] by any of the joint venturers connected with the subject-matter of the proposed issuance of participations is without the protection of the attorney-client privilege."[81] The panel relied in part upon *Tellier*, ruling that because the intent was to prepare and publish a prospectus, the targets intended the information communicated to their lawyer could be conveyed to others.[82] Thus the communication was made without the intent to keep all the information confidential. The fact that no prospectus had ever been prepared was irrelevant.[83]

78. The Fourth Circuit was not alone. *See* United States v. Skeddle, 989 F. Supp. 917, 919–20 (N.D. Oh. 1997) (noting the Sixth Circuit follows the subject matter waiver rule and discussing the Classic View rationale); *see also* United States v. Mendelsohn, 896 F.2d 1183, 1188–89 (9th Cir. 1990) (relating to subject matter waiver).

79. *In re* Grand Jury Proceedings, 727 F.2d 1352 (4th Cir. 1984). The first of these cases is the Fourth Circuit's 1982 decision in *United States v. Jones*, 696 F.2d 1069, 1072 (4th Cir. 1982), in which the Fourth Circuit held that information provided to an attorney related to preparation of a tax opinion for a private placement memo was not privileged.

80. *In re* Grand Jury Proceedings, 727 F.2d at 1354.

81. *Id.* at 1358 (emphasis added).

82. The Fourth Circuit cited *Colton v. United States* for the proposition that the privilege is waived as to matters disclosed to third parties, but the court did not address *Colton v. United States* with respect to the scope of that waiver. *Id.* at 1356.

83. *Id.* at 1358.

Later in 1984, the Fourth Circuit decided *In re Grand Jury 83-2 John Doe No. 462 (Under Seal)*,[84] in which the circuit first seemed to be shifting toward the Modern View. In that case the client had retained an attorney to investigate the possibility of preparing and filing a tax opinion that, if prepared, would have been disclosed to third parties in support of a limited partnership offering. The court held that materials related to the decision of whether to make a public filing were privileged.[85] The court distinguished *In re Grand Jury Proceedings* on the grounds that the client expected to publish a prospectus, and thus the attorney had been retained to convey information to third parties. In *In re Grand Jury 83-2*, the panel found that the client sought advice about whether publication was required, which left the client the option to choose not to proceed. The panel in part relied on the Second Circuit's *Colton* decision.[86] However, the Fourth Circuit noted the types of underlying information that would be ordered disclosed once the decision to publish is made by the client. The language employed by the court is disconcerting to those who might believe that drafts are privileged:

> The details underlying the published data are the communications relating the data, the document, if any, to be published containing the data, all preliminary drafts of the document,

84. United States v. (Under Seal) (*In re* Grand Jury 83-2 John Doe No. 462), 748 F.2d 871 (4th Cir. 1984).

85. *Id.* at 875–76.

> The situation faced by this court in *In re Grand Jury* is significantly different from the situation, posed by some of the documents in this case, in which a client employs an attorney to research the *possibility* of filing public papers. Only when the attorney has been authorized to perform services that demonstrate the client's intent to have his communications published will the client lose the right to assert the privilege as to the subject matter of those communications. This result will not be changed by a fortuity that prevents publication, such as in *In re Grand Jury* where the clients simply failed to stay in contact with their attorney, but it will be altered if the client subsequently decides not to publish his communications and tells his attorney before the release of any of the communications. In the latter situation, the attorney has counseled his client in a matter that the client was able to conclude, presumably through full and frank discussions with his attorney, should remain private. A contrary result might discourage clients who had contemplated a public course of action from fully informing their attorney of all relevant facts when the possibility of refraining from public action begins to dawn. *See* Trammel v. United States, 445 U.S. 40, 51 (1980) ("the lawyer-client privilege rests on the need for the advocate and counselor to know all that relates to the client's reasons for seeking representation if the professional mission is to be carried out").

Id.

86. United States v. (Under Seal) (*In re* Grand Jury 83-2 John Doe No. 462), 748 F.2d 871, 876 n.8 (4th Cir. 1984).

> and any attorney's notes containing material necessary to the preparation of the document. Copies of other documents, the contents of which were necessary to the preparation of the published document, will also lose the privilege.[87]

The essential feature of the Classic View, as played out in these cases, is a binary approach to privilege—if one expects to publish something, then everything provided to the lawyer thereafter is deemed to be provided with the expectation that it might be published. The willingness to permit publication is deemed to be irreconcilable with confidentiality.[88] If one has not yet decided to publish, however, information can remain confidential and privileged until the decision is made.

The Second Circuit's shift away from the Classic View is most clear in *In re Grand Jury Subpoena Duces Tecum Dated Sept. 15, 1983*.[89] In that case, Marc Rich & Co. moved to quash a grand jury subpoena *duces tecum* issued to its general counsel. The government argued that documents in the custody of the Marc Rich & Co.'s former counsel were not privileged because the documents included information about compensation plans for employees that the company had intended to publish to those employees. The Second Circuit panel disagreed, holding that the possibility that information in the documents (some of which were drafts) would be given to third parties did not vitiate the privilege or create the inference that the communications were not intended to be confidential:[90]

> [T]he the fact that certain information in the documents might ultimately be disclosed to AG employees did not mean

87. *Id.* at 875 n.7.

88. In *United States v. Jones*, the Fourth Circuit noted that in *In re Sealed Case* the D.C. Circuit opined that courts retain "discretion not to impose full waiver as to all communications on the same subject matter where the client has merely disclosed a communication to a third party, as opposed to making some use of it." United States v. Jones, 696 F.2d 1069, 1072 (4th Cir. 1982) (quoting *In re* Sealed Case, 676 F.2d 793, 809 n.54 (D.C. Cir. 1982)). However, the *Jones* court did not adopt this view. *Jones*, 696 F.2d at 1072-73. The approach has been criticized in a variety of ways. *See* Paul R. Rice, *Attorney-Client Privilege: Continuing Confusion About Attorney Communications, Drafts, Pre-existing Documents and the Source of Facts Communicated*, 48 AM. U. L. REV. 967 (1999).

89. *In re* Grand Jury Subpoena Duces Tecum Dated Sept. 15, 1983 (Marc Rich & Co. A.G. v. United States), 731 F.2d 1032 (2d Cir. 1984).

90. Of course, drafts prepared by an attorney are protected by the attorney-client privilege only if the draft contains confidential information communicated by the client, and if the draft is maintained in confidence. A draft is not privileged simply because it is prepared by an attorney. *E.g.*, SEC v. Beacon Hill Asset Mgmt., LLC, 231 F.R.D. 134, 145 (S.D.N.Y. 2004) (collecting cases).

> that the communications to Proskauer were foreclosed from protection by the privilege as a matter of law. Nor did the fact that certain information might later be disclosed to others create the factual inference that the communications were not intended to be confidential at the time they were made.[91]

Two years later, in the 1986 case *Schenet v. Anderson*,[92] the District Court for the Eastern District of Michigan issued an influential decision contrasting the Modern View's pragmatic approach with the Classic View's ideologically pure approach. The district court held that if a portion of the information conveyed to counsel is published, the privilege does not apply to that published information, but the un-published portion remains privileged. In short, the Eastern District of Michigan adopted the Modern View.

In *Schenet*, the Michigan district court was willing to perform a granular analysis of the communications and the ultimate disclosure. The court was also willing to assume that the privilege protects communications provided for the purpose of assessing and preparing public filings, such as notes made on drafts, thereby protecting the confidentiality of the process of preparing the public filings. The approach depends upon the willingness to permit clients to engage in a decision process about publication that is conditional, not binary. The client is permitted to conclude: "I am willing to disclose what confidential information I must in order to comply with law and meet my objectives, but no more than is required." On several occasions, district courts in the Second Circuit have adopted *Schenet*'s approach and expounded on the protection for draft doc-

91. *Marc Rich & Co. A.G.*, 731 F.2d at 1037 (emphasis added). The court continued, "If confidentiality were not intended, of course, the privilege would not attach, but we see no indication that confidentiality was not intended. For example, although some of the documents appear to be drafts of communications the final version of which might eventually be sent to other persons, and as distributed would not be privileged, we see no basis in the record for inferring that AG did not intend that the drafts—which reflect its confidential requests for legal advice and were not distributed—to be confidential. Confidentiality may also, of course be waived; but we see no indication that a waiver has yet occurred." *Id.* (citation omitted).

The *Marc Rich & Co. A.G.* court did note that the privilege protects only confidential communications that are reflected in such documents. In focusing on confidential communications to counsel instead of information the court returned to the core principles of the privilege. Thus, even under the Modern View drafts, communications, and data related to a public filing are not automatically privileged because they are drafts or unpublished; the proponent of the privilege must prove the documents and testimony contain confidential communications.

92. Schenet v. Anderson, 678 F. Supp. 1280 (E.D. Mich. 1988).

uments.[93] The Modern View allows the client with this contingent intent to protect information by looking to the result of the process of evaluation, and to the client's final decision about what is to be published. As a consequence, the Modern View protects the confidentiality of the process.[94]

The courts espousing the Modern View did not adopt the published/unpublished distinction in a single decision.[95] For example, *Saxholm AS v. Dynal, Inc.* held that communications from a client to attorneys conveying information or requesting advice regarding a patent application were privileged, while the draft patent application was not, in part, because the communication was made with the expectation that it would be disclosed.[96] Later decisions expressing the view that *Saxholm AS* and similar cases took too narrow a

93. Muller v. Walt Disney Prods., No. 93 Civ. 0427 (GLG), 1994 WL 801529, at *1 (S.D.N.Y. Sep. 27, 1994) ("The attorney-client privilege applies to all information conveyed by clients to their attorneys for the purpose of drafting documents to be disclosed to third parties and all documents reflecting such information to the extent that such information is not contained in the document published and is not otherwise disclosed to third parties. Preliminary drafts may reflect not only client confidences, but also the legal advice and opinions of attorneys, all of which is protected by the attorney-client privilege.") (citing *Schenet*, 678 F. Supp. at 1284). In *USPS v. Phelps Dodge Ref. Corp.*, 852 F. Supp. 156, 163 (E.D.N.Y. 1994), the district court noted that drafts of documents ultimately published to third parties "may be considered privileged if they were prepared with the assistance of an attorney for the purpose of obtaining legal advice and/or contain information a client considered but decided not to include in the final version." *See also* Diversified Grp., Inc. v. Daugerdas, 304 F. Supp. 2d 507, 514 (S.D.N.Y. 2003) (noting that drafts of marketing materials were privileged because the document specifically required the lawyer to review, in his capacity as a tax attorney, the section of the marketing materials summarizing the tax considerations of the strategy); *In re* U.S. Healthcare, Inc. Sec. Litig., No. 88-0559, 1989 WL 11068, at *2 (E.D. Pa. Feb. 8, 1989).

94. For one example, this contingent intent dichotomy is evident in *Daugerdas*, 304 F. Supp. 2d at 514, in which drafts of marketing materials were held to be privileged because the client requested the lawyer to review, in his capacity as a tax attorney, the section of the marketing materials summarizing the tax considerations of the strategy. *See also* Rattner v. Netburn, No. 88 CIV.2080(GLG), 1989 WL 223059, at *4 (S.D.N.Y. June 20, 1989) (stating that a client may intend to direct or permit the release of the final version of a document while still intending that his communications with his attorney prior to the finalization of the document will remain confidential).

95. *See* P. & B. Marina, Ltd. v. Logrande, 136 F.R.D. 50, 56 (E.D.N.Y. 1991) ("Drafts of letters or documents to third parties lack the requisite confidentiality to be protected under the attorney-client privilege. Moreover, the privilege is generally waived upon disclosure to third-parties of otherwise confidential materials.") (citations omitted).

96. Saxholm AS v. Dynal, Inc., 164 F.R.D. 331, 336–37 (E.D.N.Y 1996) (finding that the underlying communications remained privileged).

view of the attorney-client privilege have ruled that drafts of letters or documents ultimately sent to third parties lack the requisite confidentiality to be protected.[97] By 2006, in *In re Rivastigmine Patent Litigation*, the Southern District of New York felt comfortable recognizing the change in view of the privilege as applied to patent applications since *Saxholm*; the court held that drafts of patent applications, amendments, and supporting affidavits will be privileged.[98] Courts in other circuits have ruled likewise.[99] It is worth noting that some jurisdictions have justified refusing to apply the privilege to draft SEC filings based on the view that legal advice is not involved, similar to the Eastern District of New York's reasoning in *Saxholm* with respect to preparation of patent applications. For instance, the District Court for the Northern District of Illinois has held that draft proxy statements, SEC forms, comments to drafts, and letters regarding the drafts are not privileged because they are not legal advice.[100]

97. *See, e.g.*, *In re* Rivastigmine Patent Litig. (MDL No. 1661), 237 F.R.D. 69, 85–86 (S.D.N.Y. 2006); Softview Computer Prods. Corp. v. Haworth, Inc., 97 Civ. 8815 (KMW) (HBP), 2000 U.S. Dist. LEXIS 4254, at *53 (S.D.N.Y. Mar. 31, 2000).

98. *In re Rivastigmine Patent Litig. (MDL No. 1661)*, 237 F.R.D. at 85–86; *see also Daugerdas*, 304 F. Supp. at 514 (noting that drafts of marketing materials were held to be privileged because the client requested the lawyer to review the section of the marketing materials summarizing the tax considerations of the strategy in his capacity as a tax attorney).

99. *See, e.g.*, SEC v. Teo, No. 04-1815 (SDW), 2009 WL 1684467, at *12 (D.N.J. June 12, 2009) (granting motion to quash subpoena seeking drafts of Schedule 13D filings); Apex Mun. Fund v. N-Group Secs., 841 F. Supp. 1423, 1428 (S.D. Tex. 1993) (denying motion to compel discovery of the communications and documents underlying the bond offering where statements disclosed to the public); *In re* U.S. Healthcare, Inc. Secs. Litig., No. 88-0559, 1989 WL 11068, at *2 (E.D. Pa. Feb. 8, 1989) (denying motion to compel production of documents underlying financial statements and SEC filings); *see also In re* Sealed Case, 877 F.2d 976, 979 (D.C. Cir. 1989) (refusing to enforce grand jury subpoena seeking attorney-client communications relating to "adjusting entries" for an amended tax return filed with the IRS); Buford v. Holladay, 133 F.R.D. 487, 492 (S.D. Miss. 1990) (refusing to permit attorney general to be questioned on attorney-client communications even though opinion letters had been published); Ball v. U.S. Fid. & Guar. Co., No. M8-85 (RWS), 1989 WL 135903, at *2 (S.D.N.Y. Nov. 8, 1989) (denying motion to compel production of attorney-client communications where some of counsel's opinions were disclosed in industry committee meetings).

100. *See* Blau v. Harrison (*In re* JP Morgan Chase & Co. Secs. Litig.), No. 04 C 6592, 2007 U.S. Dist. LEXIS 60095, at *9 (N.D. Ill. Aug. 13, 2007) (citing Christman v. Brauvin Realty Advisors, Inc., 185 F.R.D. 251, 256–57 (N.D. Ill. 1999)). Interestingly, the Northern District of Illinois contradicted its own holdings in *JP Morgan* and *Christman* in other decisions. *See* Roth v. Aon Corp., 254 F.R.D. 538, 541 (N.D. Ill. 2009) (holding drafts of 10-K and communications related to preparation are privileged because the determination of what information should be disclosed for compliance is not merely a business operation but a legal

B. The Reaffirmation of the Classic View

The Fourth Circuit revisited the issue of privilege waiver in *United States v. Under Seal (In re Grand Jury Proceedings)*.[101] The Fourth Circuit rejected the more flexible approach crystallized in *Schenet*. In *Under Seal (In re Grand Jury Proceedings)*, counsel for the parent of a public company had refused to produce drafts, notes, and memoranda generated for an audit response to the company's auditor, and one of the company's subsidiaries had refused to produce drafts of SEC filings and "related documents" sought by a government subpoena. The appellants advocated that the Fourth Circuit should adopt *Schenet*, but the court refused.[102] The panel began its analysis by stating that privilege would not apply to communications made in connection with a proposed public disclosure.[103] The Fourth Circuit ruled that the "only way the appellants can prevail is to demonstrate to the court that they did not retain the services of the attorneys for the purpose of advice on publication."[104]

The Classic View can be particularly disconcerting to public companies. Few contemplate that because SEC filings are published, the process of preparing those filings, including drafts and comments on drafts, is not protected by the privilege. In *Neuberger Berman Real Estate Income Fund, Inc. v. Lola Brown Trust No. 1B*, a Maryland district court aptly summarized the Fourth Circuit precedent and ordered the disclosure of a panoply of materials related to

concern) (citing Ziemack v. Centel Corp., No. 92 C 3551, 1995 WL 314526, at *4 (N.D. Ill. May 19, 1995)).

101. United States v. Under Seal (*In re* Grand Jury Proceedings), 33 F.3d 342 (4th Cir. 1994).

102. *Id.* at 355.

103. *Id.* at 354. The court relied on the Fourth Circuit's decision in *In re Grand Jury Proceedings*, 727 F.2d 1352 (4th Cir. 1984), "as modified" by *United States v. (Under Seal)* (*In re* Grand Jury 83-2 John Doe No. 462), 748 F.2d 871 (4th Cir. 1984).

104. *Id.* at 355. "If a client communicates information to his attorney with the understanding that the information will be revealed to others, that information as well as 'the details underlying the data which was to be published' will not enjoy the privilege," and "the loss of the privilege [is not] confined to the particular words used to express the communication's content but extends to the substance of a communication, since the disclosure of any significant part of a communication waives the privilege and requires the attorney to disclose the details underlying the data which was to be published." *Id.* at 354–55 (citing *In re Grand Jury Proceedings*, 727 F.2d at 1356).

This controversy is discussed at length in an article that favors the approach taken in the *Schenet* decision, focusing on communications, not pre-existing information. Rice, *supra* note 88.

the defendant's Rule 13D filing, including all related emails, drafts, and documents with information necessary to prepare the filings.[105]

The question of contingent intent is particularly problematic in the context of government investigations. It is increasingly the practice that corporations conduct internal investigations with the express or implied understanding that the results of the investigation will be shared with government investigators. There is, therefore, from the outset, an argument that there is no expectation of confidentiality as to the final work product. What percentage of the underlying information, such as email, interview notes, and legal research will be disclosed is often decided only later. When the investigation is commenced under the cloak of the attorney-client privilege, but—as is often the case—the corporation reserves the decision whether to turn over some or all of the results of the investigation, the question of if privilege attaches becomes murkier. As one would expect, the fact that an investigation report is not published is not necessarily determinative.[106]

One case that illustrates the impact of expectations on whether investigation materials are privileged from the outset is *United States v. Bergonzi*.[107] In *Bergonzi*, HBO & Company ("HBOC") conducted an investigation of securities fraud as the result of misstatements of

105. Neuberger Berman Real Estate Income Fund, Inc. v. Lola Brown Trust No. 1B, 230 F.R.D. 398, 414–15 (D. Md. 2005). Furthermore, the communications via email that accompanied the drafts would not be privileged, as they constitute the details "underlying the published data" of the public disclosure. The Fourth Circuit has described these details as including "the communications relating the data, the document, if any, to be published containing the data, all preliminary drafts of the document, and any attorney's notes containing material necessary to the preparation of the document." Further, any copies of other documents, "the content of which were necessary to the preparation of the published document," will not be privileged. *Id.* Therefore, all the drafts of the section 13D filing and accompanying underlying "details" found in Volume II are not privileged. *Id.* (citation omitted).

The same result follows for the drafts and related communications of other disclosed documents for which privilege has been asserted. A public disclosure to a government agency is not required. "Public" can also mean "public action," i.e., disclosure to any third party outside the attorney-client relationship. *See* United States v. Under Seal (*In re* Grand Jury Subpoena), 204 F.3d 516, 522 (4th Cir. 2000) (concluding that no privilege existed where client, through his attorney, made disclosure via letter to another attorney). "When a client hires an attorney to take public actions on the client's behalf—in this case, sending the letter to [the other attorney] in order to avoid a lawsuit—the privilege does not extend . . . once that public action is taken." *Neuberger Berman*, 230 F.R.D. at 414–15.

106. United States v. Tellier, 255 F.2d 441, 447 (2d Cir. 1958).

107. United States v. Bergonzi, 216 F.R.D. 487 (N.D. Cal. 2003).

the financial results of HBOC and McKesson HBOC.[108] HBOC began the internal investigation after McKesson's announcement in April of 1999 that its auditors had discovered accounting irregularities.[109] Private actions alleging securities fraud followed immediately.[110] The McKesson board of directors' audit committee retained counsel, who in turn retained an accounting firm to assist in the review.[111]

After fifty-five documented interviews of thirty-seven present and former employees, a report prepared by the audit committee's counsel and the backup materials (including all of the interview memoranda) were provided to the SEC and the local United States Attorney's Office.[112] McKesson entered into a confidentiality agreement with the SEC under which the SEC and the company agreed that the documents should be kept confidential, and which took extraordinary steps to shield the documents from discovery in the private lawsuits.[113] The confidentiality agreement recited that the company had a "common interest" with the SEC in the results of the internal investigation.[114] The SEC agreed it would not argue that the voluntary submission of the information would constitute a waiver of any privilege and promised to maintain the confidentiality of the materials, "except to the extent that the [SEC] determines that disclosure is otherwise required by federal law."[115] Later, in May of 1999, the company entered into a separate, similar agreement with the United States Attorney's Office, but that agreement provided that the United States Attorney's Office could use any documentation provided in pursuit of criminal proceedings.[116]

108. *Id.* at 490.

109. *Id.*

110. *Id.*

111. *Id.*

112. *Id.* at 491.

113. *Id.*

114. *Id.*

115. *Id.* at 494. The case that gave rise to elaborate efforts between HBOC and the government is *Diversified Industries, Inc. v. Meredith*, 572 F.2d 596 (8th Cir. 1978), in which the Eighth Circuit adopted the equivalent of a "selective waiver" doctrine (not to be confused with limited waiver) for voluntary disclosures made to the government under which the delivery of the documents to the government in its regulatory or law enforcement capacity did not waive the privilege as to any other parties. This effort to facilitate regulatory functions and settlements while relieving corporations of follow-on civil liability driven by the disclosure of the investigation results has been universally rejected under both the Classic and the Modern Views.

116. United States v. Bergonzi, 216 F.R.D. 487, 491 (N.D. Cal. 2003).

In September of 2000, a grand jury indicted a number of former executives of HBOC for securities fraud, mail fraud, and wire fraud.[117] After the indictments, the U.S. Attorney's Office inadvertently produced four of the interview memoranda to the defendants.[118] The government immediately sought return of the four interview memoranda, but one of the defendants refused.[119] McKesson intervened in the criminal case for the purpose of asserting the privilege, as well as the confidentiality, of the report and the back-up material, and to obtain a return of the interview memoranda.[120] Meanwhile, the former officers sought production of the report and all backup materials, arguing that any privilege was waived, and, further, that they were entitled to the documents both under Rule 16 of the Federal Rules of Criminal Procedure and under the government's *Brady v. Maryland* obligation to provide exculpatory material.[121]

The *Bergonzi* court had little difficulty finding that the documents were not privileged, ruling that whatever the merits of a waiver argument, it was the Company's intention when the investigation began that the materials would be produced to the government.[122] As a result, the attorney-client privilege never attached in the first instance.[123] The court relied on the fact that all of the documents had been prepared after the company had agreed to provide the report and the backup materials, including the interview memoranda, to the government.[124] The *Bergonzi* court also ruled

117. *Id.* at 490.
118. *Id.* at 491.
119. *Id.* at 490.
120. *Id.* at 492.
121. *Id.*
122. *Id.* at 494.
123. *Id.*
124. *Id.* at 493 (citing United States v. Sudikoff, 36 F. Supp. 2d 1196, 1204–05 (C.D. Cal. 1999)). The requirement that employees know that information is provided in confidence and not for publication for privilege to attach can be a trap for the unwary in other contexts. In *Deel v. Bank of America*, 227 F.R.D. 456, 463 (W.D. Va. 2005), the court ruled that employee questionnaires completed for a FLSA investigation were not privileged. The bank had conducted a survey of employees to obtain information about possible FLSA violations. The court held that to secure the privilege of questionnaires sent to employees, the employer must notify the employees that it seeks their help in obtaining legal advice by the employees providing information in confidence, which was not done. Generally, the impact on arguments of work product protection has been similar under the Modern View. For example, in *In re Royal Ahold N.V. Secs. & ERISA Litig.*, 230 F.R.D. 433 (D. Md. 2005), the plaintiffs in a class action securities fraud case sought notes and memoranda prepared by the Royal Ahold companies' outside counsel describing witness interviews conducted as part of an investigation into accounting irregulari-

that the documents were not protected by the doctrine of attorney work product, or the doctrine regarding materials prepared in anticipation of litigation.[125]

The Advisory Committee Notes state that Rule 502 makes no attempt to alter federal or state law on whether information is protected under the attorney-client privilege or work product immunity as an initial matter. The rule's failure to address privilege as an initial matter leaves open the possibility that privilege law strictly applied will moot the effort to limit the scope of the implied waiver. However, when it comes to investigations, original intent and waiver can be two sides of the same coin, and which side of the coin controls is often determined by ambiguous facts with respect to when the company decided to prepare and disclose a report to the government.

IV.
IMPLIED WAIVER

A waiver of attorney-client privileged materials is implied from the actions of a client in two basic circumstances. The attorney-client privilege may be waived "by placing the subject matter of coun-

ties and alleged fraud. The volume of material was startling; there was a total of at least 827 interview memoranda, of which approximately 269 had been disclosed to government agencies and approximately 558 had not. The disclosed memoranda were disclosed under a confidentiality agreement that prohibited the government from disclosing the documents to any third party, and stated that Royal Ahold had not waived any privilege. The court ruled that, as to the memos disclosed to government agencies, there was a waiver of both fact and mental impression work product. With respect to the memos that were not disclosed directly, the court ordered disclosure to the plaintiffs of all fact information in the memos that had been the basis of disclosures to the government, to accountants or to the public, but indicated an intent to preserve mental impression work product as to those memos. *Id.* at 437–38. The trial court declined to follow *Saito v. McKesson HBOC, Inc.*, No. 18553, 2002 LEXIS 125, at *19 (Del. Ch. Oct. 25, 2000), in which the Delaware Chancery Court found that a company and the government did not share a common interest, but applied a selective waiver analysis to keep documents confidential.

125. The trial court rejected the arguments of the company and the government that the two had a common interest in enforcing the relevant regulations and statutes. Relying principally on the decision of the United States Court of Appeals for the First Circuit in *United States v. Mass. Inst. of Tech.*, 129 F.3d 681 (1st Cir. 1997), the court ruled that the common interests professed by the government and the company—their joint interest in ensuring compliance with the law and their joint interest in the government's carrying out its law enforcement responsibilities—are not the types of common interest shared by parties who are working together in the prosecution or defense of a lawsuit.

sel's advice in issue[126] or by making selective disclosure of only part of such advice."[127] We refer to the former as the "at-issue" waiver and the latter as the "disclosure" waiver. At-issue waiver results when a client takes a position in litigation, "the validity of which can only be tested by invasion of the attorney-client privilege,"[128] and the client asserts a claim or defense that requires the use of the privileged materials for proof.[129] The paradigm example of at-issue waiver is the party who asserts his good faith belief as a defense to a lawsuit on a subject about which he obtained legal advice, such as his duty to make disclosure under the securities laws. By pleading this good faith defense he puts his communications with counsel about his obligations at issue and the privilege is waived. The Classic View courts do not accept or protect limited disclosure, so the disclosure of a portion of a communication waives the privilege as to the remainder of the communications on the same subject mat-

126. United States v. Jones, 696 F.2d 1069, 1072 (4th Cir. 1982) ("[W]hen a party reveals part of a privileged communication to gain advantage in litigation, the party waives the attorney-client privilege as to all other communications relating to the same subject matter.").

127. Soho Generation v. Tri-City Ins. Brokers, 653 N.Y.S.2d 924, 925 (App. Div. 1997) (citation omitted).

128. The Modern View taken by the Second Circuit extends to recharacterizing the waiver itself: "Because the words *implied* and *waiver* both tend to suggest that the party possessing the privilege *intended* to give it up, the terms *waiver*, and *implied waiver*, are not especially appropriate designations for circumstances in which the party possessing the privilege makes no representation, express or implied, that it intends to surrender its privilege. In such circumstances, the rule is perhaps more aptly described as one of forfeiture, rather than waiver." John Doe Co. v. United States (*In re* Grand Jury John Doe Co.), 350 F.3d 299, 302 (2d Cir. 2003); Village Bd. of Pleasantville v. Rattner, 515 N.Y.S.2d 585, 586 (App. Div. 1987) (holding that municipalities relying on good faith reliance on counsel waived the attorney-client privilege regarding communications concerning the transactions for which counsel's advice was sought; the plaintiff was entitled to test assertions of good faith).

129. Deutsche Bank Trust Co. v. Tri-Links Inv. Trust, 837 N.Y.S.2d 15, 23 (App. Div. 2007) (citing *Village Bd. of Pleasantville*, 515 N.Y.S.2d at 655). Unsurprisingly, courts have also found that a party places its attorney-client communications "at issue" in legal malpractice actions. *See, e.g.*, Creditanstalt Inv. Bank AG. v. Chadbourne & Parke LLP, No. 106539/01, 2005 WL 3681657, at *3 (N.Y. Sup. Ct. Nov. 18, 2005), *aff'd*, 831 N.Y.S.2d 705 (App. Div. 2007) (finding "that the issues of what legal advice plaintiffs received regarding the Russian operation's compliance with Russian tax laws and licensure requirements are central to defendant's defense in this action [and] the application of the privilege in this case would deprive defendant of vital information"); Goldberg v. Hirschberg, 806 N.Y.S.2d 333, 336 (Sup. Ct. 2005) (finding that plaintiffs placed their representation by current counsel "at issue" by claiming as damages the legal fees paid to current counsel "to get them out of trouble directly caused by defendants' bad advice").

ter.[130] Similarly, whenever a privileged communication is made relevant and material by the assertion of a claim or defense, the privilege is waived. The underpinning of these limits of the attorney-client privilege under the Classic View is, again, the concept that the privilege protects confidentiality; any action by the client that is inconsistent with that imperative destroys the theoretical basis for the privilege and therefore the privilege itself.

It is critical to understand that Rule 502 addresses waiver by disclosure only; it is not intended to address at-issue waiver resulting from a client putting privilege at issue.[131] Thus, the conflict between the Classic View and the Modern View as to precisely when such waiver occurs, and the scope of that waiver, is theoretically unaffected by Rule 502's express provisions. However, Congress' Statement of Purpose makes clear that a subject matter waiver results:

> This subdivision [502(a)(1)] does not alter the substantive law regarding when a party's strategic use in litigation of otherwise privileged information obliges that party to waive the privilege regarding other information concerning the same subject matter, so that the information being used can be fairly considered in context. One situation in which this issue arises, the assertion as a defense in patent-infringement litigation that a party was relying on advice of counsel, is discussed elsewhere in this Note. In this and similar situations, under subdivision (a)(1) the party using an attorney-client communication to its advantage in the litigation has, in so doing, intentionally waived the privilege as to other communications concerning the same subject matter, regardless of the circumstances in which the communication being so used was initially disclosed.[132]

Until recently, many courts adopted the test set forth in the Eastern District of Washington decision *Hearn v. Rhay*[133] for deter-

130. United States v. Pollard (*In re* Martin Marietta Corp.), 856 F.2d 619, 623–24 (4th Cir. 1988) (citing United States v. (Under Seal) (*In re* Grand Jury 83-2 John Doe No. 462), 748 F.2d 871 (4th Cir. 1984)).

131. The Statement of Purpose added to the Advisory Committee Note by Congress makes this point clear: "[The] rule addresses only the effect of disclosure, under specified circumstances, of a communication that is otherwise protected by attorney-client privilege, or of information that is protected by work-product protection, on whether the disclosure itself operates as a waiver of the privilege or protection for purposes of admissibility of evidence in a federal or state judicial or administrative proceeding."

132. 154 CONG. REC. H7818 (2008) (Statement of Congressional Intent Regarding Rule 502 of the Federal Rules of Evidence).

133. Hearn v. Ray, 68 F.R.D. 574, 581 (E.D. Wash. 1975).

mining whether a party has waived the attorney-client privilege by placing the advice at issue.[134] Now, courts adopting the Modern View have retreated from the *Hearn* test, finding that it results in waiver too often and produces a waiver that is too broad in scope.[135] The *Hearn* test is derived from the court's conclusion that all cases in which there was an invasion of the privilege had these factors in common:

> (1) [A]ssertion of the privilege was a result of some affirmative act;
> (2) through this affirmative act the asserting party put the protected information at issue [for its own benefit] by making it relevant to the case; and
> (3) application of the privilege would deny the opposing party access to information vital to its case.[136]

134. *E.g.*, Cornett Mgmt. Co. v. Lexington Ins. Co., No. 5:04-CV-22, 2007 WL 1140253, at *4 (N.D. W.Va. Apr. 17, 2007) (finding *Hearn* "persuasive"); Small v. Hunt, 152 F.R.D. 509, 512 (E.D.N.C. 1994) (applying *Hearn* to find waiver where party submitted affidavit, placing privilege at issue by claiming that a commission, which included attorneys, did not foresee certain circumstances); Santrade, Ltd. v. Gen. Elec. Co., 150 F.R.D. 539, 544 (E.D.N.C. 1993) (same); 1050 Tenants Corp. v. Lapidus, 817 N.Y.S.2d 491, 495 (Civ. Ct. 2006).

135. *E.g.*, Pritchard v. County of Erie, 546 F.3d 222, 227–28 (2d Cir. 2008) (finding that courts have criticized *Hearn* and have applied its tests unevenly); Pereira v. United Jersey Bank, Nos. 94 Civ. 1565(LAP) & 94 Civ. 1844(LAP), 1997 WL 773716, at *3 (S.D.N.Y. Dec. 11, 1997) ("*Hearn* is problematic insofar as there are very few instances in which the *Hearn* factors, taken at face value, do not apply and, therefore, a large majority of claims of privilege would be subject to waiver."); Allen v. West Point-Pepperell, Inc., 848 F. Supp. 423, 429–30 (S.D.N.Y. 1994) (noting that district courts within the Second Circuit have reached conflicting decisions in the application of *Hearn*, and rejecting reliance "upon a line of cases in which courts have unhesitatingly applied a variation of the *Hearn* balancing test"); Koppers Co. v. Aetna Cas. & Sur. Co., 847 F. Supp. 360, 363 (W.D. Pa. 1994) (rejecting the third *Hearn* factor for vagueness and overbreadth); Connell v. Bernstein-Macaulay, Inc., 407 F. Supp. 420, 422 (S.D.N.Y. 1976) ("The actual holding in [*Hearn*] is not in point because the party there asserting the privilege had expressly relied upon the advice of counsel as a defense to the plaintiff's action.").

The test also has been subject to academic criticism. *See, e.g.*, Richard L. Marcus, *The Perils of Privilege: Waiver and the Litigator*, 84 MICH. L. REV. 1605, 1628–29 (1986); Note, *Developments in the Law: Privileged Communications*, 98 HARV. L. REV. 1629, 1641–42 (1985) ("[T]he faults in the Hearn approach are (1) that it does not succeed in targeting a type of unfairness that is distinguishable from the unavoidable unfairness generated by every assertion of privilege, and (2) that its application cannot be limited."). In view of the foregoing, it seems to us that there is a need for clarification of the scope of the at-issue waiver and the circumstances under which it should be applied.

136. *Hearn*, 68 F.R.D. at 581; *see also* United States v. Bilzerian, 926 F.2d 1285, 1292 (2d Cir. 1991); United States v. Jones, 696 F.2d 1069, 1072 (4th Cir. 1982) (holding that "selective disclosure for tactical purposes" waives the privilege);

In the Second Circuit's view, contrary to the approach taken in *Hearn,* the mere fact that a privileged communication contains information *relevant* to issues that the parties are litigating does not put privileged communication sufficiently at issue to require an implied at-issue waiver.[137] In *In re the County of Erie*,[138] the Second Circuit made express its rejection of the *Hearn* test, holding that waiver occurs only when counsel's advice is *relied upon* in asserting a claim or defense. The court wrote, "[w]e agree with its critics that the *Hearn* test cuts too broadly. . . . Nowhere in the *Hearn* test is found the essential element of reliance on privileged advice in the assertion of the claim or defense in order to affect a waiver. . . . We hold that a party must rely on privileged advice from his counsel to make his claim or defense."[139] The *County of Erie* Court distinguished *United States v. Bilzerian*,[140] a criminal prosecution for securities fraud in which the Second Circuit found a waiver when the defendant proposed to testify that he thought his actions were legal based upon advice of counsel, which put at issue his communications with his lawyers regarding what the law required.[141] However, the Second Circuit said that the assertion of the defense did not result in

Small, 152 F.R.D. at 512 (applying *Hearn* to find waiver where party submitted affidavit, placing privilege at issue by claiming that a commission, which included attorneys, did not foresee certain circumstances); Garfinkle v. Arcata Nat'l Corp., 64 F.R.D. 688, 689 (S.D.N.Y. 1974) ("[P]rivilege may be waived if the privileged communication is injected as an issue in the case by the party which enjoys its protection."); Orco Bank, N.V. v. Proteinas Del Pacifico, S.A., 577 N.Y.S.2d 841, 842 (App. Div. 1992); Hayes v. Ricard, 93 S.E.2d 540, 548 (N.C. Sup. Ct. 1956).

137. *Cf.* Deutsche Bank Trust Co. of Ams. v. Tri-Links Inv. Trust, 837 N.Y.S.2d 15, 23 (App. Div. 2007); Veras Inv. Partners, LLC v. Akin Gump Strauss Hauer & Feld LLP, 860 N.Y.S.2d 78 (App. Div. 2008).

138. *Pritchard*, 546 F.3d at 228.

139. *Id.* at 227 (holding that the qualified immunity defense did not put protected communications "at issue" because the defense is based on an objective evaluation of relevant case law to determine whether a right was "clearly established" and not the defendant's subjective reliance on advice of counsel); *In re* Keeper of the Records, 348 F.3d 16, 24 (1st Cir. 2003) (adopting the view expressed in *In re von Bulow*, 828 F.2d 94, 102–03 (2d Cir. 1987)).

140. *Bilzerian*, 926 F.2d at 1285.

141. *Id.* at 1292; *see also* Allen v. West Point-Pepperell, Inc., 848 F. Supp. 423, 429–30 (S.D.N.Y. 1994) (finding that assertion of a laches defense did not waive the privilege because legal advice was not "in issue," as the defense did not depend on confidential communications with counsel, but on plaintiffs' knowledge of the facts). *But see In re* Leslie Fay Co. Secs. Litig., 161 F.R.D. 274, 283 (S.D.N.Y. 1995) (holding that production of the audit committee's report to the SEC waived any protection as to the report itself in that subsequent events in the litigation made it "manifestly unfair" to maintain the privilege as to interview notes and financial documents summarized in that report when the company sought to use the investigative report to incriminate an outside auditor in a securities fraud class action).

automatic waiver: instead, waiver would result only if he chose to testify as to his good faith belief. Since *County of Erie*, courts in the Second Circuit have continued to enforce this narrower waiver standard, under which the communication must itself be at issue, or essential to fair litigation before the fact-finder.[142] Similarly, if a party does not need a privileged communication to sustain its cause of action, no at-issue waiver results.[143]

A. Importance of Context to Implied Waiver

A court can use fairness to imply a waiver because Rule 502's approach to the question of whether undisclosed information has lost the attorney-client privilege in a federal proceeding or to a government regulator is deceivingly simple: disclosure is required if the undisclosed information concerns the same subject matter as the disclosed information and "they ought in fairness to be considered together."[144] The context—the circumstances under which a waiver occurs—is also critical. The Classic View and the Modern View take dramatically different positions on the role of context in deciding whether waiver has occurred, and in determining the scope of any such waiver. Context has become a key element in the fairness calculus of the Modern View; the Classic View applies virtually the same implied waiver rules, regardless of context. The implications for whether an implied waiver has occurred and its scope are drastic. Indeed, as Rule 502 is applied, it is possible that the most important difference between the Modern View and the Classic View is the context in which a waiver of privilege occurs. Under the Classic View, any disclosure, anytime, anywhere, results in implied waiver as to the subject matter. By contrast, the focus of the

142. For example, in *Bodega Invs., LLC v. United States*, No. 08 Civ. 4065(RMB)(MHD), 2009 U.S. Dist. LEXIS 78217, at *9 (S.D.N.Y. Aug. 21, 2009), the court applied *County of Erie* and found a waiver of attorney-client privilege by the plaintiff, who contested interest payments based on his alleged reliance on an IRS agent's statement to his counsel that the interest would be suspended, because, the court held, the substance of counsel's communications about what the IRS agent said and counsel's communications about whether to rely on the representation were "vital to assessing plaintiff's claim that he in fact relied, as well as whether such reliance was reasonable." *See also* Morande Auto. Grp., Inc. v. Metro. Grp., Inc., No. 3:04CV918(SRU), 2009 WL 650444, at *5 (D. Conn. Mar. 12, 2009) (citing *County of Erie*) (holding that although the plaintiff's privileged communications were relevant, the court found no waiver because the claim did not depend on plaintiff's reliance on legal advice, but on defendants' alleged negligent misrepresentations).

143. Carl v. Cohen, No. 117043/06, slip op. at 3 (N.Y. Sup. Ct. Apr. 7, 2009) (citation omitted).

144. FED. R. EVID. 502(a)(2)–(3).

Modern View is the protection of the integrity of the judicial process, so implied waiver of undisclosed privileged information will not occur in an extra-judicial context because there is nothing of sufficient import to protect.

The Second Circuit articulated the Modern View with regard to context in *In re von Bulow*,[145] where the court made clear its substitution of the "fairness doctrine" for the *Hearn* test.[146] Claus von Bulow was famously convicted of murdering his wife, appealed that conviction successfully, and was acquitted at the re-trial. He was represented by Alan Dershowitz on appeal and in the second trial. Dershowitz wrote a book about the case, "Reversal of Fortune—Inside the von Bulow Case," which disclosed confidential communications between von Bulow and his lawyers. After the book was published, Ms. von Bulow's children by her first marriage, who had sued their stepfather for their mother's death, sought discovery of all privileged materials from the criminal case, arguing that the publication, and related public appearances by Dershowitz and von Bulow to tout the book, affected a subject matter waiver of the attorney-client privilege. The Second Circuit held that the rule requiring disclosure of the subject matter of the communication "does not come into play when . . . the privilege-holder or his attorney has made extrajudicial disclosures, and those disclosures have not subsequently been placed at issue during litigation."[147]

The Second Circuit had no trouble concluding that von Bulow had waived the privilege as to every portion of a communication

145. *In re* von Bulow, 828 F.2d 94 (2d Cir. 1987).

146. With the rise of the need to make public disclosure, including press releases, about investigations and litigation, corporations have found the need to consult their in-house or outside public relations consultants. They have also felt the need to make public statements about the underlying allegations, the corporation's response to the allegations, or the corporation's defense to the allegations. These public statements have given rise to another layer of ambiguity or uncertainty about the attorney-client privilege. There is an increasing trend in the case law to accommodate these realities without forcing the corporation to waive the attorney-client privilege (as might be predicted by the Classic View of the privilege). The fairness doctrine has played a prominent role in these analyses. *See* Calvin Klein Trademark Trust v. Wachner, 198 F.R.D. 53, 54–55 (S.D.N.Y. 2000) (finding disclosure to PR firm waives privilege for information disclosed); *see also* Deniza Gertsberg, Comment, *Should Public Relations Experts Ever Be Privileged Persons?*, 31 FORDHAM URB. L.J. 1443 (2004) (addressing the privilege extended to attorney communications with public relations experts). *Contra* Viacom, Inc. v. Sumitomo Corp. (*In re* Copper Mkt. Antitrust Litig.)., 200 F.R.D. 213, 219–20 & n.4 (S.D.N.Y. 2001) (interpreting *Upjohn Co. v. United States*, 449 U.S. 383 (1981) and holding no waiver of privileged information given to PR professionals because they are the functional equivalent of an employee for purposes of the case).

147. *In re von Bulow*, 828 F.2d at 102.

disclosed in the book; however, at the outset, the court defined the purpose of the fairness doctrine narrowly: "to prevent prejudice to a party and distortion of the judicial process that may be caused by the privilege-holder's selective disclosure during litigation of otherwise privileged information."[148] Then, applying the fairness doctrine, the court concluded first that "this rule protecting the party, the fact-finder, and the judicial process from selectively disclosed and potentially misleading evidence does not come into play when, as here, the privilege-holder or his attorney has made extrajudicial disclosures, and those disclosures have not subsequently been placed at issue during litigation."[149] The court held that the extrajudicial disclosures of the attorney-client communications in the book and on the promotional appearances did not result in implied waiver because von Bulow did not subsequently use the disclosed communications to his adversary's prejudice in a judicial proceeding.[150] Finding no prejudice to the children in the litigation, the court held that the fairness doctrine did not demand disclosure.[151]

The Modern View's narrow approach to fairness conflicts with the Classic View. In *United States v. Under Seal (In re Grand Jury Subpoena)*,[152] the defendant told FBI agents during a non-custodial interview that he had made a misrepresentation on an INS form "under the advice of an attorney." After the grand jury subpoenaed the attorney, the client attempted to prevent the attorney from disclosing the advice to the grand jury investigating him for immigration fraud. The Fourth Circuit held that the defendant had waived the privilege when he told the FBI agents that he acted on counsel's advice.[153] Similarly, in *United States v. Mendelsohn*,[154] Mendelsohn was convicted of transporting, aiding, and abetting the interstate

148. *Id.* at 101.

149. *Id.* at 102.

150. *Id.* The court had one caveat: "[I]t is conceivable that assertions before trial may mislead or prejudice an adversary at trial and thereby impede the proper functioning of the judicial system. For that reason plaintiff is entitled to attempt to demonstrate in subsequent proceedings that von Bulow's assertion of his attorney-client privilege is misleading or otherwise prejudicial. At such time, the district court may, in its discretion, reevaluate the scope of petitioner's waiver." *Id.* at 102 n.1.

151. *Id.* at 102–03 (citing Weil v. Investment/Indicators, Research & Mgmt., Inc., 647 F.2d 18 (9th Cir. 1981)) (finding no subject matter waiver from statement made to opposing counsel early in the proceedings and not demonstrably prejudicial to the opposing party).

152. United States v. Under Seal (*In re* Grand Jury Subpoena), 341 F.3d 331 (4th Cir. 2003).

153. *Id.* at 336.

154. United States v. Mendelsohn, 896 F.2d 1183 (9th Cir. 1990).

sale of bookmaking paraphernalia after he sold a computer program used to record and analyze sports bets to an undercover agent. Mendelsohn told the agent that his attorney had said that selling the software was legal, but later qualified his statement, saying that the attorney had not opined about interstate sales. The Fifth Circuit ruled that the attorney could be compelled to testify at trial because Mendelsohn's statement to the agent was a waiver of the privilege.[155]

The Second Circuit has refused to find waiver based on disclosures to the government in the context of an investigation, in part because such statements are extra-judicial. In *In re Grand Jury John Doe Co.*,[156] the "John Doe Company" was the target of an investigation into alleged firearms law violations arising from the company allowing firearms sellers to use its facilities without the company having the necessary licenses. The Company's attorneys submitted a forty-six page letter to the United States Attorney's Office professing its willingness to cooperate. Counsel argued that the company had acted in the good faith belief that its actions were legal based upon consultations with identified Alcohol, Tobacco, and Firearms ("ATF") personnel, who (counsel claimed) had repeatedly advised the company that Doe did not need a license. Counsel also stated that nothing in the letter was "intended to waive any applicable privilege or protection available under law." When the U.S. Attorney's Office sought the attorneys' notes of conversations with the ATF officials and the attorneys' interviews of employees about their conversations with those officials,[157] the trial court and the appellate court held that whether there was an implied waiver, and the scope of any such waiver, were functions of if Doe's use of the privilege created an unfair disadvantage to an opposing party, and that the answer to that question depended very much on the context:

> The crucial issue is whether the unfairness affects the result of the judicial process, when a party uses an assertion of fact to influence the decision-maker while denying its adversary access to privileged material potentially capable of rebutting the presumption.[158]

155. *Id.* at 1189. The district court, affirmed by the circuit, did not find a full subject matter waiver, but did find a waiver to the extent Mendelsohn had purported to disclose specific advice. United States v. Plache, 913 F.2d 1375, 1379–80 (9th Cir. 1990) (asserting reliance on counsel's advice in grand jury testimony constituted waiver of privilege and requiring attorney to testify at trial).

156. John Doe Co. v. United States (*In re* Grand Jury John Doe Co.), 350 F.3d 299 (2d Cir. 2003).

157. *Id.* at 301.

158. *Id.* at 306.

The Second Circuit found no implied waiver because there was no unfairness to an adversary, as the prosecutors were free to disbelieve the company and continue to investigate.[159]

The Fourth Circuit's *In re Martin Marietta* decision explains the justification for implied waiver outside of the courtroom. In that case, the Fourth Circuit held that the use of privileged information to persuade the government to enter into a settlement was *testimonial use* of privileged information and affected a full subject matter waiver. The fact that no litigation had commenced was not relevant.[160]

In *Martin Marietta*, one of the company's former employees, William Pollard, was indicted for a scheme in which Department of Defense ("DOD") travel cost rebates were improperly accounted for as fees instead of credits, with the result that Martin Marietta's subsidiary was able to overbill the DOD.[161] Martin Marietta entered into a settlement agreement with the government that included a guilty plea to reduced charges.[162] In the course of the settlement negotiations, Martin Marietta submitted a fifteen page position paper to the government in which it made representations about the testimony that its employees would give and about the evidence that it had found. Privileged information from files and interviews formed one basis for the representations to the government by Martin Marietta. After the settlement agreement was reached, Pollard was indicted. In support of his defense that he was made a scapegoat by his former employer, Pollard subpoenaed fifteen categories of documents, including all witness statements taken by the company's counsel and counsel's notes of meetings with the government.

159. *Contra* Joy v. North, 692 F.2d 880, 893–94 (2d Cir. 1982) (holding that the use of special litigation committee report to move to terminate derivative litigation waives the privilege regarding the underlying material).

160. United States v. Pollard (*In re* Martin Marietta Corp.), 856 F.2d 619, 623–24 (4th Cir. 1988); *see also In re* Royal Ahold N.V. Sec. & ERISA Litig., 230 F.R.D. 433 (D. Md. 2005) (finding waiver of attorney-client privilege and non-opinion work product as to the subject matter disclosed in a Form 20-F filed with the SEC and investigation reports produced to the SEC and DOJ in a securities fraud class action). Other jurisdictions have ruled similarly in the context of disclosures made during SEC investigations. *See, e.g.*, SEC v. Microtune, Inc., 258 F.R.D. 310, 317 (N.D. Tex. 2009) (finding that voluntary disclosure of documents and results concerning company's internal investigation of stock option practices to the SEC and other third parties during a government investigation resulted in waiver of the attorney-client privilege with respect to all documents "relating to the internal investigation").

161. *Pollard*, 856 F.2d at 620.

162. *Id.*

The question was the extent of the implied waiver. Deciding the scope of the waiver, the Fourth Circuit panel was comfortable stating at the time that "[m]ost courts continue to state the rule of implied waiver in absolute form—any disclosure of a confidential communication outside a privileged relationship will waive the privilege as to all information related to the same subject matter . . . ."[163] The Fourth Circuit found implied waiver of attorney-client privilege as to all details underlying the position paper as a result of the company's submission to the government during settlement discussions.[164] The Fourth Circuit's application of the concept of "testimonial use" of privileged matters recalls the paramount position given to the search for truth: "[The] privilege of secret communication is intended only as an incidental means of defense, and not as an independent means of attack, and to use it in the latter is to abandon it in the former."[165]

In *In re Grand Jury John Doe Co.*, the Second Circuit acknowledged the conflict with the Classic View epitomized by *Martin Marietta*:

> We acknowledge that *Martin Marietta* does not conform to our description of the characteristic circumstances in which factual assertions will cause forfeiture of the privileges of the party making the assertions. The corporation's assertions in that case were not made to an adjudicative authority, as was the case in *Nobles* and *Bilzerian*. They were made only to the U.S. Attorney.

163. *Id.* at 623; *see also* SEC v. Brady, 238 F.R.D. 429, 440–41 (N.D. Tex. 2006) (providing report to the SEC waives privilege); United States v. Woodall, 438 F.2d 1317, 1324 (5th Cir. 1970) (en banc) ("[A] client's offer of his own or his attorney's testimony as to a specific communication constitutes a waiver as to all other communications on the same matter [because] 'the privilege of secret communication is intended only as an incidental means of defense, and not as an independent means of attack, and to use it in the latter character is to abandon it in the former.'") (quoting WIGMORE ON EVIDENCE § 2327 (1961)).

164. *Pollard*, 856 F.2d at 623–24; *see also* United States v. Reyes, 239 F.R.D. 591, 603 (N.D. Cal. 2006) (relying on *Martin Marietta* and finding that counsel's verbal disclosure of the substance of investigation interviews, reports and conclusions to the government resulted in a broad subject matter waiver, stating "it smacks of injustice when, as in this case, the favored party seizes upon the disclosed information to exercise legal leverage against the obstructed one").

165. WIGMORE ON EVIDENCE § 2327 (1961); *see also SEC v. Brady*, 238 F.R.D. at 441. The court also relied on the D.C. Circuit's position rejecting limited waiver. *In re* Subpoenas Duces Tecum, 738 F.2d 1367 (D.C. Cir. 1984) (finding that Tesoro Petroleum provided the SEC with information about illegal foreign bribes in return for more lenient treatment, and that Tesoro shareholders brought a derivative suit seeking the documents in discovery, but rejecting the limited waiver argument).

> At the same time, however, as compared with our facts, the facts of *Martin Marietta* better supported a claim that it would be unfair to the government to allow the corporation to continue thereafter to assert its privileges. In *Martin Marietta*, the government had entered into a global settlement with the corporation in reliance on the corporation's representations. Having irrevocably compromised its legal position in reliance on the corporation's representations, the government could more plausibly assert that the corporation forfeited its legal right to withhold disclosure of privileged documents that might impeach the representations upon which the government relied.[166]

Although it acknowledges the argument that there is some unfairness to the government, the *John Doe* opinion leaves the firm impression that the Second Circuit would have decided *Martin Marietta* differently.[167] However, the distinction the Second Circuit drew between the *In re Grand Jury John Doe Co.* decision and the *Martin Marietta* decision is an awkward one. The court implies that if the government had rejected the positions taken by Martin Marietta in its position paper, there could be no waiver, but that since the government had accepted those arguments, a waiver resulted. In contrast, regardless of the outcome, the Fourth Circuit would focus on the *use* of the privileged communications. Appreciating this contrast is crucial to understanding the conflict between the two views.

It seems clear that the Second Circuit's awkward effort to distinguish the *Martin Marietta* decision should not be given weight. The fact is that the Modern View and the Classic View take different approaches to the issue, and the Second Circuit's position with respect to waiver is not based on reliance, or lack of reliance, by an adversary, but only on the effort to have a final decision-maker rely, whether court, arbitrator, or other tribunal. The Second Circuit stated: "The crucial issue is not merely some connection to a judicial process but rather the type of unfairness to the adversary that results in litigation circumstances when a party uses an assertion of fact to influence the decision maker while denying its adversary ac-

166. John Doe Co. v. United States (*In re* Grand Jury John Doe Co.), 350 F.3d 299, 305 (2d Cir. 2003).

167. After enactment of the amendments to Fed. R. Evid. 502, the issue of what is the waiver as decided in *Martin Marietta* might well change, but the issue of whether a settlement is within the scope of the fairness doctrine, when the regulator can choose not to rely and settle nonetheless, may not change.

cess to privileged material potentially capable of rebutting the assertion."[168]

The differences between the Classic View and the Modern View are illustrated further by decisions from the Fourth and Second Circuits with respect to deposition testimony about confidential communications. In *Hawkins v. Stables*,[169] the Fourth Circuit found that subject matter waiver resulted from the party testifying in a deposition that she had not discussed an illegal wiretap with her counsel. This ruling was based expressly on the Fourth Circuit's refusal to adopt the concept of limited waiver.[170] In direct contrast, the Second Circuit held in *County of Erie* that the defendant's deposition testimony that its counsel had rewritten the policy at issue in the lawsuit did not constitute a waiver, in part because the testimony did not occur "before a decision-maker or fact finder," and the defendants had "not been placed in a disadvantage at trial."[171] Similarly, in *Sims v. Blot*,[172] the Second Circuit held that deposition testimony did not result in a waiver of his psychotherapist-patient privilege, because the testimony did not occur "before a decision-maker or fact finder."[173] There is enormous uncertainty resulting from the clash between the Modern View and the Classic View as to the context in which implied waiver will apply.

The text of Rule 502 contemplates that implied waiver may occur outside of the courtroom. Rule 502 states that "[w]hen the disclosure is made in a Federal proceeding *or to a Federal office or agency* and waives the attorney-client privilege or work-product protection, the waiver extends to an undisclosed communication or information in a Federal or State proceeding only if: the waiver is intentional; the disclosed and undisclosed communications or information concern the same subject matter; and they ought in fairness to be considered together." Had the drafters intended to adopt the Second Circuit's view on the context in which implied waiver might occur, one would expect to see a clear statement to that effect in the Advisory Committee Note. The text of 502(a) certainly contemplates that implied waiver may occur by disclosure to a federal agency outside of an adversary proceeding. The rule and

168. *In Re Grand Jury John Doe Co.*, 350 F.3d at 306.

169. Hawkins v. Stables, 148 F.3d 379, 384 n.4 (4th Cir. 1998).

170. *Id.*

171. Pritchard v. County of Erie, 546 F.3d 222, 230 (2d Cir. 2008) (citing Sims v. Blot, 534 F.3d 117 (2d Cir. 2008)).

172. *Sims*, 534 F.3d at 140.

173. *Id.* (holding that the fairness inquiry focuses on whether there is a risk that a "decisionmaker" will accept the privilege-holder's statements without his opponent having an adequate opportunity to present rebutting evidence).

the Advisory Committee Note with Congress' Statement of Intent unambiguously favor the Modern View, but do not foreclose a continuing struggle between the Classic and Modern Views as to the context in which implied waiver will occur. The Advisory Committee Note to 502(a) reads in part:

> The rule provides that a voluntary disclosure in a federal proceeding *or to a federal office or agency, if a waiver*, generally results in a waiver only of the communication or information disclosed; a subject matter waiver (of either privilege or work product) is reserved for those unusual situations in which fairness requires a further disclosure of related, protected information, in order to prevent a selective and misleading presentation of evidence to the disadvantage of the adversary. *See, e.g.*, *In re* United Mine Workers of America Employee Benefit Plans Litig., 159 F.R.D. 307, 312 (D.D.C. 1994) (waiver of work product limited to materials actually disclosed, because the party did not deliberately disclose documents in an attempt to gain a tactical advantage). Thus, *subject matter waiver is limited to situations in which a party intentionally puts protected information into the litigation in a selective, misleading and unfair manner.*[174]

This portion of the Advisory Committee Note, referencing putting protected information "into litigation" might be read as a reference to adopt *von Bulow* and the cases that follow which hold that disclosure outside of the lawsuit is never a waiver. However, the phraseology is very awkward for that reading. In addition, the original Advisory Committee Note included in the Committee's May 6, 2006 report to the Standing Committee included a cite to the *von Bulow* case, but that reference was removed in the final version of the note.[175] Cases like *Hawkins v. Stables* that find implied waiver

174. FED. R. EVID. 502 advisory committee's note (emphasis added). *In re United Mine Workers of Am. Emp. Benefit Plans Litig.*, 159 F.R.D. 307, 312 (D.D.C. 1994), cited in the note, does not speak to the context in which implied waiver occurs. It speaks only to the extent of the waiver.

175. In the May 6, 2006 report to the Standing Committee the Advisory Committee Note read: "The rule provides that a voluntary disclosure generally results in a waiver only of the communication or information disclosed; a subject matter waiver (of either privilege or work product) is reserved for those unusual situations in which fairness requires a further disclosure of related, protected information, in order to protect against a selective and misleading presentation of evidence to the disadvantage of the adversary." FED. R. EVID. 502 advisory committee's note. In the final version sent to, and adopted by Congress, the reference to *von Bulow* is eliminated: "The rule provides that a voluntary disclosure in a federal proceeding or to a federal office or agency, if a waiver, generally results in a waiver only of the communication or information disclosed; a subject matter waiver (of either privilege or work product) is reserved for those unusual situations in which fairness

only in statements before the fact-finder are clearly not supported by this text. The first clause of this comment section, which references disclosure to a federal agency as reflected by the text of Rule 502, suggests that such a disclosure might result in an implied waiver, with the scope of that waiver to be determined by the fairness doctrine. Read as a whole, the Advisory Committee Note does not distinguish between disclosures governed by federal or state law. This is clearly not the intent of the drafters, who, in resolving the issue of when State or federal privilege law should control, "determined that the proper solution for the federal court is to apply the law that is most protective of privilege and work product."[176]

B. In re Kidder Peabody Securities Litigation *and Arbitrage under the Modern View*

Kidder Peabody Sec. Litig. illustrates the transition to the Modern View and the flexibility (though some would say unpredictability) of available outcomes when using "fairness" as the criterion to determine the scope of waiver.[177] The decision raises questions for the application of Rule 502 and the use of the rule to keep inconvenient facts from government or public scrutiny. The *Kidder Peabody* court, which predates the Second Circuit's direct attack on *Hearn*, and so can be said to be "less modern" than cases that follow, addresses a number of issues that highlight the possible vices of the Modern View, even while continuing to rely on *Hearn*.

In 1994, Kidder Peabody discovered that one of its most prominent traders, Joseph Jett, had allegedly perpetrated a fraud against the firm by which he allegedly created phantom profits totaling hundreds of millions of dollars in an effort to secure larger bonuses and hide losses.[178] When Kidder Peabody announced the discovery, shareholder suits followed quickly against Kidder's parent General Electric Company ("GE").[179] Likewise, the United States Attorney's Office and the SEC launched investigations of GE, Kidder, and Jett.

---

requires a further disclosure of related, protected information, in order to prevent a selective and misleading presentation of evidence to the disadvantage of the adversary."

Note that the final version adds the clause "if a waiver" when describing the effect of a voluntary disclosure, thereby giving some weight to the argument that the Advisory Committee contemplated that some voluntary disclosures would not result in waiver.

176. *Id.*

177. *In re* Kidder Peabody Sec. Litig., 168 F.R.D. 459, 474 (S.D.N.Y. 1996).

178. *Id.* at 461.

179. *Id.* at 463.

Kidder quickly brought an arbitration against Jett.[180] In response, Jett brought a lawsuit to compel Kidder to indemnify him for his defense costs related to the SEC and the criminal probes, and brought an arbitration asserting other claims.[181] Kidder and GE were taking a beating in the press and there were questions as to whether Kidder would survive the scandal.[182] The company moved quickly to retain counsel (former SEC Enforcement Chief Gary Lynch) to conduct an investigation and to defend the various litigations and investigations.[183] Lynch and his team conducted approximately 120 interviews of sixty-five present and former employees.[184] In addition, they looked at a significant number of documents.

Judge Dolinger of the Southern District of New York found that, since the discovery of the scandal, Kidder and GE had engaged in an extensive public relations campaign based largely on Lynch's investigation. In interviews and press releases, Kidder and GE emphasized to the public that they were using the Lynch investigation to determine how the scheme had gone undetected for approximately twenty-eight months and whether others were involved. From the outset, Kidder indicated that the investigation would be shared with the SEC and would probably be made public. In August of 1994, Kidder sent a draft of the Lynch report to the SEC asking for comments. GE then released the final Lynch report to the public, accompanied by announcements of personnel actions and institutional changes to carry out the reforms suggested by the report. The report found that the losses were the work of Jett acting alone. Kidder and GE used the report extensively to influence the U.S. Attorney and SEC investigations.

Kidder invoked attorney-client privilege to protect the interview notes and the report. Plaintiffs and Kidder's co-defendant, Jett, moved to compel Kidder to produce the documents. The court found as a fact that, although Kidder had hired Davis Polk to represent it in bringing the arbitration against Jett and defending lawsuits, Kidder would nonetheless have hired counsel to perform the inquiry if there had been no litigation because Kidder still would have needed to address the business and public relation issues created by the scandal.[185] Interestingly, the court noted that Kidder's intent to use the investigation and the report raised the question of

180. *Id.*
181. *Id.*
182. *Id.*
183. *Id.* at 464.
184. *Id.* at 462.
185. *Id.* at 465.

if any of the investigation materials were ever protected by the privilege, due to the fact that Kidder and GE had intended to publish the investigation results—to the SEC, at a minimum—from the beginning. However, the district court chose to analyze the problem as one of implied waiver.

The court began its analysis with *von Bulow* and the "fairness doctrine." The court found that the publication of the Lynch report was "functionally indistinguishable from the book publication in *von Bulow*." Because both were extra-judicial disclosures, the court concluded that the publication of the Lynch report constituted a waiver only as to the report and those portions of the interview documents that were specifically alluded to or paraphrased in the Lynch report.

Kidder Peabody highlights a vice of the Modern View's limiting of implied waiver to actions in front of the final adjudicator.[186] By limiting an evaluation of fairness to the judicial context, the Modern View allows the client to use the selective disclosure of privileged information to its material advantage and to the disadvantage of a party with adverse or just different interests, as Kidder Peabody did in this case. One may assume that portions of the interview memoranda that underlie the Lynch report were unflattering to Kidder and possibly to its parent, but by allowing Kidder to publish the report without disclosing the underlying information, the Modern View allowed Kidder to go into the public markets and the broader public marketplace and to argue that it had engaged in no wrongdoing, while withholding evidence inconsistent with that position, in a transparent effort to influence civil and criminal federal investigations, the public perception of Jett, and GE's stock price. Kidder and GE consciously sought to have their shareholders, customers, and others (including presumably state or federal regulators who were not actively investigating Kidder) rely. Essentially, Kidder and GE engaged in privilege arbitrage, or disclosing privileged information to the general public where it may prove beneficial and seeking to conceal privileged information in legal proceedings where the information may be harmful. One can ask courts espousing the Modern View what societal good comes from allowing such arbitrage, and how does allowing such arbitrage foster those aspects of the attorney-client relationship that are of value to society.

186. *See In re* von Bulow, 828 F.2d 94, 103 (2d Cir. 1987) ("Although it is true that disclosures in the public arena may be 'one-sided' or 'misleading', so long as such disclosures are and remain extrajudicial, there is no *legal* prejudice that warrants a broad court-imposed subject matter waiver.").

Having decided, however, that the publication of the report to the public did not result in a waiver of the privilege as to the interview memoranda, the court held that the privilege as to the interview memoranda was waived by Kidder's repeated use of the report in *other* litigation and in arbitration against Jett. The district court did not rely on Kidder's use of the report in the litigation before the court. Separately evaluating Kidder's use of the report with the SEC and U.S. Attorney's Office in an effort to persuade them to take no action, the court, relying on *Hearn v. Rhay* and the decision of the D.C. Circuit Court of Appeals in *In re Sealed Case*,[187] held that the use of the report to persuade the SEC and the U.S. Attorney's Office to do nothing was the use of the report in a "litigative context," and that use resulted in a waiver of the attorney-client privilege.[188] Obviously, this holding cannot be reconciled with the Second Circuit's subsequent decisions in *In re Grand Jury John Doe Co.*,[189] *Stable*, and *County of Erie*. Contrasting *Kidder* with *In re Grand Jury John Doe Co.* highlights the vice of the Modern View, which allows a litigant to extensively arbitrage the privilege, so long as the corporation or individual does not do so in a courtroom or in arbitration. Again, the Classic View courts ask how this fosters good attorney-client communications and relations, and at what cost.

The cases display the inherent tension between the attorney-client privilege on the one hand, and confidence that a governmental body, the public, or the court has been given a completely accurate picture of the events and intent of the client, on the other. The assumption of the Modern View is that a court will be able to distinguish between the situations in which privileged information is used fairly and those in which privileged information is used in a selective, misleading, and unfair manner, or to gain technical advantage. The Classic View assumes any limited disclosure is unfair. When that tension results in a misleading disclosure, or a sufficiently unfair disclosure, is not answered by Rule 502.

We should note that under the fairness doctrine, as applied by the Second Circuit, even if the government argued that fairness entitles it to privileged information to test the representations of a company during pre-charging or plea negotiations, a third party would not be able to claim the benefits of implied waiver in order to demand access to privileged information, because that third party would have no fairness argument to make as the result of

187. *In re* Sealed Case, 676 F.2d 793, 818–22 (D.C. Cir. 1982).

188. *In re* Kidder Peabody Sec. Litig., 168 F.R.D. 459, 473 (S.D.N.Y. 1996).

189. The court sidestepped the issue, choosing to use the implied waiver analysis. *Id.* at 468.

statements to the government. In contrast, under the Classic View, the waiver may be asserted by anyone with an interest. In *Martin Marietta*, Pollard, the company's former employee, was not engaged in litigation against the company. Under the Modern View, he would not have had a claim to implied waiver of materials not conveyed to the government; by contrast, under the Classic View, once the privilege had been waived by Martin Marietta's testimonial use of privileged materials with the government, the privilege was lost for all materials on the subject matter as against everyone. In *Kidder,* the court found waiver based upon Kidder's use of the Lynch Report against Jett in other lawsuits, not based upon the federal class action in which the motion was decided. Under the Second Circuit's view, unless the unfairness results in the proceeding in which the privilege is attached, there should be no waiver. *Kidder* illustrates how a Modern View circuit can use the fairness doctrine to arrive at much the same result as the Classic View.

V.
THE SCOPE OF WAIVER AND SUBJECT MATTER WAIVER

The determination that a waiver of privilege has occurred is not the end of the inquiry. The exercise of discretion is inherent in defining what communications a waiver reaches,[190] and, as we have seen, the courts are not uniform in defining the subject matter and scope of a waiver. No precise test has been developed for determining what constitutes "the same subject matter."[191] However, the sub-

190. *See* United States v. Doe (*In re* Grand Jury Proceedings), 219 F.3d 175, 190 (2d Cir. 2000) ("We leave it to the district court, in the first instance, to define the subject matter of the waiver."); Electro Scientific Indus. v. Gen. Scanning, Inc., 175 F.R.D. 539, 544 (N.D. Cal. 1997), *aff'd*, 247 F.3d 1341 (Fed. Cir. 2001) (determining how broad a waiver extends is an issue that the law commits in some measure to the court's discretion); Koster v. Chase Manhattan Bank, No. 81 Civ. 5018, 1984 WL 883, at *3 (S.D.N.Y. Sept. 18, 1984) (citations omitted) ("[I]n view of the policies underlying the concept of waiver and the evident ambiguity of the phrase 'the same subject,' a court necessarily has some discretion in determining the appropriate scope of the waiver.").

191. *See, e.g.*, Rambus, Inc. v. Infineon Techs., AG, 220 F.R.D. 264, 289 (E.D. Va. 2004) ("The Fourth Circuit holds that the scope of the subject-matter waiver rule is limited to other communications relating to the same subject matter. This, of course, leaves open for interpretation what exactly constitutes the same subject matter. Indeed: Despite the centrality of the term, same subject matter, to this inquiry, courts have not defined its meaning and content precisely."); *see also* Jennifer A. Hardgrove, Note, *Scope of Waiver of Attorney-Client Privilege: Articulating a Standard That Will Afford Guidance to Courts*, 1998 U. ILL. L. REV. 643, 662 (1998) (noting that few courts have "developed a specific list of criteria to apply when determining the appropriate breadth of waiver; instead courts seem to analyze

ject matter of a waiver generally includes both a temporal component and a topical component. The time element is frequently overlooked until a litigant faces waiver of materials prepared after commencement of the litigation.[192] But both time and subject are important. Decisions, including many discussed here, suggest that predicting the scope of a subject matter waiver is almost always difficult, given courts' discretion under the Modern View to determine what fairness requires in a given circumstance. According to the Advisory Committee, Rule 502 was not intended to affect the scope of waiver when a party puts a privileged communication into dispute,[193] but one of the rule's primary objectives is to narrow the scope of waiver resulting from disclosure of a portion of the privileged communications on any subject.[194] How the Classic View courts will implement this new instruction, particularly given the room to exercise discretion when a concept as broad as "fairness" defines the scope of the waiver, will determine whether inter-circuit uniformity increases.

each case on an ad hoc basis") (some citations omitted). *See also* Broun & Capra, *supra* note 47, at 16–24 (discussing cases regarding the scope of waiver and determination of the scope); SEC v. Microtune, Inc., 258 F.R.D. 310, 317 (N.D. Tex. 2009) (finding that voluntary disclosure of documents and results of internal investigation of stock option practices to the SEC and other third parties during a government investigation resulted in waiver of the attorney-client privilege with respect to all documents "relating to the internal investigation," without further explanation); SNK Corp. of Am. v. Atlus Dream Entm't Co., 188 F.R.D. 566, 571 (N.D. Cal. 1999) (stating that in determining scope of any subject matter waiver, courts must be "guided by 'the subject matter of the documents disclosed, balanced by the need to protect the frankness of the client disclosure and to preclude unfair partial disclosures' ") (quoting Starsight Telecast v. Gemstar Dev. Corp., 158 F.R.D. 650, 655 (N.D. Cal. 1994)).

192. *In re* Keeper of the Records, 348 F.3d 16, 21–26 (1st Cir. 2003) (rejecting the argument that waiver was to operate without limit of time "so long as people are talking about the same subject" since no ends would be served by implying a broad prospective waiver when disclosure was made in the extrajudicial context); Winbond Elecs. Corp. v. Int'l Trade Comm'n, 262 F.3d 1363, 1375–76 (Fed. Cir. 2001) (holding time period of waiver was appropriate because company's assertion that its understanding of the law changed after a court opinion put at issue the inventor's, and thus its attorneys', understanding of the law before and after the Commission's initial decision). *Contra In re* Buspirone Patent Litig., 210 F.R.D. 43, 53–54 (S.D.N.Y. 2002) (refusing to permit "arbitrary" date cutoff of waiver where advice of counsel defense was raised with respect to antitrust claims); Gabriel Capital, L.P. v. Natwest Fin., Inc., No. 99 Civ. 10488(SAS), 2001 WL 1132050, at *1 (S.D.N.Y. Sept. 21, 2001) (finding prospective waiver effected by adoption of advice-of-counsel defense and communications revealing substance of advice given by counsel to be discoverable, regardless of when they occurred).

193. FED. R. EVID. 502 advisory committee's note.

194. *Id.*; *see also supra* notes 58–62.

A. The Topical Scope of Waiver

The principle adopted by Modern View courts, and by Rule 502 as it applies to waiver by disclosure, is that the subject of a waiver should be interpreted narrowly, so the subject matter information must "be *directly related*" to the disclosed subject. As we have noted, the rule and the Advisory Committee Note are not ambiguous:

> The rule provides that a voluntary disclosure in a federal proceeding or to a federal office or agency, if a waiver, generally results in a waiver only of the communication or information disclosed; a subject matter waiver (of either privilege or work product) is reserved for those unusual situations in which fairness requires a further disclosure of related, protected information, in order to prevent a selective and misleading presentation of evidence to the disadvantage of the adversary.[195]

However, the application of this principle is almost necessarily imprecise.[196] For example, in *United States v. Bilzerian*,[197] the Second Circuit ruled that if Bilzerian testified at trial as to his good faith belief that his actions were legal, he would open the door to cross-examination about privileged communications with his attorney on the subject.[198] Neither the district court nor the circuit court defined the precise scope of the waiver; rather, the Second Circuit ruled that the district court had correctly indicated that the scope of the waiver would be based on Bilzerian's testimony.[199] This left counsel and client faced with tremendous risk as to the scope of the waiver that the court would deem "fair."

There is an enormous amount of practical and theoretical territory between a full subject matter waiver as contemplated in the Fourth Circuit's *Under Seal* (*In re Grand Jury Proceedings*) decision,[200]

195. FED. R. EVID. 502 advisory committee's note; *see also* Church & Dwight Co., Inc. v. Mayer Lab's, Inc., No. C10-4429-EMC (JSC), 2011 U.S. Dist. LEXIS 141315, at *1 (Dec. 8, 2011) (finding waiver as to all documents provided to "a government agency during an investigation . . . even when the documents are produced pursuant to a confidentiality agreement").

196. *E.g.*, Alpex Computer Corp. v. Nintendo Co., Ltd., No. 86 Civ. 1749, 1994 WL 330381, at *2 (S.D.N.Y. July 11, 1994) ("Courts should simply take care to extend the scope of the waiver only so far as necessary to ensure fairness to the litigants.").

197. United States v. Bilzerian, 926 F.2d 1285 (2d Cir. 1991).

198. *Id.* at 1291.

199. *Id.*

200. United States v. Under Seal (*In re* Grand Jury Proceedings), 33 F.3d 342 (4th Cir. 1994).

in which an exhaustive body of material was subject to implied waiver, and cases such as *von Bulow*,[201] which limit waiver to those confidences actually disclosed. This is the territory of a limited subject matter waiver, in which the court, whether using in-camera review or other means, extends the waiver to some, but not all, materials regarding the subject matter, as the prevention of "unfairness" dictates. This imprecision, and the vast territory between the extremes, is the space that the Second Circuit's *In Re Marc Rich & Co. A.G* decision[202] and the Eastern District of Michigan's *Schenet* decision[203] attempt to map. This will also be the space in which Classic View courts and Modern View courts must try to reach some common ground.

Fairness may still include, in certain cases, an implied waiver that extends to the entire transaction where the disclosure is deemed "part-and-parcel" of the whole, and a court deems disclosure necessary to avoid unfairness.[204] *Rambus, Inc. v. Infineon Tech. AG*[205] illustrates a result that one would not expect to survive Rule 502, but it might. The company's disclosure of its document retention policy and various details regarding the development of the policy resulted in a far more extensive waiver. The *Rambus* court considered "several flexible non-exhaustive factors" in determining if communications related to the same subject matter. Among these-factors were:

> (1) [T]he general nature of the lawyer's assignment;
> (2) the extent to which the lawyer's activities in fulfilling that assignment are undifferentiated and unitary or are distinct and severable;
> (3) the extent to which the disclosed and undisclosed communications share, or do not share, a common nexus with a distinct activity;
> (4) the circumstances in and purposes for which disclosure originally was made;
> (5) the circumstances in and purposes for which further disclosure is sought;

201. *In re* von Bulow, 828 F.2d 94 (2d Cir. 1987).

202. Marc Rich & Co. A.G. v. United States, 731 F.2d 1032 (2d Cir. 1984).

203. Schenet v. Anderson, 678 F. Supp. 1280, 1283 (E.D. Mich. 1988).

204. *See, e.g.*, *In re* Wilkerson, 393 B.R. 734 (Bankr. D. Colo. 2008) (holding that implied waiver applied to all communications with bankruptcy attorney concerning disclosure of legal actions, duty to disclose such actions, and asking for information about such actions, including general communications regarding her duty of full disclosure).

205. Rambus, Inc. v. Infineon Techs., 220 F.R.D. 264 (E.D. Va. 2004).

(6) the risks to the interests protected by the privilege if further disclosure were to occur; and

(7) the prejudice which might result if disclosure were not to occur.[206]

The *Rambus* court concluded that the subject matter of the waiver should extend beyond the document retention policy to include the company's patent litigation strategy because "the text of one document rather strongly indicates that Rambus' document retention policy was part-and-parcel of its intellectual property and litigation strategy."[207] The court reasoned: "[B]ecause Rambus has disclosed information respecting the substance of its document retention policy, it has also waived the privilege with respect to those documents that pertain to its patent litigation strategy." Repackaged as fairness with an enumeration of fairness factors, this result is not foreclosed by Rule 502.

The fairness doctrine as articulated in Rule 502 has a narrow waiver as a default position, but even in a Modern View circuit, the fairness doctrine does not *per se* result in a narrow scope of waiver. Deciding the scope of implied waiver arising from an at-issue waiver in 2002 in *In re Buspirone Patent Litigation*, the Southern District of New York wrote: "A party should not be permitted to define selectively the subject matter of the advice of counsel on which it relied in order to limit the scope of the waiver of the attorney-client privilege and therefore the scope of discovery. To do so would undermine the very purpose behind the exception to the attorney-client privilege at issue here—fairness."[208] This position is unaffected by Rule 502, which makes clear that at-issue waiver is not limited by the rule.[209]

206. *Id.* at 289 (citing United States v. Skeddle, 989 F. Supp. 917, 919 (N.D. Oh. 1997)); SNK Corp. of Am. v. Atlus Dream Entm't Co., 188 F.R.D. 566, 571 (N.D. Cal. 1999) (stating that in determining scope of any subject matter waiver, courts must be "guided by 'the subject matter of the documents disclosed, balanced by the need to protect the frankness of the client disclosure and to preclude unfair partial disclosures' ") (quoting Starsight Telecast v. Gemstar Dev. Corp., 158 F.R.D. 650, 655 (N.D. Cal. 1994)).

207. *Rambus, Inc.*, 220 F.R.D. at 290.

208. *In re* Buspirone Patent Litig., 210 F.R.D. 43, 53 (S.D.N.Y. 2002) (citation omitted).

209. FED. R. EVID. 502 advisory committee's note ("The rule governs only certain waivers by disclosure. Other common-law waiver doctrines may result in a finding of waiver even where there is no disclosure of privileged information or work product. The rule is not intended to displace or modify federal common law concerning waiver of privilege or work product where no disclosure has been made.") (citations omitted).

Decisions since enactment of Rule 502 have provided little guidance regarding the impact of the rule on the scope of subject matter waiver. For example, in a 2011 decision, *Enns Pontiac, Buick & GMC Truck v. Flores*, the Eastern District of California found that the defendant's disclosure of a report to the government during a prior investigation did not result in waiver of the entire subject matter in subsequent civil litigation.[210] The court concluded that the waiver was "limited to production of the underlying data referenced or contained in the report" and did not extend to "drafts or notes regarding the report."[211] The court was not persuaded by the plaintiffs' reliance on Rule 502(a), stating only that "[t]he Court questions the applicability of [Rule 502] in this instance. Even applying the statutory provision, however, a broad waiver is not indicated."[212] Similarly, in *SEC v. Berry*, the Northern District of California found waiver as a result of counsel's verbal briefings with the government concerning facts learned through five witness interviews.[213] As counsel "[p]resumably . . . used the final interview memoranda to refresh their recollections when discussing the five Key Witnesses with the government," the court concluded only final interview memoranda were discoverable and that the waiver did not extend to internal notes or draft reports.[214] The *Berry* decision made no mention of Rule 502 and relied on a pre-Rule 502 decision, *United States v. Reyes*.[215] There is no indication from the *Flores* or *Berry* decisions that Rule 502 had any influence on the waiver analysis.

210. Enns Pontiac, Buick & GMC Truck v. Flores, No. 1:07cv01043, 2011 U.S. Dist. LEXIS 77113, at *3 (E.D. Cal. July 13, 2011).

211. *Id.* at *3.

212. *Id.* at *4.

213. SEC v. Berry, No. C07-4431, 2011 WL 825742, at *6 (N.D. Cal. Mar. 7, 2011).

214. *Id.*

215. United States v. Reyes, 239 F.R.D. 591, 604 (N.D. Cal. 2006). *Reyes* similarly involved waiver as a result of government briefings regarding witness interviews. The *Reyes* court held that, to the extent counsel disclosed information contained in written interview summaries, notes, and memoranda, counsel waived any right to protect those writings under the attorney-client or work product privileges. *Id.* Although not addressed in *Berry*, *Reyes* can be construed as resulting in a broader waiver than *Berry* because it did not expressly limit the waiver to final drafts of witness memoranda. Rather than reflecting a shift in the analysis, this may have been due to the fact that the attorneys in *Reyes* did not prepare a final investigation report and retained solely handwritten notes and other unspecified work product relating to the investigation. *Id.* at 596. In the absence of final interview summaries or specific evidence of available writings related to the investigation, it may have been more difficult for the *Reyes* court to delineate the scope of the waiver.

By contrast, in *Bear Republic Brewing Co. v. Central City Brewing Co.*, the District of Massachusetts relied on Rule 502 in finding that the intentional production of work product materials resulted in waiver as to the materials produced and "all the circumstances involved with respect to this material, including how it came to be obtained, at whose direction it was obtained, and the manner in which it was obtained."[216] The court, however, provided no further explanation as to how it arrived at the scope of the subject matter waiver other than consideration of "fairness" as mandated by Rule 502.[217] Similarly, in *US Airline Pilots Assoc. v. Pension Benefit Guaranty Corp.*, the District of Columbia applied Rule 502 and found that reliance on an investigation in litigation and disclosure of counsel's investigative report resulted in a broad waiver as to the subject matter of the report, including the scope and methods of the investigation, the documents reviewed during the investigation, and findings of the investigation.[218] Even when courts apply Rule 502, no clear guidelines have emerged as to determination of the scope of the subject matter waiver.

B. The Temporal Scope of Waiver

The temporal component can have a greater impact on the scope of a waiver than the topic of the waived privileged communications.[219] Courts from a variety of jurisdictions have held that the temporal scope of a waiver should be defined based on considerations of fairness and the materiality of events occurring during the time when the advice put at issue in the case guides the conduct of the client.[220] The cases in which courts have found broad temporal

216. Bear Republic Brewing Co. v. Central City Brewing Co., 275 F.R.D. 43, 49 (D. Mass. 2011).

217. *Id.*

218. US Airline Pilots Assoc. v. Pension Benefit Guaranty Corp., 274 F.R.D. 28, 33 (D.D.C. 2011).

219. *Cf. In re* Seagate Tech., LLC, 497 F.3d 1360, 1365 (Fed. Cir. 2007) (reversing lower court ruling applying waiver through any period of patent infringement, including trial); United States v. Weissman, No. S1 94 CR. 760 (CSH), 1996 WL 751384, at *3 (S.D.N.Y. Dec. 26, 1996) (limiting subject matter waiver to privileged communications regarding the same subject matter that took place on only one date).

220. *In re* Keeper of the Records, 348 F.3d 16, 21–26 (1st Cir. 2003) (rejecting the argument that waiver was to operate without limit of time "so long as people are talking about the same subject," since no ends would be served by implying a broad prospective waiver when disclosure was made in the extrajudicial context); Winbond Elecs. Corp. v. Int'l Trade Comm'n, 262 F.3d 1363, 1375–76 (Fed. Cir. 2001); *In re* Papst Licensing GMBH & Co. KG Litig., 250 F.R.D. 55, 59–60 (D.D.C. 2008) ("In balancing competing interests at issue . . . a temporal limitation is ap-

waivers, including waiver as to communications made after the lawsuit is filed, are generally cases in which an advice-of-counsel defense has been raised[221] and situations where the disclosed communication continues to be placed at issue for the advantage of its proponent.[222] When an advice-of-counsel defense is raised, the temporal scope of waiver may be broader than the more limited waiver imposed when there is disclosure of a particular privileged communication. *Buspirone Patent Litigation*,[223] *Gabriel Capital, L.P. v. Natwest Finance, Inc.*,[224] and *United States v. Weissman*[225] illustrate different approaches. The court in *Weissman*, relying on the fairness doctrine, refused to extend the waiver to communications that took place on a date after the privileged communications had already been disclosed in a deposition, even though the communications dealt with the same topic, because the court did not deem the later communications material to the issues in dispute.[226] By contrast,

---

propriate."); *In re* Commercial Fin. Servs., Inc., 247 B.R. 828, 851–52 (Bankr. N.D. Okla. 2000) ("When communications made between a party and prior counsel have been made available, a party should not be found to have waived the benefits of its privileges with current counsel about the same subject matter unless the communications with current counsel are *at issue* in the litigation."); Katz v. AT&T Corp., 191 F.R.D. 433, 441 (E.D. Pa. 2000) ("It would be fundamentally unfair to extend the scope of the waiver beyond the date of the patent application because the issue of the conception date of an invention and its necessary elements is constrained to the time period prior to the filing of the application."); *Weissman*, 1996 WL 751384, at *3 (refusing to extend the waiver to include privileged communications that took place on a different date than privileged communications already disclosed due to fairness considerations, even though communications dealt with the same topic).

221. *In re Buspirone Antitrust Litig.*, 210 F.R.D. at 53–54 (refusing to permit "arbitrary" date cutoff of waiver where advice-of-counsel defense raised with respect to antitrust claims); Gabriel Capital, L.P. v. Natwest Fin., Inc., No. 99 Civ. 10488, 2001 WL 1132050, at *1 (S.D.N.Y. Sept. 21, 2001) (finding prospective waiver effected by adoption of advice-of-counsel defense discoverable regardless of when they occurred); *see also In re Keeper of the Records*, 348 F.3d at 26 ("Courts have generally allowed prospective waivers in discrete and limited situations, almost invariably involving advice of counsel defenses.").

222. *E.g.*, Bd. of Trustees of the Leland Stanford Junior Univ. v. Roche Molecular Sys., 237 F.R.D. 618, 627 (N.D. Cal. 2006) (holding "once a waiver has occurred, it is inappropriate to limit waiver on a temporal basis" where party "continues to place the issue of inventorship of the [invention] before the PTO as it files continuation and divisional applications [derived] from the [invention]").

223. *In re Buspirone Patent Litig.*, 210 F.R.D. at 53–54.

224. *Gabriel Capital, L.P.*, 2001 WL 1132050, at *1.

225. *Weissman*, 1996 WL 751384, at *3.

226. *Id.* at *3. Similarly, in *In re Commercial Fin. Servs., Inc.*, 247 B.R. 828, 848–52 (Bankr. N.D. Okla. 2000), the court refused to extend a proposed privilege waiver to communications with current bankruptcy counsel because the communications that the debtor proposed to disclose were made prior to the engagement

defendant BMS in *Buspirone* was considering raising an advice-of-counsel defense against antitrust claims. Plaintiffs alleged that, as part of its antitrust activity, BMS had instituted patent infringement actions against them in order to trigger an automatic stay of the Food and Drug Administration's approval of their generic products for up to 30 months. BMS sought to limit the temporal scope of any waiver to privileged communications occurring prior to BMS's filing the patent infringement lawsuit, but the court refused, again relying on fairness considerations and the *possible* relevancy of communications with counsel after institution of the suits. The court reasoned:

> It is also clear that BMS cannot in fairness try to tailor an advice-of-counsel defense in an unfairly prejudicial way, so as to provide all the documents relating to the advice of counsel up until a certain arbitrary date, which may tend to support BMS's positions in this litigation, while categorically excluding any such documents thereafter, which may tend to show that BMS thereafter pursued the challenged conduct in bad faith or against the advice of counsel.
>
> The Magistrate Judge was also correct that some of the allegedly illicit conduct in this case spanned into the period in which BMS was prosecuting its patent infringement claims, and that some documents from this period would in fairness likely need to be subject to discovery to test any factual assertions BMS might make in raising an advice-of-counsel defense. There are clearly circumstances in which documents that . . . were created after the patent infringement actions were filed would be relevant to assessing the factual basis for a reliance-on-counsel defense or to placing the disclosures that were made to BMS into an accurate context.[227]

These precedents, as they are outside the ambit of Rule 502, may not affect post-502 decisions, but the "fairness" principles that they espouse will appeal to Classic View courts. However, the Classic View approach to fairness is not necessarily boundless subject matter waiver. For instance, a growing number of courts have followed the view espoused by the Federal Circuit in *In re Seagate Technology*,

of current counsel and "fairly disclose[d] facts regarding the debtor's operations and financial status during the time periods relevant to creditors and investors interested in the case." Communications with current counsel were not at issue in any litigation and the debtor was not seeking to use the privileged communications to gain a tactical advantage. *Id.* at 849.

227. *In re Buspirone Patent Litig.*, 210 F.R.D. at 53–54.

LLC,[228] that "as a general proposition, asserting the advice of counsel defense and disclosing opinions of patent counsel do not constitute waiver of the attorney-client privilege for communications with trial counsel."[229] The Federal Circuit, which has espoused the Classic View in the past,[230] concluded that the interference with the trial counsel relationship, balanced against the risk of loss of relevant evidence, weighed in favor of limiting waiver to non-trial counsel.

CONCLUSION

Contrasting the cases under the Modern View and the Classic View, one can foresee a court articulating a fairness justification for

228. *In re* Seagate Tech., LLC, 497 F.3d 1360, 1372 (Fed. Cir. 2007).

229. *Id.* at 1374–76. *Compare* Akeva LLC v. Mizuno Corp., 243 F. Supp. 2d 418, 422–23 (M.D.N.C. 2003) (assertion of advice of counsel defense against willful infringement waived privilege as to "all opinions received by the client relating to infringement . . . even if they come from defendants' trial attorneys, and even if they pre-date or post-date the advice letter of opinion counsel"), *with* Parker-Hannifin Corp. v. Seiren Co., No. Civ. 07-104-MPT, 2009 WL 840594, *4 (D. Del. Mar. 31, 2009) (citing *In re Seagate Tech., LLC* for the proposition that waiver of attorney-client privilege and work product generally does not extend to trial counsel); New River, Inc. v. Newkirk Products, Inc., No. 1:08-MISC-61(GLS/RFT), 2008 WL 5115244, at *2–3 (N.D.N.Y. Dec. 4, 2008) ("It is clear that the scope of discovery is relegated to pre-litigation and gaining trial counsel's opinions is not relevant. The accuser needs to focus on whose opinion the defendant is relying upon and then needs to separate from all other privileged communications those communications made and created before litigation, which were shared with the alleged infringer.") (citation omitted); Reedhycalog UK, Ltd. v. Baker Hughes Oilfield Operations Inc., 251 F.R.D. 238, 241–42 (E.D. Tex. 2008) (citing *In re Seagate Technology, LLC* for proposition that generally, disclosure of opinions of opinion counsel do not constitute waiver of attorney-client privilege as to communications between the client and trial counsel or work product of trial counsel); SPX Corp. v. Bartec USA, LLC, 247 F.R.D. 516, 525–27 (E.D. Mich. 2008) (acknowledging the holding *In re Seagate Technology, LLC* that communications with trial counsel that post-date the filing of the lawsuit most likely have little bearing on the question of willfulness or reliance upon counsel's advice to shape accused pre-suit conduct, but noting that communications with trial counsel that occurred *befor*e the lawsuit was filed are not presumptively beyond the scope of relevancy identified by the *In re Seagate Technology, LLC* court); Convolve, Inc. v. Compaq Computer Corp., 2007 WL 4205868, at *5 (S.D.N.Y. Nov. 26, 2007) ("Whether the waiver applies to post-litigation communications by in-house counsel concerning their opinions is a closer question. . . . However, in *Seagate*, the court found that post-litigation opinions even of outside opinion counsel are of 'marginal value' because the willfulness analysis focuses on pre-litigation conduct. Given that reasoning, there is no basis here for taking discovery of in-house counsel's communications after the litigation was commenced.").

230. *E.g.*, Fort James Corp. v Solo Cup Corp., 412 F.3d 1340, 1349 (Fed. Cir. 2005).

the decisions taken under the Classic View. This has at least two consequences. First, courts that have adopted the Classic View may proceed under Rule 502 with their approach to waiver unchanged, while using the nomenclature of fairness. Second, whatever the court, the fairness test is so flexible that Rule 502 may bring less certainty to the application of the attorney-client privilege than the drafters and Congress had hoped, at least when the question concerns waiver arising from an intentional disclosure of a portion of privileged communications or information.

Even after the enactment of Rule 502, application of the attorney-client privilege may remain inconsistent and uncertain. This uncertainty undermines the attorney-client relationship, because clients and counsel remain unsure of whether and when their communications will be protected. The inconsistency, and the uncertainty that it creates, can only be resolved by a federal law that demands uniformity. Ultimately, a uniform body of federal privilege law should provide certainty as to what will and will not be considered privileged.

AVOIDING "MAGIC MIRRORS"[1]—A POST-*PADILLA* CONGRESSIONAL SOLUTION TO THE 28 U.S.C. § 2254 "CUSTODY" AND "COLLATERAL" SANCTIONS DILEMMA

*RUTH A. MOYER**

1. Lefkowitz v. Fair, 816 F.2d 17, 20 (1st Cir. 1987).

* Associate, Gerald A. Stein, P.C.; Temple University Beasley School of Law (J.D., *cum laude*, 2008). The author expresses her sincere appreciation to Professor Michael E. Libonati, Professor Louis M. Natali, and Professor James A. Strazzella of the Temple University Beasley School of Law for their kind support. The author is immensely grateful to the Honorable Joseph J. O'Neill (Philadelphia Municipal Court) for his valuable encouragement with her experience representing criminal defendants in misdemeanor trials. The author also truly thanks Assistant District Attorney Richard J. Sax (Philadelphia District Attorney's Office, Homicide Unit) for his criminal justice perspective. Finally, the author wishes to acknowledge the assistance of the Temple Law School library.

INTRODUCTION

In the winter of 1806, Chief Justice John Marshall termed the writ of habeas corpus a "great Constitutional privilege."[2] He observed that the First Congress, in drafting the Judiciary Act of 1789, "must have felt, with peculiar force, the obligation of providing efficient means by which [the writ] should receive life and activity; for if the means be not in existence, the privilege itself would be lost."[3] Throughout the more than two centuries that have passed since Chief Justice Marshall's characterization of the "Great Writ," Congress and the federal courts have demonstrated a remarkable ability to adapt habeas corpus to the equally compelling demands of individual liberty and governmental authority.

One important facet of contemporary federal habeas relief is that those convicted of state crimes may challenge in federal court the validity of their convictions under the U.S. Constitution. Congress has codified this federal court review at 28 U.S.C. § 2254.[4] As an essential jurisdictional element, § 2254 requires that a petitioner seeking federal habeas relief must be in "custody" pursuant to a state court judgment at the time that he files the petition.[5]

Federal courts have repeatedly instructed that § 2254 "custody" does not require "physical restraints." As a general matter, a petitioner satisfies the § 2254 custody requirement when he is under "some type of continuing governmental supervision" and is "subject

2. *Ex parte* Bollman, 8 U.S. 75, 95 (4 Cranch) (1807).
3. *Id.*
4. *See infra* Part I.B for a general discussion of habeas relief.
5. *See infra* Part I.C for a discussion of the "custody" requirement.

to restraints not shared by the public generally" that significantly restrict his liberty.[6]

The increasing use of "collateral" sanctions as an integral part of criminal penalties, however, has rendered the "custody" requirement far more convoluted than it has been in previous decades.[7] "Collateral" consequences are generally not an explicit component of the punishment that the sentencing court imposes; they result from the defendant's convicted status, rather than from the sentence itself. Collateral sanctions include, for example, deportation, loss of public benefits, deprivation of the right to vote, and the loss of the right to engage in certain occupations. Direct consequences, by contrast, include those sanctions resulting directly from a conviction, such as imprisonment, probation, or fines.

In 1968, the Supreme Court held in *Carafas v. LaVallee* that these collateral consequences may prevent a § 2254 petition from becoming moot—so long as the petitioner filed it while he was "in custody."[8] In its 1989 *Maleng v. Cook* decision, however, the Court held that these same consequences do not by themselves satisfy the requirement that the § 2254 petitioner be in custody at the time that he files his petition.[9] As the First Circuit has noted, "There are no magic mirrors: even grievous collateral consequences stemming directly from a conviction cannot, without more, transform the absence of custody into the presence of custody for the purpose of habeas review."[10] Consistent with this formalistic custody framework,[11] courts do not differentiate between those collateral consequences that are potentially serious—such as deportation—and those that are likely trivial—such as anxiety resulting from a conviction.

Importantly, state convicts may not challenge their sentence through a § 2254 petition until they have exhausted all available state court remedies.[12] This rule can create difficulties for state convicts who are no longer in custody by the time that they have

6. *See infra* Part I.C.1 for a discussion how the federal courts devised this definition of "custody."

7. *See infra* Part I.A for a discussion of "collateral" sanctions.

8. 391 U.S. 234, 237–38 (1968). *See infra* Part I.C.2 for a discussion of how "collateral" sanctions prevent a § 2254 petition from becoming "moot."

9. 490 U.S. 488, 492 (1989). *See infra* Part I.C.3 for a discussion of authority holding that "collateral" sanctions do not sufficiently establish "custody."

10. Lefkowitz v. Fair, 816 F.2d 17, 20 (1st Cir. 1987).

11. *See infra* Part II.A for a discussion of the *Carafas/Maleng* paradox.

12. *See infra* notes 80–85 and accompanying text for a brief discussion of the exhaustion requirement.

finished exhausting state remedies.[13] To illustrate, a state court may impose a sentence of six months of incarceration followed by two years of probation on a defendant convicted of a misdemeanor. Due to a crowded appellate court docket, however, the entire state direct appeal and post-conviction review process takes five years. Under these circumstances, the defendant would no longer satisfy the traditional conception of custody so as to permit him to file a § 2254 petition. Thus, contrary to the axiomatic justification for § 2254 review, "short-sentence" state convicts will rarely have the opportunity to challenge their convictions in federal court, regardless of (1) the underlying merit of their federal constitutional claims, and (2) any significant collateral sanctions that they may suffer as a result of their now-expired sentence. In short, under the current legal framework, these collateral consequences are sufficient to prevent mootness, but fail to establish custody where the sentence has already expired.

Collateral sanctions have created equally vexing issues within the Sixth Amendment ineffective assistance of counsel context.[14] In both the § 2254 custody and the Sixth Amendment contexts, courts have often relied upon a formalistic distinction between the direct and collateral results of a conviction.[15] Just as *Maleng* held that collateral consequences were insufficient to establish § 2254 custody,

13. *See infra* Part II.B for a discussion of this difficulty.

14. The Supreme Court has created a two-part test that a criminal defendant must satisfy in order to prove that his attorney provided "ineffective assistance" of counsel in violation of the Sixth Amendment of the U.S. Constitution. *See, e.g.*, Strickland v. Washington, 466 U.S. 668, 687 (1984). First, the defendant must demonstrate that his attorney's performance failed to satisfy an "objective standard of reasonableness." *Id.* at 687–88. Second, the defendant must demonstrate that a reasonable probability exists that if his attorney had performed adequately, the result of the proceeding would have been different. *Id.* at 694.

15. A considerable amount of scholarly discussion concerning "collateral consequences" has centered on the issue of whether an attorney, in order to provide constitutionally effective assistance of counsel, must advise his client of the collateral consequences of a guilty plea or conviction. *See, e.g.*, Gabriel J. Chin & Richard W. Holmes, Jr., *Effective Assistance of Counsel and the Consequences of Guilty Pleas,* 87 CORNELL L. REV. 697, 698 (2001); Michael Pinard, *An Integrated Perspective on Collateral Consequences of Criminal Convictions and Reentry Issues Faced by Formerly Incarcerated Individuals*, 86 B.U. L. Rev. 623, 633 (2006) [hereinafter Pinard, *Collateral Consequences*]; Jenny Roberts, *Ignorance is Effectively Bliss: Collateral Consequences, Silence, and Misinformation in the Guilty-Plea Process,* 95 IOWA L. REV. 119, 147 (2009) [hereinafter Roberts, *Guilty-Plea Process*] (arguing that "exploring the potential collateral consequences of a particular guilty plea" with defendant may be equally important as "basic investigation" of the case itself). Many of these observers have argued that the Supreme Court should "overcome the mythical divide between direct and collateral consequences." Roberts, *Guilty-Plea Process, supra*, at 120.

courts had held that an attorney's failure to advise a client about the collateral consequences of a conviction was insufficient to establish "constitutionally deficient counsel" under the Sixth Amendment.[16] The Ninth Circuit's decision in *Resendiz v. Kovensky* illustrates the analogous relationship between these two concepts.[17] Holding that the collateral sanction of deportation was insufficient to establish § 2254 custody, the Ninth Circuit cited the then-prevailing rule that "the failure of counsel to advise his client of the potential [collateral] immigration consequences of a conviction does not violate the Sixth Amendment right to effective assistance of counsel."[18]

Yet, in its pivotal 2010 *Padilla v. Kentucky* decision, the Supreme Court held that criminal defense attorneys must advise non-citizen clients about the deportation risks of guilty pleas.[19] In reaching this conclusion, *Padilla* emphasized the seriousness of deportation as a collateral consequence and noted that it is "intimately" related to the criminal process. The Court also remarked that the "collateral versus direct distinction" was "ill-suited" to determining whether an attorney provided constitutionally sufficient assistance of counsel under the Sixth Amendment. Applying and extending *Padilla*, some federal and state courts have held that attorneys are constitutionally ineffective if they fail to advise their clients about collateral consequences other than those related to immigration.[20]

By elevating collateral sanctions to the level of Sixth Amendment scrutiny, *Padilla* represents a paradigm shift in how courts view collateral sanctions. This Article argues that the *Padilla* analysis can be applied to the analogous issue of collateral sanctions as a

16. *See, e.g.,* Broomes v. Ashcroft, 358 F.3d 1251, 1256 (10th Cir. 2004) (holding that counsel was not ineffective for failing to advise client of possibility of deportation as collateral sanction of criminal conviction), *abrogated by* Padilla v. Kentucky, 130 S. Ct. 1473 (2010).

17. 416 F.3d 952, 956–57 (9th Cir. 2005).

18. *Id.* at 957.

19. 130 S. Ct. 1473, 1478 (2010). *See infra* Part II.D.1 for an overview of the *Padilla* decision.

20. See *infra* Part I.D.2 for an overview of how courts have extended the *Padilla* holding in the Sixth Amendment "ineffective assistance of counsel" context. *See also* Webb v. State, 334 S.W.3d 126, 138 (Mo. 2011) (Wolff, J., concurring) (explaining why *Padilla* should apply to collateral consequences other than those related to immigration); Travis Stearns, *Intimately Related to the Criminal Process: Examining the Consequences of a Conviction after* Padilla v. Kentucky *and* State v. Sandoval, 9 SEATTLE J. SOC. JUST. 855, 863–64 (2011) (arguing that *Padilla* analysis should apply to "other consequences").

basis for § 2254 custody.[21] As a result of the *Padilla* paradigm shift, federal courts may eventually decide that some collateral sanctions also sufficiently establish custody for purposes of § 2254 claims.

Nevertheless, an indiscriminate expansion of the definition of "custody" based on the existence of collateral sanctions would prove to be problematic. This Article recognizes that an inappropriately broad definition of § 2254 custody frustrates not only the inherently limited purpose of federal habeas review, but also the compelling societal interest in finality.[22] On a superficial level, when applied to the custody context, the *Padilla* analysis seems to create the undesirable result of an overly expansive meaning of § 2254 custody. When examined more critically and in its proper context, however, the *Padilla* analysis provides a necessary new perspective on the "collateral" sanctions and § 2254 custody dilemma.

Thus it is necessary to resolve the arduous question of which collateral sanctions are substantial enough to constitute custody. Some scholars have contended that "just how 'substantial' restraints must be . . . is a question of degree and one that ultimately falls to the judiciary.' "[23] However, this Article contends that it is Congress—and not the federal judiciary—that must ultimately provide statutory guidance for determining whether a collateral consequence is sufficient to establish custody under § 2254. Under Article III of the U.S. Constitution, Congress possesses the authority to determine the jurisdiction of the federal courts. Given the inherently jurisdictional nature of custody, Congress should provide a statutory framework to aid courts in determining whether a collateral sanction sufficiently establishes custody under § 2254. Indeed, since the Judiciary Act of 1789, Congress has exercised a vital role in remedying disagreements concerning the scope and availability of federal habeas review.[24] Furthermore, considered as a collective whole, the collateral consequences that may result from convictions in each of the fifty states are numerous and vary widely in their degree of severity. Therefore, clear Congressional guidance would have the practical effect of ensuring uniformity in federal case law concerning custody and collateral consequences. Consistent with

21. *See infra* Part III.C for a discussion of how the *Padilla* analysis abandons formalistic reliance on the distinction between "collateral" and "direct" consequences of a conviction.

22. *See infra* Part II.C for a discussion of this societal interest.

23. Wayne A. Logan, *Federal Habeas in the Information Age*, 85 MINN. L. REV. 147, 208 (2000).

24. *See infra* Part III.A for a discussion of why Congressional intervention is necessary.

the *Padilla* analysis and traditional understandings of custody, this Article proposes statutory language that both avoids a formalistic reliance on the direct and collateral distinction, and enables federal courts to meaningfully determine whether § 2254 custody exists.

Pursuant to this proposed statutory solution, the § 2254 petitioner whose sentence has expired bears the burden of establishing that his purported collateral harm satisfies the custody requirement. Drawing upon the *Padilla* language as well as established § 2254 custody jurisprudence,[25] the statutory proposal provides factors relevant to determining whether a collateral sanction sufficiently establishes custody. The petitioner must demonstrate that he is subject to "significant restraints not shared by the public generally" and that "governmental action" created these restraints. In determining whether a petitioner has established "significant restraints," courts should consider: (1) the permanency of the restraints, (2) the degree to which the restraints are "intimately related to the criminal process," (3) the relative severity of the restraint to the petitioner, (4) the degree to which the severity of the restraint exceeds the severity of the sentence itself, and (5) the availability of relief from the restraint through means other than § 2254 review.[26]

In contrast to the underlying premise of this Article, some federal district courts have squarely rejected the contention that *Padilla* alters the habeas "custody" analysis.[27] Most existing scholarship, however, has asserted that the collateral sanctions of offender registration and immigration satisfy the definition of custody.[28] Yet, as a critical matter, this scholarship not only predates

25. *See infra* Parts III.B.1–4 for a discussion of how traditional definitions of "custody" would affect proposed statutory analysis.

26. *See infra* Part III.D for this proposed statutory solution.

27. *See, e.g.*, Fenton v. Ryan, No. 11–2303, 2011 WL 3515376, at *2 (E.D. Pa. Aug. 11, 2011) (holding that *Padilla* did not "expand" custody requirement to "encompass" conviction's immigration consequences). *See infra* Part I.D.3 for a discussion of case law declining to extend *Padilla* to the custody context.

28. *See, e.g.*, Tina D. Santos, Note, Williamson v. Gregoire: *How Much is Enough? The Custody Requirement in the Context of Sex Offender Registration and Notification Statutes,* 23 SEATTLE U. L. REV. 457, 459 (1999) ("The registration and notification provisions operate to constructively restrain the liberty of a convicted sex offender and, therefore, [render him] 'in custody' for purposes of habeas corpus relief."); Kerri L. Arnone, Note, *Megan's Law and Habeas Corpus Review: Lifetime Duty with No Possibility of Relief?,* 42 ARIZ. L. REV. 157, 158-59 (2000); Garrett Ordower, Comment, *Gone But Not Forgotten? Habeas Corpus for Necessary Predicate Offenses,* 76 U. CHI. L. REV. 1837, 1873 (2009); Joshua D. Smith, Comment, *Habeas Corpus: Expired Conviction, Expired Relief: Can the Writ of Habeas Corpus Be Used to Test the Constitution-*

the watershed *Padilla* decision but also embraces an inapt willingness to automatically establish custody merely because a petitioner encounters a collateral sanction.

Two hundred years after Chief Justice Marshall's assertion, federal courts and Congress must now confront which restraints can enable a state offender to claim the benefit of the "great Constitutional privilege" of federal habeas review. Consistent with their historic dialogic relationship in the federal habeas context, the legislative and judicial branches should collaboratively resolve whether a given collateral restraint constitutes § 2254 custody.

I.
OVERVIEW OF EXISTING LAW

Determining the appropriate extent to which § 2254 petitioners satisfy the custody requirement based on collateral consequences requires a basic understanding of these sanctions as well as federal habeas law.

A. Collateral Sanctions of Convictions

Criminal convictions result in two general types of consequences: direct and collateral.[29] However, the boundaries of these categories are not always apparent.[30] Even though collateral consequences have become an increasingly key component of the criminal process,[31] courts often disagree on how to precisely distinguish between direct and collateral consequences.[32]

As a general matter, direct consequences are largely limited to the penal sanction that will be imposed as a result of a conviction,

ality of a Deportation Based on an Expired Conviction?, 58 OKLA. L. REV. 59, 61 (2005); Logan, *supra* note 24, at 147 (arguing that laws mandating that sex offenders provide information to law enforcement satisfy custody requirement).

29. Jenny Roberts, *The Mythical Divide Between Collateral and Direct Consequences of Criminal Convictions: Involuntary Commitment of "Sexually Violent Predators,"* 93 MINN. L. REV. 670, 678 (2008) [hereinafter Roberts, *The Mythical Divide*].

30. *Id.* at 678–80 ("It is . . . far from clear exactly where the line between direct and collateral consequences falls.").

31. Chin & Holmes, *supra* note 15, at 699.

32. Padilla v. Kentucky, 130 S. Ct. 1473, 1481 n.8 (2010). Some observers have asserted that the "doctrine draws a sharp but false distinction between 'direct' consequences of criminal proceedings (such as incarceration) and 'collateral' consequences (such as deportation)." McGregor Smyth, *From "Collateral" to "Integral": The Seismic Evolution of* Padilla v. Kentucky *and Its Impact on Penalties Beyond Deportation*, 54 HOW. L.J. 795, 796 (2011).

such as imprisonment, probation, or fines.[33] Direct consequences may also include parole eligibility in some jurisdictions.[34] Collateral consequences, on the other hand, are the penalties or civil disabilities that indirectly result from a conviction and that do not constitute part of the "explicit punishment handed down by the court; they stem from the fact of conviction rather than from the sentence of the court."[35] To illustrate, immigration consequences of a state conviction are deemed collateral because they "arise from the action of an independent [federal] agency—indeed, in the case of a state conviction, an independent *sovereign*—and are consequences over which the state trial judge has no control whatsoever."[36] Often functioning as a "secret sentence,"[37] collateral consequences encompass the "vast network" of civil sanctions that constrain a convicted person's "social, economic, and political access."[38] Many collateral consequences are codified in a variety of federal and state statutes and regulations.[39]

These collateral consequences may include, for example, the loss of the right to vote, to engage in some occupations, to hold

33. Roberts, *The Mythical Divide*, *supra* note 29, at 679–80; Stearns, *supra* note 20, at 869.

34. Pinard, *Collateral Consequences*, *supra* note 15, at 634 ("Direct consequences include the duration of the jail or prison sentence imposed upon the defendant as well as, in some jurisdictions, the defendant's parole eligibility or imposition of fines.").

35. *Id.*; Arnone, *supra* note 28, at 166 n.80; *see also* North Carolina v. Rice, 404 U.S. 244, 247 (1971) ("A number of disabilities may attach to a convicted defendant even after he has left prison . . . ."); Michael Pinard, *Broadening the Holistic Mindset: Incorporating Collateral Consequences and Reentry into Criminal Defense Lawyering*, 31 FORDHAM URB. L. J. 1067, 1074 (2004) [hereinafter Pinard, *Broadening the Holistic Mindset*] (noting that collateral consequences are "legally separate" from the criminal sentence).

36. Resendiz v. Kovensky, 416 F.3d 952, 957 (9th Cir. 2005).

37. Chin & Holmes, *supra* note 15, at 700.

38. Pinard, *Collateral Consequences*, *supra* note 16, at 634–35; *see also* Fiswick v. United States, 329 U.S. 211, 222 (1946) (noting that because of a criminal conviction, an individual would "carry through life the disability of a felon; and by reason of that fact he might lose certain civil rights"). For a helpful overview of the significance of collateral consequence in contemporary society, see generally Michael Pinard, *Reflections and Perspective on Reentry and Collateral Consequences*, 100 J. CRIM. L. & CRIMINOLOGY 1213 (2010) [hereinafter Pinard, *Perspective on Reentry*].

39. Webb v. State, 334 S.W.3d 126, 138 (Mo. 2011) (Wolff, J., concurring) (noting that "consequences can be broad-ranging and scattered" throughout both federal and state statutes); Roberts, *The Mythical Divide*, *supra* note 30, at 678; *see also* Chin & Holmes, *supra* note 15, at 704.

public office, or to serve as a juror, as well as the possibility of deportation.[40] As Justice Samuel Alito has recently observed:

> [C]riminal convictions can carry a wide variety of consequences other than conviction and sentencing, including civil commitment, civil forfeiture, the loss of the right to vote, disqualification from public benefits, ineligibility to possess firearms, dishonorable discharge from the Armed Forces, and loss of business or professional licenses. A criminal conviction may also severely damage a defendant's reputation and thus impair the defendant's ability to obtain future employment or business opportunities.[41]

Although governmental entities acting independently of the criminal justice system impose some collateral sanctions as a matter of discretion, other collateral sanctions "attach automatically upon the conviction by operation of law."[42] For example, after a physician is convicted of a crime, the state agency responsible for issuing physicans' licenses may choose to revoke his license.[43] In many jurisdictions, a felony conviction automatically bars a defendant from possessing a firearm.[44] Also, collateral consequences accompany misdemeanor as well as felony convictions.[45] As one scholar has observed, "[M]isdemeanor convictions can render defendants ineligible for several employment related licenses. Perhaps most critically,

40. Arnone, *supra* note 28, at 166 n.80; McGregor Smyth, *Holistic is Not a Bad Word: A Criminal Defense Attorney's Guide to Using Invisible Punishments as an Advocacy Strategy*, 36 U. TOL. L. REV. 479, 481 (2005).

41. Padilla v. Kentucky, 130 S. Ct. 1473, 1488 (2010) (Alito, J., concurring); *see also* Chin & Holmes, *supra* note 16, at 705–06 (listing potential "collateral" sanctions); Roberts, *Guilty-Plea Process*, *supra* note 16, at 124 (noting other severe collateral consequences "including deportation, sex-offender registration, post-sentence involuntary commitment as a 'sexually violent predator,' the loss of voting rights, and the loss of housing and employment").

42. Pinard, *Collateral Consequences*, *supra* note 15, at 635. As another scholar has explained: "Collateral sanctions—often referred to as collateral consequences or civil disabilities—are not part of the sentence calculus, even though they derive from a criminal conviction. Some of them follow automatically upon a conviction; others must be imposed by an administrative agency or another regulatory body." Nora V. Demleitner, *Smart Public Policy: Replacing Imprisonment with Targeted Nonprison Sentences and Collateral Sanctions*, 58 STAN. L. REV. 339, 356 (2005).

43. *See, e.g.*, Lefkowitz v. Fair, 816 F.2d 17, 20 (1st Cir. 1987) (noting that the state agency revoked the physician's medical license after his criminal conviction).

44. *See, e.g.*, 18 U.S.C. § 922(g) (2006) (barring certain felons from possessing firearms as a matter of federal law); 18 PA. CONS. STAT. ANN. § 6105 (West 2000) (proscribing persons convicted of certain felonies from possessing firearms in Pennsylvania).

45. Pinard, *Collateral Consequences*, *supra* note 15, at 635; Pinard, *Broadening the Holistic Mindset*, *supra* note 35, at 1073, 1078.

for non-citizen defendants certain misdemeanor convictions constitute 'aggravated felonies' under federal law. Consequently, numerous convictions that are misdemeanors under state law can result in deportation."[46] Notably, scholarship concerning collateral consequences has focused primarily on felony convictions and has largely neglected to address the effect of collateral consequences in the misdemeanor context.[47]

Although criminal convictions have historically been accompanied by collateral sanctions,[48] the number and severity of these consequences, including limitations on employment and housing, have substantially increased over the past two decades.[49] For example, Congress has expanded the number of conviction-related justifications for the deportation of non-citizens.[50] Both federal and state sentencing frameworks have increasingly imposed more serious sentences on defendants with prior convictions.[51] In an attempt to accommodate the ever-increasing costs of prison systems, many states have imposed "non-prison sanctions" on low-risk, non-violent offenders.[52] Furthermore, the enforcement of collateral sanctions has become much more stringent.[53] As a result, "the successful inte-

46. Pinard, *Broadening the Holistic Mindset*, *supra* note 35, at 1077–78; *see also, e.g.*, Andrew Moore, *Criminal Deportation, Post-Conviction Relief and the Lost Cause of Uniformity,* 22 GEO. IMMIGR. L.J. 665, 667 (2008) (noting that the current "crimmigration system" depends upon both state and federal criminal proceedings to "form the basis of grounds for deportation in the federal immigration system").

47. Pinard, *Broadening the Holistic Mindset*, *supra* note 35, at 1076–77.

48. *See, e.g.*, Sibron v. New York, 392 U.S. 40, 55 (1968) ("[It is an] obvious fact of life that most criminal convictions do in fact entail adverse collateral legal consequences.").

49. Chin & Holmes, *supra* note 15, at 699–700; Demleitner, *supra* note 42, at 356; Pinard, *Collateral Consequences*, *supra* note 15, at 638 (explaining that collateral consequences "have reached unprecedented breadth in recent decades"); Roberts, *Guilty-Plea Process*, *supra* note 29, at 701–02 (asserting that "collateral consequences have mushroomed").

50. Maureen A. Sweeney, *Fact or Fiction: The Legal Construction of Immigration Removal for Crimes,* 27 YALE J. ON REG. 47, 63–66 (2010). As Justice Stevens noted in *Padilla*, "While once there was only a narrow class of deportable offenses and judges wielded broad discretionary authority to prevent deportation, immigration reforms over time have expanded the class of deportable offenses and limited the authority of judges to alleviate the harsh consequences of deportation." Padilla v. Kentucky, 130 S. Ct. 1473, 1478 (2010).

51. Alan C. Smith, Note, *More Than a Question of Forum: The Use of Unconstitutional Convictions to Enhance Sentences Following* Custis v. United States, 47 STAN. L. REV. 1323, 1324 (1995).

52. Demleitner, *supra* note 42, at 340–41.

53. Chin & Holmes, *supra* note 15, at 699–700; Pinard, *Broadening the Holistic Mindset*, *supra* note 35, at 1075–76; *see* Roberts, *The Mythical Divide, supra* note 29, at 673–74.

gration of individuals convicted of crimes while burdened with civil liabilities is of major social and economic concern."[54]

The collateral consequences of a conviction may adversely affect a convicted person for the rest of his life.[55] Especially where the defendant committed a minor offense, the collateral consequences frequently overshadow and outlast his direct penal sentence:[56]

> [I]t is fairly typical for an individual pleading guilty for the first time to felony possession or sale of hard drugs to walk out of court[ ] receiving a sentence of time served and probation. The collateral consequences are a far more meaningful result of such a conviction. By virtue of the conviction, the offender may become ineligible for federally funded health care benefits, food stamps and Temporary Assistance for Needy Families, and housing assistance. She is ineligible for federal educational aid. Her driver's license will probably be suspended and she will be ineligible to enlist in the military, receive a security clearance, or possess a firearm. If an alien, she will be deported; if a citizen, she will be ineligible to serve on a federal jury and in some states will lose her right to vote. In cases like these, traditional sanctions such as a fine or imprisonment are comparatively insignificant. The real work of the conviction is performed by the collateral consequences.[57]

As one scholar has noted, "At no point in United States history have collateral consequences been as expansive and entrenched as they are today."[58]

B. Habeas Corpus Relief under 28 U.S.C. § 2254

The "Great Writ of Liberty," federal habeas corpus review provides a "procedural device for subjecting executive, judicial, or private restraints on liberty to judicial scrutiny."[59] In the United States, the writ has evolved into what Chief Justice Earl Warren termed a

54. Roberts, *Guilty-Plea Process*, *supra* note 15, at 126–27, 193.

55. *See id.* at 127 (noting permanent effect of some consequences, such as lifelong civil commitment).

56. *Id.* at 147; Pinard, *Broadening the Holistic Mindset*, *supra* note 35, at 1078; *see also* Smyth, *supra* note 40, at 482 (asserting that "hidden sanctions" can wield "more severe" consequences than "the immediate criminal sentence" does).

57. Chin & Holmes, *supra* note 15, at 699–700.

58. Pinard, *Perspective on Reentry*, *supra* note 38, at 1214–15.

59. Peyton v. Rowe, 391 U.S. 54, 58 (1968); *see also* Prigg v. Pennsylvania, 41 U.S. (16 Pet.) 539, 619 (1842); ERIC M. FREEDMAN, HABEAS CORPUS: RETHINKING THE GREAT WRIT OF LIBERTY 1 (2001); Preiser v. Rodriguez, 411 U.S. 475, 484 (1973) (noting that habeas corpus writ allows petitioner to challenge and secure release from unlawful custody).

"symbol and guardian of individual liberty."[60] Habeas corpus constitutes "an effective and speedy instrument by which judicial inquiry may be had into the legality of the detention of a person."[61] In short, the writ is a fundamental means for protecting "individual freedom against arbitrary and lawless state action."[62]

As the Supreme Court has repeatedly explained, the writ "is not now and never has been a static, narrow, formalistic remedy; its scope has grown to achieve its grand purpose—the protection of individuals against erosion of their right to be free from wrongful restraints upon their liberty."[63] Interpretations of the scope of the Great Writ "must retain the 'ability to cut through barriers of form and procedural mazes.'"[64] At the same time, however, the "sole function" of the writ traditionally has been to grant relief from unlawful custody—and "it cannot be used properly for any other purpose."[65]

Congress has codified federal habeas corpus review in 28 U.S.C. §§ 2241–2255.[66] Section 2241 statutorily recognizes the general authority of the federal courts to grant the habeas corpus writ.[67] Section 2241 provides, for example, the means for a non-

60. *Peyton*, 391 U.S. at 58; Jones v. Cunningham, 371 U.S 236, 243 (1963). The writ is probably the most "extolled, and storied, aspect of the Anglo-American legal tradition . . . . Tracing its lineage back to the Magna Carta, the Great Writ was so revered by the Framers of the U.S. Constitution that they expressly prohibited its suspension except in times of extreme governmental distress." Logan, *supra* note 24, at 147. For a historical overview of the habeas corpus writ, see generally PAUL D. HALLIDAY, HABEAS CORPUS: FROM ENGLAND TO EMPIRE (2010).

61. Carafas v. LaVallee, 391 U.S. 234, 238 (1968) ("[The writ is] shaped to guarantee the most fundamental of all rights . . . .").

62. Harris v. Nelson, 394 U.S. 286, 290–91 (1969).

63. *Peyton,* 391 U.S. at 66; *Jones,* 371 U.S. at 243; Hensley v. Mun. Court, San Jose Milpitas Judicial Dist., 411 U.S. 345, 349–50 (1973).

64. *Hensley,* 411 U.S. at 350 (quoting *Harris*, 394 U.S. at 291).

65. Pierre v. United States, 525 F.2d 933, 935–36 (5th Cir. 1976).

66. 28 U.S.C. §§ 2241–2255 (2006).

67. Section 2241 provides that as a general matter, the writ "shall not extend to a prisoner" unless he is in "custody." 28 U.S.C. § 2241(c) (2006). Section 2241(d) additionally provides:

> Where an application for a writ of habeas corpus is made by a person in custody under the judgment and sentence of a State court of a State which contains two or more Federal judicial districts, the application may be filed in the district court for the district wherein such person is in custody or in the district court for the district within which the State court was held which convicted and sentenced him and each of such district courts shall have concurrent jurisdiction to entertain the application. The district court for the district wherein such an application is filed in the exercise of its discretion and in furtherance of justice may transfer the application to the other district court for hearing and determination.

citizen facing removal to challenge his final order of deportation; this final order of deportation sufficiently establishes "custody" for purposes of § 2241.[68]

By contrast, § 2254 constitutes the principal means through which state criminal offenders can challenge the constitutionality of their convictions in a federal forum.[69] Once a judgment of conviction has been entered in state court, it is "subject to review in multiple forums."[70] Each state has a framework that permits a defendant to challenge his conviction through appellate and post-conviction review by the state's courts.[71] After unsuccessfully litigating his federal constitutional claims through the state appellate and post-conviction process, a state criminal offender may petition for habeas corpus relief in a federal district court.[72] Should the federal district court deny relief, the § 2254 petitioner can appeal to the federal courts of appeals and the Supreme Court.[73]

As a result, through § 2254, Congress granted federal courts the authority to hear a habeas claim alleging that a state conviction violated the U.S. Constitution. Section 2254 provides that:

> The Supreme Court, a Justice thereof, a circuit judge, or a district court shall entertain an application for a writ of habeas corpus in behalf of a person in custody pursuant to the judgment of a State court only on the ground that he is in custody

§ 2241(d).

68. Zadvydas v. Davis, 533 U.S. 678, 687 (2001); Rosales v. Bureau of Immigration & Customs Enforcement, 426 F.3d 733, 735 (5th Cir. 2005). Similarly, a military personnel who is seeking discharge from service may petition for a writ of habeas corpus pursuant to § 2241. *See, e.g.,* Kanai v. McHugh, 638 F.3d 251 (4th Cir. 2011) (holding that a conscientious objector seeking discharge from military service satisfies § 2241 custody requirement). Additionally, some circuits allow inmates who allege defects in the "execution" of their sentences—rather than a constitutional error in the underlying conviction or sentence—to seek relief pursuant to § 2241. *See, e.g.*, Montez v. McKinna, 208 F.3d 862, 865 (10th Cir. 2000) (holding that claim was properly brought under § 2241).

69. CARY FEDERMAN, THE BODY AND THE STATE: HABEAS CORPUS AND AMERICAN JURISPRUDENCE 1 (2006); Larry W. Yackle, *Explaining Habeas Corpus*, 60 N.Y.U. L. REV. 991, 997 (1985) (noting that § 2254 review provides a crucial means for state defendants to challenge the federal constitutionality of their convictions). Section 2255, by contrast, allows those convicted of *federal offenses* to challenge their federal convictions and/or sentences. 28 U.S.C. § 2255 (2006).

70. Lackawanna Cnty. Dist. Attorney v. Coss, 532 U.S. 394, 402 (2001).

71. *See generally* 1 DONALD E. WILKES, JR., STATE POSTCONVICTION REMEDIES AND RELIEF HANDBOOK (2009-2010 ed. 2009) (discussing state post-conviction remedies).

72. *See* 28 U.S.C. § 2254(a), (b)(1) (2006).

73. FEDERMAN, *supra* note 69, at ix.

> in violation of the Constitution or laws or treaties of the United States.[74]

In short, a § 2254 challenge seeks invalidation of the state court judgment authorizing the petitioner's constraints.[75]

Crucially, under § 2254, a one-year "period of limitation" applies to a habeas corpus petition filed by a person in custody pursuant to a state court judgment.[76] The limitation period runs from the latest of:

> (A) the date on which the judgment became final by the conclusion of direct review or the expiration of the time for seeking such review; (B) the date on which the impediment to filing an application created by State action in violation of the Constitution or laws of the United States is removed, if the applicant was prevented from filing by such State action; (C) the date on which the constitutional right asserted was initially recognized by the Supreme Court, if the right has been newly recognized by the Supreme Court and made retroactively applicable to cases on collateral review; or (D) the date on which the factual predicate the claims or claims presented could have been discovered through the exercise of due diligence.[77]

As an additional matter, the "time during which a properly filed application for State post-conviction or other collateral review with respect to the pertinent judgment or claim is pending shall not be counted toward any period of limitation under this subsection."[78]

Furthermore, before a federal court may review the merits of a § 2254 claim, the petitioner must "exhaust" available remedies in state court.[79] Specifically, a § 2254 petitioner "must give the state courts an opportunity to act on his claims before he presents those claims to a federal court in a habeas petition."[80] Thus a state petitioner may not obtain § 2254 relief unless he has properly litigated his "claims through 'one complete round of the State's established

74. § 2254(a).

75. Wilkinson v. Dotson, 544 U.S. 74, 83 (2005).

76. 28 U.S.C. § 2244(d)(1) (2006); Rhines v. Weber, 544 U.S. 269, 274 (2005) (noting this "one year" statute of limitations requirement).

77. § 2244(d)(1)(A)–(D).

78. § 2244(d)(2).

79. 28 U.S.C. § 2254(b)(1) (2006); O'Sullivan v. Boerckel, 526 U.S. 838, 842 (1999).

80. *Boerckel*, 526 U.S. at 842.

appellate review process.'"[81] As Justice Sandra Day O'Connor explained in *O'Sullivan v. Boerckel*, exhaustion is premised on comity; its primary goal is to reduce the "friction between the state and federal court systems by avoiding the 'unseem[liness]' of a federal district court's overturning a state court conviction without the state courts having had an opportunity to correct the constitutional violation in the first instance."[82] Procedural default constitutes the "sanction" for a § 2254 petitioner's failure to exhaust state remedies properly, meaning that the petitioner loses his right to present his habeas claim in federal court.[83] In other words, if state court remedies are no longer available because the petitioner failed to comply with procedural requirements for state-court review, "those remedies are technically exhausted, but exhausted in this sense does not automatically entitle the habeas petitioner to litigate his or her claims in federal court"; instead, if the petitioner "procedurally defaulted," the petitioner *may not* obtain federal review of the merits of his claims.[84]

C. The § 2254 Custody Requirement

The primary justification for § 2254 review is that the individual's interest in freedom from unlawful detention "warrants a second look at federal claims already rejected by the state courts."[85] The subject matter jurisdiction of federal courts reviewing § 2254 claims, however, is "limited explicitly to petitions from applicants who allege they are in 'custody' in violation of federal law."[86]

A petitioner seeking relief pursuant to § 2254 must demonstrate "that he is in 'custody pursuant to the judgment of a State

81. Woodford v. Ngo, 548 U.S. 81, 92 (2006) (quoting *Boerckel*, 526 U.S. at 845); *see also, e.g.*, Dill v. Holt, 371 F.3d 1301, 1303 (11th Cir. 2004) (holding that where petitioner has failed to pursue state remedies he "has not met § 2254(b)(1)(A)'s exhaustion requirement"); Jones v. Jones, 163 F.3d 285, 296 (5th Cir. 1998) (noting that in order to exhaust state remedies, § 2254 petitioner "must have fairly presented the substance of his claim to the state courts").

82. *Boerckel,* 526 U.S. at 845.

83. *Id.* at 848.

84. *Ngo* 548 U.S. at 93. Relief under § 2254 "may be granted on a procedurally defaulted claim only if the petitioner 'can demonstrate cause for the default and actual prejudice' from the alleged violation of federal law or demonstrate that failure to consider the claim[ ] will result in a fundamental miscarriage of justice." *Jones,* 163 F.3d at 296.

85. Yackle, *supra* note 69, at 998; *see also* FEDERMAN, *supra* note 69, at ix (explaining that § 2254 review provides "an important avenue of appeal for state prisoners with unresolved federal questions").

86. Yackle, *supra* note 69, at 999 (emphasis added).

court.'"[87] As a historical matter, the federal habeas corpus statutes, including the present version of § 2254 contained in the Anti-Terrorism and Effective Death Penalty Act (AEDPA),[88] have all required that those seeking relief through the Great Writ must be "in custody."[89]

The fundamentality of "custody" to the purpose and meaning of federal habeas review cannot be overemphasized. The writ "lies to examine the lawfulness of the *custody*; if unlawful, it must be invalidated, and the petitioner must consequently be released from its restraints."[90]

Not only does § 2254 statutorily require "custody," but a petitioner must also demonstrate that he is in custody in order to satisfy the "case or controversy" requirement of Article III of the U.S. Constitution.[91] The case or controversy requirement demands that, throughout the litigation, the party seeking relief must have suffered, or have "be[en] threatened with, an actual injury . . . [that is] likely to be redressed by a favorable judicial decision."[92] Therefore, a party's challenge to the validity of a conviction, in order to satisfy the case or controversy requirement, must demonstrate a "concrete injury caused by the conviction and redressable by invalidation of the conviction."[93]

Because the custody requirement of § 2254 is jurisdictional, it constitutes a critical threshold question that courts must resolve

87. McCormick v. Kline, 572 F.3d 841, 847 (10th Cir. 2009) (quoting Lackawanna Cnty. Dist. Attorney v. Coss, 532 U.S. 394, 401 (2001)).

88. As the First Circuit noted in *Lefkowitz v. Fair*, the requirement of "custody" originated "on this side of the Atlantic in the very first federal habeas statute[, the Judiciary Act of 1789,] and derives from the historic practice in England." 816 F.2d 17, 19 (1st Cir. 1987).

89. WAYNE R. LAFAVE ET AL., CRIMINAL PROCEDURE § 28.3(a), (b), at 1342, 1344 (5th ed. 2009); *see also, e.g.*, Carafas v. LaVallee, 391 U.S. 234, 238 (1968) (noting that the requirement that applicant be "in custody" when habeas corpus petition is filed is "required not only by the repeated references in the statute, but also by the history of the Great Writ"); Mays v. Dinwiddie, 580 F.3d 1136, 1138–39 (10th Cir. 2009) (noting that § 2254 relief is appropriate only where state prisoner is "in custody in violation of the Constitution or laws of the United States"); Logan, *supra* note 23, at 150–51 ("[D]ating back to its common law origins and throughout its extended statutory history . . . federal habeas has required that a petitioner be in 'custody . . . .').

90. *Developments in the Law, Federal Habeas Corpus*, 83 HARV. L. REV. 1072, 1079–80 (1970) [hereinafter *Developments*] (emphasis added) (noting that courts have "traditionally" held that habeas corpus is available only to obtain petitioner's "discharge from custody").

91. U.S. CONST. art. III, § 2.

92. Spencer v. Kemna, 523 U.S. 1, 7 (1998).

93. *Id.*

before hearing the merits of any habeas challenge.[94] The pivotal point on the time continuum for determining whether a § 2254 petitioner is in custody is the time at which he files the petition.[95]

Notably the text of § 2254(a) uses the phrase "in custody" twice.[96] First, the text requires that the petition be "filed in behalf" of a person "in custody." Second, federal courts can hear the petition only on the ground that the petitioner is "in custody" in violation of the "Constitution or laws or treaties of the United States."[97] Consequently, although case law often refers generally to a single "in custody" requirement, "this statutory requirement has two distinct aspects."[98] The language of § 2254(a) "explicitly requires a nexus between the petitioner's claim and the unlawful nature of the custody."[99] Although it limits federal habeas review to those "in custody," § 2254 does not otherwise "attempt to mark the bounda-

94. Bailey v. Hill, 599 F.3d 976, 978 (9th Cir. 2010); *see also, e.g.*, Lackawanna Cnty. Dist. Attorney v. Coss, 532 U.S. 394, 401 (2001) (instructing that § 2254 petitioner must first demonstrate that he is "in custody" pursuant to state court judgment); Jones v. Cunningham, 371 U.S. 236, 238 (1963) (noting, as a "jurisdictional" matter, that the habeas corpus statute "implements the constitutional command that the writ of habeas corpus be made available" to those "in custody"); McCormick v. Kline, 572 F.3d 841, 848 (10th Cir. 2009) (terming the "in custody" requirement of § 2254 "jurisdictional"); Erlandson v. Northglenn Mun. Court, 528 F.3d 785, 788 (10th Cir. 2008) (terming the "in custody" requirement "jurisdictional"); *Developments, supra* note 91, at 1072 (noting that contemporary federal habeas statutes retain custody as "jurisdictional requirement").

95. *Spencer*, 523 U.S. at 7 (noting that petitioner must satisfy custody requirement at the time that habeas petition is filed); Carafas v. LaVallee, 391 U.S. 234, 239–40 (1968) (explaining that the filing of a petition is the relevant point the on time continuum to determine custody); *Bailey*, 599 F.3d at 978–79 (quoting Resendiz v. Kovensky, 416 F.3d 952, 956 (9th Cir. 2005) (internal quotation marks omitted) ("Section 2254(a)'s 'in custody' requirement has been interpreted to mean that federal courts lack jurisdiction over habeas corpus petitions unless the petitioner is under the conviction or sentence under attack at the time his petition his filed.")); Tinder v. Paula, 725 F.2d 801, 803 (1st Cir. 1984) ("Custody is determined from the date that the habeas petitioner is first filed."); Pringle v. Court of Common Pleas, 744 F.2d 297, 300 (3d Cir. 1984) (holding that determination of whether petitioner is in § 2254 custody "is made at the time" petition is filed).

96. *Bailey*, 599 F.3d at 978.

97. *Id.*

98. *Id.* Other observers have similarly explained that to obtain federal habeas corpus review pursuant to § 2254, the petitioner "must satisfy two jurisdictional requirements—(1) the status requirement that the petition be 'in behalf of a person in custody pursuant to the judgment of a State court' and (2) the substance requirement that the petition challenge the legality of that custody on the ground that it is in violation of the Constitution or laws or treaties of the United States.'" 1 RANDY HERTZ & JAMES S. LIEBMAN, FEDERAL HABEAS CORPUS PRACTICE AND PROCEDURE 425 (6th ed. 2011).

99. *Bailey*, 599 F.3d at 980.

ries of 'custody'" or "the situations in which the writ can be used."[100]

1. Extending the Concept of Custody

Throughout most of America's history, federal courts restrictively interpreted the custody requirement to mean actual physical restraint.[101] Yet, beginning in the 1960's, the Supreme Court broadened the definition of "custody" under § 2254.[102] Consistent with this jurisprudential "metamorphosis," § 2254 custody for the past half-century has not required physical restraints or confinement.[103]

In its pivotal 1963 *Jones v. Cunningham* decision, the Supreme Court held that a § 2254 petitioner who had been paroled was in custody for purposes of federal habeas corpus review.[104] *Jones* concluded that a petitioner remained in custody even though he had changed status from a prisoner to a parolee. While a prisoner, the petitioner had been under the custody of the Virginia prison superintendent, but once he became a parolee, the Virginia Parole Board assumed custody over him. As the *Jones* Court critically observed, the parole order (1) placed the petitioner in the control of the Virginia Parole Board; (2) confined him to a particular community, house, and job; (3) required that he periodically report to a parole officer and follow the officer's advice; and (4) required that he allow the officer to visit him at any time.[105] Because of these

100. Jones v. Cunningham, 371 U.S. 236, 238 (1963); *see also* Smith, *supra* note 51, at 1338 ("Despite the jurisdictional importance of the custody requirement, Congress has never defined its meaning.").

101. *See* Barry v. Bergen Cnty. Prob. Dep't, 128 F.3d 152, 160 (3d Cir. 1997) (noting that an early Supreme Court decision held that incarceration was necessary to establish "custody"); Westberry v. Keith, 434 F.2d 623, 624 (5th Cir. 1970); *see also, e.g.*, HERTZ & LIEBMAN, *supra* note 98, at 426 (noting that as a historical matter, federal courts had limited federal habeas corpus review to prisoners "who, at the moment the federal petition was actually adjudicated, were seeking release from actual physical confinement by challenging the particular judgment of conviction or sentence responsible for the confinement").

102. Logan, *supra* note 23, at 148–49, 151. *See generally* FEDERMAN, *supra* note 69, at 95–124 (discussing the expansion of habeas corpus during the 1960's as well as the limits that the Supreme Court subsequently imposed beginning in the 1970's).

103. *Westberry*, 434 F.2d at 624; *Barry,* 128 F.3d at 160; Lefkowitz v. Fair, 816 F.2d 17, 19 (1st Cir. 1987); *see also* Brian M. Hoffstadt, *The Deconstruction and Reconstruction of Habeas*, 78 S. CAL. L. REV. 1125, 1183–84 (2005) ("[T]he Court has held that habeas petitioners satisfy the 'custody' requirement as long as they are 'in custody' at the time they initially file their habeas petitions . . . .").

104. *Jones,* 371 U.S. at 243; *see also* HERTZ & LIEBMAN, *supra* note 98, at 426 n.6 (terming *Jones* a "watershed" decision).

105. *Jones*, 371 U.S. at 240–41.

conditions, the Court held that the *Jones* petitioner was in the custody of the Virginia Parole Board. Concluding that "[a]ctual physical custody" is sufficient but not necessary to satisfy the custody requirement, *Jones* reasoned:

> It is not relevant that conditions and restrictions such as these may be desirable and important parts of the rehabilitative process; what matters is that they *significantly restrain* petitioner's liberty to do those things which in this country free men are entitled to do. Such restraints are enough to invoke the help of the Great Writ. Of course, that writ always could and still can reach behind prison walls and iron bars. But it can do more. . . . While petitioner's parole releases him from immediate physical imprisonment, it imposes conditions which significantly confine and restrain his freedom; this is enough to keep him in the "custody" of . . . the Virginia Parole Board within the meaning of the habeas corpus statute . . . .[106]

Importantly, *Jones* asserted, "History, usage, and precedent can leave no doubt that, besides physical imprisonment, there are other restraints on a man's liberty, restraints not shared by the public generally, which have been thought sufficient in the English-speaking world to support the issuance of habeas corpus."[107] Abandoning a constrained definition of custody, *Jones* recognized that "legal as well as physical restraints could constitute 'custody.'"[108]

Subsequent interpretations of the "in custody" language have consequently not required that a prisoner be physically confined in order to obtain § 2254 review.[109] To illustrate, the Supreme Court in *Peyton v. Rowe* relied on the *Jones* holding to expand the concept of custody.[110] Decided five years after *Jones, Peyton* involved a state prisoner who was sentenced to two consecutive sentences and who raised a habeas challenge as to his second conviction and sentence while he was still serving his first sentence.[111] The Court held that a prisoner serving consecutive sentences was "in custody" under any one of them for purposes of § 2254 review.[112] In reaching this con-

106. *Id.* at 242–43 (emphasis added) (footnote omitted).

107. *Id.* at 240.

108. *Developments*, *supra* note 91, at 1074.

109. *See, e.g.*, Maleng v. Cook, 490 U.S. 488, 491 (1989); Bailey v. Hill, 599 F.3d 976, 979 (9th Cir. 2010).

110. *Peyton v. Rowe*, 391 U.S. 54, 66 (1968).

111. *Id.* at 55.

112. *Id.* at 55, 64–65 (overturning *McNally v. Hill*, 293 U.S. 131, 138 (1934)); *see also, e.g.*, Yale L. Rosenberg, *The Federal Habeas Corpus Custody Decisions: Liberal Oasis or Conservative Prop?*, 23 AM. J. CRIM. L. 99, 101 (1995) (asserting that the Supreme Court read in custody in *Peyton*).

clusion, *Peyton* rejected the assertion that "immediate physical release was the only remedy under the federal writ of habeas corpus."[113]

Consistent with the *Peyton* rule, a habeas petitioner may challenge his state sentences even though he is not presently serving them due to his incarceration in federal prison on federal charges.[114] Likewise, a state prisoner who is serving consecutive state sentences is "in custody" and may challenge the sentence scheduled to run first, even after it has expired, until all sentences have been served—at least as long as they "continue to postpone the date for which he would be eligible for [release]" under the expired sentence.[115] Thus, when determining whether a petitioner has satisfied the § 2254 custody requirement, courts must view "consecutive sentences in the aggregate, not as discrete segments."[116]

In addition, a § 2254 petitioner may satisfy the custody requirement even if he has served no part of his sentence. In *Hensley v. Municipal Court, San Jose Milpitas Judicial District*, the Supreme Court held that a petitioner who was released on his own recognizance pending execution of his state sentence but who (1) was subject by court order to appear at specified times and places, (2) remained at large only by "grace of a stay," and (3) had exhausted all available state court remedies was "in custody" for purposes of § 2254 review.[117]

---

113. *Peyton*, 391 U.S. at 67.

114. *Maleng*, 490 U.S. at 493–94. The federal circuit courts are divided on the question of whether a conviction is subject to habeas corpus challenge by a prisoner incarcerated in another jurisdiction on a different conviction when the demanding state has not formally filed a detainer. *See, e.g.*, Vargas v. Swan, 854 F.2d 1028, 1031–32 (7th Cir. 1988); Frazier v. Wilkinson, 842 F.2d 42, 45 (2d Cir. 1988); Stacey v. Warden, 854 F.2d 401, 403 (11th Cir. 1988).

115. Garlotte v. Fordice, 515 U.S. 39, 43, 47 (1995). In *Garlotte*, invalidation of the petitioner's marijuana conviction, for which the sentence had expired, would have advanced "the date of his eligibility for release from present incarceration." *Id.* Consistent with this rule, a prisoner who was still serving the longer of two concurrent sentences, but who had completed the shorter sentence, was not "in custody" for the purpose of raising a § 2254 challenge to the conviction underlying the shorter sentence. Mays v. Dinwiddie, 580 F.3d 1136, 1137–38 (10th Cir. 2009); *see also* Rosenberg, *supra* note 113, at 101 (terming situation in *Garlotte* the "mirror image" of that in *Peyton*).

116. *Garlotte,* 515 U.S. at 47.

117. 411 U.S. 345, 351–53 (1973); *see also* Lawrence v. 48th Dist. Court, 560 F.3d 475, 480 (6th Cir. 2009) (concluding that a habeas petitioner on recognizance bond at the time of filing a habeas petition met the custody requirement); McVeigh v. Smith, 872 F.2d 725, 727 (6th Cir. 1989) (holding that a stayed one-year probation sentence on recognizance bond satisfied "custody" requirement).

As a practical matter, the Court observed that the *Hensley* petitioner was "subject to restraints 'not shared by the public generally . . . .' He cannot come and go as he pleases. His freedom of movement rests in the hands of state judicial officers who may demand his presence at any time and without a moment's notice. Disobedience is itself a criminal offense."[118] Notably, although the *Hensley* petitioner had been convicted only of a misdemeanor and received a sentence of one year of imprisonment as well as a $625 fine,[119] the Court stated:

> The custody requirement of the habeas corpus statute is designed to preserve the writ of habeas corpus as a remedy for severe restraints on individual liberty. Since habeas corpus is an extraordinary remedy whose operation is to a large extent uninhibited by traditional rules of finality and federalism, its use has been limited to cases of special urgency, leaving more conventional remedies for cases in which the restraints on liberty are neither severe nor immediate. Applying that principle, we can only conclude that petitioner is in custody for purposes of the habeas corpus statute.[120]

As a final consideration, the Court explained that its conclusion that the *Hensley* petitioner was "in custody" did not interfere "with any significant interest of the State."[121]

Importantly, the Supreme Court has held that a petitioner may satisfy the § 2254 custody requirement even where the state court

118. *Hensley*, 411 U.S. at 351 (quoting Jones v. Cunningham, 371 U.S. 236, 240 (1963)). The Supreme Court cautioned, however, that its decision did not "open the door of the [federal] district courts to the habeas corpus petitions of all persons released on bail or on their own recognizance." *Id.* at 353. Instead, where a state petitioner is "released on bail or on his own recognizance pending trial or pending appeal, he must still contend with the requirements of the exhaustion doctrine if he seeks habeas corpus relief in the federal courts. Nothing in today's opinion alters the application of that doctrine to such a defendant." *Id.*

119. *Id.* at 346.

120. *Id.* at 351. Relying upon the reasoning in *Hensley*, the Supreme Court held in *Justices of Boston Municipal Court v. Lydon* that a petitioner was in § 2254 "custody" where he had been released on personal recognizance between (1) a vacated bench trial conviction and (2) the beginning of a jury "trial de novo." 466 U.S. 294, 300–301 (1984). The conditions of release for the *Lydon* petitioner subjected him to "restraints not shared by the public generally." *Id.* at 301 (quoting *Hensley*, 411 U.S. at 351) (internal quotation marks omitted). The Court concluded, "Although the restraints on [the *Lydon* petitioner's] freedom are not identical to those imposed on Hensley, we do not think that they are sufficiently different to require a different result." *Id.*

121. *Hensley,* 411 U.S. at 352; *see also infra* Part II.B (discussing the governmental and societal interest in the finality of state convictions).

judgment from which he seeks relief is not a criminal conviction. In *Duncan v. Walker*, the Court observed that incarceration pursuant to a state criminal conviction constitutes the "most common and most familiar basis for satisfaction of the 'in custody' requirement in § 2254 cases."[122] Yet other types of state court judgments, such as a state court order of civil commitment or civil contempt, also satisfy the § 2254 custody requirement.[123]

Likewise, some courts have held that community service satisfies the § 2254 custody requirement.[124] For example, in *Barry v. Bergen County Probation Department*, the Third Circuit held that a § 2254 petitioner who was serving a sentence of 500 hours of community service satisfied the custody requirement.[125] *Barry* emphasized that the petitioner was subject to a "severe restraint on his liberty" that was not "shared by the public generally" and to some "level of supervision" by the government.[126]

2. "Collateral" Sanctions May Prevent Mootness

The concept of mootness derives from Article III of the U.S. Constitution "under which the exercise of judicial power depends upon the existence of a case or controversy."[127] As the Seventh Circuit has explained, "jurisdiction is a matter of satisfying the statutory 'in custody' requirement, whereas mootness is a question of whether there is any relief the court can grant once it has determined that it indeed has jurisdiction."[128] Thus, federal courts do not possess the "power to decide questions that cannot affect the rights of litigants in the case before them."[129]

The 1968 Supreme Court decision in *Carafas v. LaVallee* held that for purposes of § 2254 petitions, once federal jurisdiction has attached, the expiration of a state prisoner's sentence and his un-

122. 533 U.S. 167, 176 (2001).

123. *Id.* ("[T]hese state court judgments neither constitute nor require criminal convictions.").

124. Barry v. Bergen Cnty. Prob. Dep't, 128 F.3d 152, 159–62 (3d Cir. 1997); Lawrence v. 48th Dist. Court, 560 F.3d 475, 481 (5th Cir. 2009) (holding that § 2254 petitioner's probation and 500 hours of community service imposed "significant restraints" on his liberty).

125. *Barry,* 128 F.3d at 161.

126. *Id.* at 159–61; *see also* Tinder v. Paula, 725 F.2d 801, 803 (1st Cir. 1984) (noting this "two-element" requirement).

127. North Carolina v. Rice, 404 U.S. 244, 246 (1971) (quoting Liner v. Jafco, Inc., 375 U.S. 301, 306 n.3 (1964)); U.S. CONST. art. III, § 2; *see also* St. Pierre v. United States, 319 U.S. 41, 42 (1943) (observing that federal courts may not decide moot issues).

128. Harrison v. Indiana, 597 F.2d 115, 118 (7th Cir. 1979).

129. *Rice*, 404 U.S. at 246.

conditional release from state prison prior to the final adjudication of federal habeas proceedings do not terminate federal jurisdiction.[130] The *Carafas* petitioner had been convicted of burglary and grand larceny in New York state court; because of his conviction, he could not serve as a labor union official, vote in state elections, or serve as a juror.[131] The Supreme Court concluded, "On account of these 'collateral consequences,' the *Carafas* petitioner's habeas claim [was] not moot."[132] The *Carafas* Court reasoned that, because of the "disabilities or burdens" that may have resulted from the petitioner's conviction, he possessed "a substantial stake in the judgment of conviction which survives the satisfaction of the sentence imposed on him."[133] *Carafas* elaborated that a habeas petitioner "should not be . . . required to bear the consequences of [an] assertedly unlawful conviction simply because the path has been so long that he has served his sentence."[134]

Consequently,"[a]s long as the habeas corpus petition was filed in federal court at a time when the petitioner was in custody, an action challenging that custody is not necessarily mooted by the petitioner's release from custody prior to final trial and appellate ad-

130. Carafas v. LaVallee, 391 U.S. 234, 236, 238–39 (1968) (overturning Parker v. Ellis, 362 U.S. 574 (1960)); *see also, e.g.*, Bailey v. Hill, 599 F.3d 976, 979 (9th Cir. 2010) (holding that petitioner must be in custody at the time that the petition is filed, but petitioner's "subsequent release from custody does not itself deprive federal habeas court of its statutory jurisdiction").

131. *Carafas*, 391 U.S. at 235, 237.

132. *Id.* at 237–38 (quoting Ginsberg v. New York, 390 U.S. 629, 633 n.2 (1968)).

133. *Id.* (quoting Fiswick v. United States, 329 U.S. 211, 222 (1946)); *see also, e.g.*, Ginsberg v. New York, 390 U.S. 629, 633 n.2 (1968) (holding that the fact that defendant was given suspended sentence did not render the case moot where certain disabilities, including the possibility of ineligibility for licensing, resulted from conviction); Carty v. Nelson, 426 F.3d 1064, 1071 (9th Cir. 2005) (holding that the release of § 2254 petitioner from state civil commitment did not moot a petition challenging initial commitment where (1) petitioner suffered continuing and concrete injury due to collateral risk of incarceration if he failed to comply with reporting requirements and (2) "meaningful relief" was available from consequences); Nakell v. Attorney Gen. of N.C., 15 F.3d 319, 322–23 (4th Cir. 1994) (holding that a $500 fine and possible disciplinary proceedings as a result of an attorney's state direct criminal contempt conviction were collateral consequences which prevented the attorney's § 2254 petition from becoming moot upon his release from custody after serving his sentence); Zal v. Steppe, 968 F.2d 924, 926 (9th Cir. 1992) (holding that case had not become moot where attorney, cited for contempt of court, faced state disciplinary sanctions); Minor v. Dugger, 864 F.2d 124, 127 (11th Cir. 1989) (holding that § 2254 petition was not rendered moot by defendant's unconditional release, given the possibility that the conviction might later be used for sentence enhancement purposes).

134. 391 U.S. at 240.

judication of the petition."[135] Once the convict's sentence has expired, courts may "presume" that the criminal conviction has continuing collateral consequences that prevent mootness.[136]

In short, a § 2254 petitioner's "subsequent and unfettered release from custody" fails to render his claim moot under Article III.[137] Endowing a degree of "elasticity" to the custody requirement, the Supreme Court has consequently allowed § 2254 petitioners "to attack convictions for which they have yet to be in custody, and for which they have already completed their custody. For persons in any of these situations, habeas is not simply about 'release . . . from unlawful confinement.'"[138] Furthermore, "*Carafas* not only broadened beyond release from custody the relief available by means of habeas, [but] it also recognized that habeas is appropriate to remedy the collateral consequences of a conviction."[139]

3. Collateral Sanctions Are Insufficient to Establish Custody

The Supreme Court and lower federal courts, however, have rejected the argument "that a habeas petitioner may be 'in custody' under a conviction whose sentence has fully expired at the time his petition is filed."[140] The Court has asserted that "this interpretation

135. HERTZ & LIEBMAN, *supra* note 98, at 428–29 (internal citations omitted).

136. Spencer v. Kenna, 523 U.S. 1, 8–12 (1998) (noting this presumption); HERTZ & LIEBMAN, *supra* note 98, at 430 n.11. Notably, however, in *Spencer v. Kemna*, the Supreme Court held that the presumption that a "wrongful criminal conviction has continuing collateral consequences" does not extend to a petitioner's challenge to his parole revocation. *Spencer*, 523 U.S. at 8, 14. The *Spencer* petitioner had been imprisoned as a result of a parole violation; after he was released from prison, he attempted to challenge the legality of the parole revocation—but not his conviction. *Id.* at 5–7. He alleged that he would suffer "collateral consequences" from the alleged improper revocation of his parole—for example, the parole revocation could be used against him in a later proceeding. *Id.* at 14. The Court concluded that such "collateral consequences" were not "concrete injuries" sufficient to avoid "mootness." *Id.*

137. Hoffstadt, *supra* note 103, at 1184.

138. *Id.* (quoting Allen v. McCurry, 449 U.S. 90, 98 n.12 (1980)).

139. Harrison v. Indiana, 597 F.2d 115, 118 (7th Cir. 1979).

140. Maleng v. Cook, 490 U.S. 488, 491 (1989); *see also* Resendiz v. Kovensky, 416 F.3d 952, 956 (9th Cir. 2005) (noting that once a sentence imposed for conviction has "completely expired," the collateral consequences of the conviction do not render an individual "in custody" for purposes of § 2254 review); Tinder v. Paula, 725 F.2d 801, 803 (1st Cir. 1984) ("[A] sentence that has been fully served does not satisfy the custody requirement of the habeas statute, despite the collateral consequences that generally attend a criminal conviction."); Furey v. Hyland, 395 F. Supp. 1356, 1360 (D.N.J. 1975) ("The discussion of the [collateral] disabilities in *Carafas* was necessary only to decide the mootness issue." (citation omitted)).

stretches the language 'in custody' too far."[141] The Court acknowledged that collateral consequences prevent a § 2254 petition that a petitioner filed while in custody from becoming moot upon his subsequent release from custody; yet, these very same consequences "do *not* by themselves satisfy the requirement that the petitioner be 'in custody' when the petition is first filed—even if those consequences have actually materialized."[142]

The 1989 Supreme Court decision in *Maleng v. Cook* held that a habeas petitioner does not remain in custody under a conviction after the sentence imposed for it has fully expired "merely because of the possibility that the prior conviction will be used to enhance the sentences imposed for any subsequent crimes of which he is convicted."[143] The Court distinguished *Carafas*, reasoning that the *Carafas* holding "rested . . . *not* on the collateral consequences of the conviction, but on the fact that the petitioner had been in physical custody under the challenged conviction at the time the petition was filed."[144] Consequently, *Maleng* asserted, "The negative implication of this holding is, of course, that once the sentence imposed for a conviction has completely expired, the collateral consequences of that conviction are not themselves sufficient to render an individual 'in custody' for the purposes of a habeas attack upon it."[145]

141. *Maleng*, 490 U.S. at 491.

142. HERTZ & LIEBMAN, *supra* note 98, at 433–34 n.12; *see also* Williamson v. Gregoire, 151 F.3d 1180, 1183 (9th Cir. 1998) (noting that once a sentence has completely expired "collateral consequences" of that conviction are insufficient to "render an individual in custody" (quoting *Maleng*, 490 U.S. at 492)); *Harrison*, 597 F.2d at 118 ("[The] collateral consequences, although insufficient to establish custody and thus jurisdiction, are enough to keep a petition from becoming moot by the petitioner's release from custody."). As one observer has concisely explained:

> The arguments for an expansive interpretation of the "custody" requirement stemmed from *Carafas'* holding that the "collateral consequences" of a conviction prevented mootness upon release. By the same logic, the argument went, a whole host of "collateral consequences" flowing from convictions warranted the availability of habeas, even post-release.

Ordower, *supra* note 28, at 1843–44.

143. *Maleng*, 490 U.S. at 492; McCormick v. Kline, 572 F.3d 841, 848 (10th Cir. 2009) (repeating the rule that a habeas petitioner does not remain in custody under a conviction after the sentence imposed for it has fully expired merely because the prior conviction could be used to enhance sentences imposed for any later convictions).

144. *Maleng*, 490 U.S. at 492.

145. *Id.*

Thus *Maleng* squarely forecloses § 2254 relief where the convicted party files his petition after his state conviction has expired.[146] As one observer cogently noted:

> [I]n deciding [*Carafas*] as it did, the Court gave several indications that it was consciously eschewing its opportunity to extend the custody concept to include petitioner's disabilities. First, by refusing to consider the collateral restraints themselves a custody, *Carafas* results in an apparent, arbitrary distinction. The negative pregnant of the not-moot holding is that habeas relief is not available to those who apply for the writ after their release, even though they may be subject to the same disabilities which the Court saw as sufficient to warrant a remedy in *Carafas*.[147]

In sum, beginning with *Jones*, the Supreme Court has clearly expanded the concept of custody "beyond immediate physical restraint."[148] *Jones* relied on several different factors in establishing "the necessary degree of restraint, [but it left unclear] what, short of actual physical confinement, remains essential to a finding of custody."[149] Nevertheless, one clear limitation on the § 2254 custody definition is the line between a "restraint on liberty" and a collateral consequence of a conviction.[150] As one scholar has aptly noted:

> [Federal courts have been reluctant to] recognize the lasting effects of criminal convictions as a basis for habeas jurisdiction. This is so despite the obvious inconsistency thereby created: that collateral consequences can preserve habeas jurisdiction, once established, against a claim of mootness, but do not suffice in and of themselves, in the absence of initial jurisdiction.[151]

The Eighth Circuit's decision in *Harvey v. South Dakota* illustrates the unwillingness of federal courts to view collateral consequences as sufficient to establish § 2254 custody.[152] Convicted of grand larceny in 1968, the *Harvey* petitioner was released from

146. Smith, *supra* note 28, at 83.

147. *Developments*, *supra* note 90, at 1077.

148. *Id.* at 1078. A broadened definition of "custody" accompanied the general expansion of the scope of federal habeas corpus review during the 1960s; "[s]ubsequent contractions of habeas review have left this expanded definition intact." LAFAVE, *supra* note 89, at 1342.

149. *Developments*, *supra* note 90, at 1075.

150. Williamson v. Gregoire, 151 F.3d 1180, 1183 (9th Cir. 1998); Bailey v. Hill, 599 F.3d 976, 979 (9th Cir. 2010) (noting this boundary).

151. Logan, *supra* note 23, at 157.

152. 526 F.2d 840, 841 (8th Cir. 1975).

prison in March of 1971.[153] In 1969, the conviction was affirmed on state direct appeal and the *Harvey* petitioner initiated the state post-conviction relief process.[154] Eventually exhausting his state post-conviction remedies, the *Harvey* petitioner finally filed a § 2254 claim in 1974—three years after his release from prison.[155]

In an attempt to satisfy the threshold jurisdictional requirement of custody, the *Harvey* petitioner contended that "the disabilities which [arose from his] conviction" established custody within the meaning of § 2254.[156] He noted that not only was he "unable to pursue certain professions" or possess a firearm, but also that he would "occup[y] the status of a recidivist if he commits another crime."[157]

The Eighth Circuit concluded that the *Harvey* petitioner failed to satisfy the custody requirement of § 2254 and thus was barred from seeking federal habeas review. *Harvey* found *Carafas* to be distinguishable because the *Carafas* petitioner had "filed his habeas corpus petition while he was still incarcerated."[158] The Eighth Circuit elaborated that the collateral consequences in *Carafas* "kept the case from becoming moot; they did not suffice to give the federal courts jurisdiction."[159] Furthermore, the Eighth Circuit observed that finding custody in a case such as *Harvey* would render Congress's "in custody" requirement meaningless and that "[t]he restraints on [the *Harvey* petitioner's] liberty are 'neither severe nor immediate.'"[160]

Lastly, the *Harvey* petitioner had contended that he had not filed his federal habeas petition while he was in prison because of "state delay in processing his attempts to obtain post-conviction relief."[161] In response, the Eighth Circuit crucially instructed, "It would appear that much of the delay in [the *Harvey* petitioner's] case in state court was attributable to his own inaction. But even if this were not so, we are powerless to grant the federal courts subject matter jurisdiction; only Congress or the Constitution may do so."[162]

153. *Id.*
154. *Id.*
155. *Id.*
156. *Id.*
157. *Id.*
158. *Id.*
159. *Id.*
160. *Id.*
161. *Id.*
162. *Id.*

Additionally, federal courts have limited the § 2254 "custody" definition within the context of sentence-enhancing prior convictions. The Supreme Court has held that § 2254 relief is unavailable where a prisoner challenges a current sentence on the ground that it was enhanced based on an allegedly unconstitutional prior conviction for which the petitioner is no longer in custody.[163] In *Lackawanna District Attorney v. Coss*, a state prisoner brought a § 2254 petition in which he sought to attack a current state sentence that had been enhanced by an earlier, expired state conviction.[164] The Court concluded that the petitioner was in custody pursuant to the later court judgment, but not with regard to the expired conviction—even though this enhanced sentence was a collateral consequence of the earlier, expired, and purportedly unconstitutional conviction. Thus once a state conviction is no longer open to direct or collateral attack in its own right, the conviction may be regarded as conclusively valid.

Furthermore, federal courts have held that a non-citizen may not "collaterally attack the legitimacy of a state criminal conviction in a deportation proceeding."[165] In reaching this conclusion courts have applied the reasoning in *Coss*, asserting that "no meaningful difference" exists "between a collateral attack on an expired state conviction underlying removal proceedings and a collateral attack on an expired state criminal conviction underlying an enhanced sentence."[166] Because immigration consequences, "such as deporta-

163. Lackawanna Cnty. Dist. Attorney v. Coss, 532 U.S. 394, 396–97 (2001). *Coss* thus resolved the issue of the "extent to which the prior expired conviction itself may be subject to challenge in the attack on the current sentence which it was used to enhance." *Id.* at 402 (quoting Maleng v. Cook, 490 U.S. 488, 494 (1989)). The Supreme Court suggested an exception to this rule (1) where the petitioner did not receive counsel in violation of the Sixth Amendment for the adjudication of his prior conviction, or (2) in those rare cases in which no channel of review was actually available to a defendant with respect to a prior conviction due to no fault of his own. *Id.* at 404–06.

164. *Id.* at 401–02.

165. Trench v. INS, 783 F.2d 181, 184 (10th Cir. 1986); *see also* Contreras v. Schiltgen, 151 F.3d 906, 908 (9th Cir. 1998) (noting that federal habeas review is "limited" where the habeas petitioner challenges the use of an expired conviction that forms the basis for the immigration custody). As the Fifth Circuit has reasoned, a collateral challenge to a criminal conviction in deportation proceedings "could not, as a practical matter, assure a forum reasonably adapted to ascertaining the truth of the claims raised. It could only improvidently complicate the administrative process. Once the conviction becomes final, it provides a valid basis for deportation unless it is overturned in a judicial post-conviction proceeding." Zinnanti v. INS, 651 F.2d 420, 421 (5th Cir. 1981).

166. *Drakes v. INS*, 330 F.3d 600, 604 (3d Cir. 2003); Neyor v. INS, 155 F. Supp. 2d 127, 133 (D.N.J. 2001).

tion, have long been viewed as 'collateral,'" they are "not themselves sufficient to render an individual 'in custody'" for purposes of § 2254 review.[167] As the Third Circuit reasoned in *Drakes v. INS*,:

> [R]emoval of an alien who violates the immigration laws does not constitute punishment but, rather, is a civil action necessary to enforce the country's immigration laws. Even if removal involves a greater potential injury to a petitioner than an enhanced sentence, such an injury does not outweigh the interests of finality and ease of administration. Allowing [a noncitizen in removal proceedings] to collaterally attack his prior convictions would "sanction an end run around statutes of limitations and other procedural barriers that would preclude the movant from attacking the prior conviction directly."[168]

The reasoning in *Neyor v. INS* is illustrative.[169] The *Neyor* petitioner, convicted of New Jersey drug offenses, was sentenced to fifteen months of imprisonment; his convictions were affirmed through state appellate and post-conviction relief proceedings.[170] After the *Neyor* petitioner completed his sentence, federal immigration authorities began proceedings to deport him to Liberia.[171] Seeking federal habeas relief, the *Neyor* petitioner did not "attack the validity of the deportation proceeding itself" but claimed only that the "underlying conviction [was] invalid."[172]

Neyor observed that individuals in immigration custody might challenge their deportation pursuant to § 2241,[173] whereas § 2254 constitutes the "principal means" for state criminal offenders to raise a federal challenge to the constitutionality of their convictions.[174] *Neyor* instructed that a federal court may not review the validity of an expired state conviction pursuant to § 2241 even if that conviction serves as a predicate for immigration detention.[175] *Neyor* reasoned that removal statutes "do not allow [immigration au-

167. Resendiz v. Kovensky, 416 F.3d 952, 956 (9th Cir. 2005); Ogunwomoju v. United States, 512 F.3d 69, 74–75 (2d Cir. 2008) (holding that one held in immigration detention is not "in custody" for purpose of challenging state conviction under § 2254).

168. *Drakes,* 330 F.3d at 605.

169. 155 F. Supp. 2d at 127.

170. *Id.* at 130.

171. *Id.* at 132.

172. *Id.*

173. *See supra* notes 68–69 and accompanying text for a general overview of § 2241.

174. *Neyor*, 155 F. Supp. 2d at 131, 134; *see also* 28 U.S.C. § 2241(c)(3) (2006) (using similar language in federal cases).

175. *Neyor*, 155 F. Supp. 2d at 137, 139.

thorities] to inquire into the validity of the underlying convictions during removal proceedings but allow only an inquiry as to the fact of conviction."[176] Therefore, *Neyor* concluded that § 2241 is a means to challenge only final administrative determination of immigration authorities and not the "validity of the conviction."[177]

The *Neyor* court explained that the petitioner should have sought § 2254 relief "directly against the state during the time he was incarcerated for his state conviction and before the 1-year statute of limitations [had]. . .expired."[178] By the time that he had filed his habeas petition, however, the *Neyor* petitioner had completed his sentence, thus rendering the state conviction "expired."[179] Therefore, *Neyor* held that the petitioner was not in "custody" under the expired state conviction and could not attack the conviction through a § 2254 petition.[180]

Notably, although often a "direct" consequence of a conviction,[181] a fine or a restitution order does not establish § 2254 custody—particularly where incarceration or the threat of incarceration for non-payment is assent. As some courts have reasoned, a fine or restitution, on its own, does not constitute a "significant restraint on liberty."[182]

In *Bailey v. Hill*, for example, the § 2254 petitioner, who had been convicted in state court of kidnapping and attempted murder, received a sentence of imprisonment and was ordered to pay $6,606.65 in restitution.[183] The *Bailey* petitioner, while serving his

176. *Id.* at 139.

177. *Id.*; *see also* Smith, *supra* note 28, at 83 ("The merits of a conviction are not subject to litigation in deportation itself.").

178. *Neyor*, 155 F. Supp. 2d at 138.

179. *Id.* at 134.

180. *Id.*

181. *See supra* notes 34–35 and accompanying text for a brief discussion concerning "direct consequences" of convictions.

182. Bailey v. Hill, 599 F.3d 976, 979–981 (9th Cir. 2010); Obado v. New Jersey, 328 F.3d 716, 717 (3d Cir. 2003); *see also, e.g.*, Erlandson v. Northglenn Mun. Court, 528 F.3d 785, 787–88 (10th Cir. 2008) (declining to find that petitioner convicted of littering was in custody where his only punishment was a fine); Barnickel v. United States, 113 F.3d 704, 706 (7th Cir. 1997); United States v. Michaud, 901 F.2d 5, 6–7 (1st Cir. 1990) (holding that an outstanding $60,000 fine and the possibility of imprisonment for nonpayment failed to establish custody); Duvallon v. Florida, 691 F.2d 483, 485 (11th Cir. 1982) (holding that a § 2254 petitioner subject only to a fine did not satisfy the "custody" requirement).

183. 599 F.3d at 977. The 2009 Seventh Circuit decision in *Washington v. Smith* is similarly instructive. 564 F.3d 1350 (7th Cir. 2009). A Wisconsin jury convicted the petitioner of forgery, and "[h]e was sentenced to serve two and a half years in prison and three years of [probation]. Additionally, the trial court ordered [him] to pay restitution in the amount of $15,000, as well as other fines and costs." *Id.* at

term of imprisonment, challenged only the restitution aspect of his conviction.[184] In holding that the *Bailey* petitioner failed to satisfy the § 2254 custody requirement, the Ninth Circuit reasoned that the desired remedy, "the elimination or alteration of a money judgment, does not directly impact—and is not directed at the source of the restraint—on his liberty."[185] Thus the "physical custody" component of the *Bailey* petitioner's sentence was insufficient to confer jurisdiction over his § 2254 petition that challenged only the "restitution" aspect of his sentence.[186] As some courts have reasoned, fines do not constitute custody because "such sentences implicate only property, not liberty."[187]

Likewise, a $250 fine and a driver's license suspension of one year following the petitioner's conviction for failure to yield the right of way did not satisfy the § 2254 custody requirement.[188] As the Fifth Circuit has reasoned, "To allow such circumstances to form the basis of a claim that [a § 2254 petitioner] was in custody would go far beyond that degree of confinement found sufficient in *Carafas* and *Jones*."[189]

Furthermore, the loss of a professional license following a criminal conviction fails to establish § 2254 "custody." To illustrate, in *Lefkowitz v. Fair*, the state medical board revoked the medical license of a physician who had been convicted of rape.[190] In holding that this revocation was insufficient to establish § 2254 custody, the First Circuit stated: "Insofar as the [*Lefkowitz* petitioner] urges that 'custody' remains attached to a degree sufficient to warrant the exercise of federal habeas jurisdiction even after the expiration of his state sentence, he is whistling past the graveyard."[191] *Lefkowitz* rea-

1350. After exhausting his state remedies, the *Washington* petitioner sought § 2254 relief on the basis that his attorney provided ineffective assistance of counsel with respect to the restitution amount. *Id.* The Seventh Circuit, noting that the petitioner attacked only the restitution aspect of his sentence, concluded that he did not state a claim for relief under § 2254. *Id.* at 1351. The court reasoned that a § 2254 petitioner must attack the "fact or duration of one's sentence; if it does not, it does not state a proper basis for relief." *Id.*

184. *Bailey*, 599 F.3d at 978.

185. *Id.* at 981.

186. *Id.* at 980–81.

187. *See* Barry v. Bergen Cty. Prob. Dep't, 128 F.3d 152, 161 (3d. Cir. 1997).

188. Westberry v. Keith, 434 F.2d 623, 624 (5th Cir. 1970); *see also* Lillios v. New Hampshire, 788 F.2d 60, 61 (1st Cir. 1986) (holding that fines and a driver's license suspension were not "severe restraints" and thus did not establish "custody").

189. *Westberr*y, 434 F.2d at 625.

190. Lefkowitz v. Fair, 816 F.2d 17, 20 (1st Cir. 1987).

191. *Id.*

soned that habeas jurisprudence "has traditionally been concerned with liberty rather than property, with freedom more than economics. . . . '[S]ome vaguely defined discrimination or some sort of economic duress' resulting from a conviction does not, by itself, confer standing to invoke the remedy of habeas corpus."[192] *Lefkowitz* concluded that "revocation of a medical license . . . does not constitute the type of grave restraint on liberty or . . . ongoing governmental supervision which are unavoidable prerequisites of actionable 'custody.'"[193] The First Circuit elaborated:

> Adverse occupational and employment consequences are a frequent aftermath of virtually any felony conviction. Government regulation, in the nature of the imposition of civil disabilities—say, loss of voting rights or disqualification from obtaining a gun permit—often follows a defendant long after his sentence has been served. To hold that the custody requirement is so elastic as to reach such sequellae would be to stretch the concept of custody out of all meaningful proportion, to render it limp and shapeless—in the last analysis, to make habeas corpus routinely available to all who suffer harm emanating from a state conviction, regardless of actual custodial status. We abjure such an expansive rule.[194]

Lefkowitz added, "There are *no magic mirrors*; even grievous collateral consequences stemming directly from a conviction cannot, without more, transform the absence of custody into the presence of custody for the purpose of habeas review."[195]

Unsurprisingly, reputational restraints on a person's ability to pursue his occupation following a criminal conviction also fail to establish § 2254 custody. To illustrate, in *Edmunds v. Won Bae Chang*, an attorney who had received a state court contempt citation contended that it constituted a "severe restraint" on his "liberty" because it harmed "his reputation as a lawyer" and "adversely affect[ed]" his relationship with the courts.[196] In holding that the *Edmunds* petitioner failed to satisfy the § 2254 custody requirement, the Ninth Circuit emphasized that the $25 contempt fine consti-

192. *Id.*; *see also* Ginsberg v. Abrams, 702 F.2d 48, 49 (2d Cir. 1983) (holding that a habeas petitioner's removal from bench of family court, revocation of his law license, and disqualification from being licensed as real estate broker or insurance agent did not "so greatly limit[ ] his economic mobility as to constitute 'custody'").

193. *Lefkowitz*, 816 F.2d at 20.

194. *Id.*

195. *Id.* (emphasis added).

196. 509 F.2d 39, 41 (9th Cir. 1975).

tuted the only demonstrable harm that the *Edmunds* petitioner had suffered.[197]

Anxiety resulting from the "unconstitutional delay" of a state appeal also fails to establish custody. The Second Circuit instructed in *Diaz v. Henderson* that "[a]nxiety, even when unconstitutionally inflicted, is harm of a personal nature."[198] Finally, offender registration requirements do not create custody.[199] As the Ninth Circuit has reasoned, these obligations are merely a "collateral consequence" of a conviction, thus barring federal courts from exercising § 2254 jurisdiction.[200]

D. The 2010 Supreme Court Decision in Padilla v. Kentucky

The watershed 2010 Supreme Court decision in *Padilla v. Kentucky* marked a paradigm shift in how courts view the collateral sanctions of a conviction. As one observer has noted, *Padilla* "suggests the potential for a more general blurring of the line between 'direct' and 'collateral' consequences of convictions."[201]

1. The *Padilla* Analysis

Prior to *Padilla v. Kentucky*, courts were divided as to whether a criminal defense attorney's failure to inform a non-citizen client about the immigration consequences of his guilty plea constituted ineffective assistance of counsel under the Sixth Amendment of the U.S. Constitution.[202] *Padilla* resolved that question, holding that in order to provide constitutionally adequate representation, criminal defense counsel must inform his client if the guilty plea and resulting conviction might result in deportation.[203]

197. *Id.*

198. 905 F.2d 652, 654 (2d Cir. 1990).

199. Williamson v. Gregoire, 151 F.3d 1180, 1184 (9th Cir. 1998); *see also, e.g.*, Resendiz v. Kovensky, 416 F.3d 952, 958 (9th Cir. 2005) (holding that a "narcotics offender" registration requirement did not establish custody under § 2254); Leslie v. Randle, 296 F.3d 518, 522–23 (6th Cir. 2002) (holding that because registration requirements were merely a "collateral consequence," the § 2254 petitioner was not "in custody").

200. *Williamson*, 151 F.3d at 1183.

201. Rachel E. Rosenbloom, *Will* Padilla *Reach Across the Border?*, 45 NEW ENG. L. REV. 327, 328 (2011); *see also* Smyth, *supra* note 32, at 796 (asserting that *Padilla* "shocked commentators and practitioners alike").

202. *See, e.g.*, Trench v. INS, 783 F.2d 181, 184 (10th Cir. 1986) (noting division among courts); *see also supra* note 15 for a brief definition of a Sixth Amendment "ineffective assistance of counsel claim."

203. Padilla v. Kentucky, 130 S. Ct. 1473, 1478, 1483.

A native of Honduras, Jose Padilla had been a lawful permanent resident of the United States for more than 40 years.[204] He faced deportation after pleading guilty in Kentucky state court to the transportation of a large amount of marijuana.[205] Padilla claimed that his counsel was constitutionally ineffective because he had failed to warn him that his guilty plea would potentially result in deportation.[206]

In affirming the validity of Padilla's conviction, the Kentucky Supreme Court concluded that "collateral consequences," such as deportation, were outside "the scope of representation required by the Sixth Amendment" and therefore that the "failure of defense counsel to advise the defendant of possible deportation consequences is not cognizable as a claim for ineffective assistance of counsel."[207] Having exhausted available state post-conviction remedies, Padilla sought § 2254 review of his conviction.[208]

In concluding that Padilla's § 2254 claim was meritorious, the Supreme Court explained, "[D]eportation is an integral part—indeed, sometimes the most important part—of the penalty that may be imposed on noncitizen defendants who plead guilty to specified crimes."[209] Justice John Paul Stevens, writing the majority opinion for the Court, further observed:

> We have long recognized that deportation is a particularly severe "penalty;" but it is not, in a strict sense, a criminal sanction. Although removal proceedings are civil in nature, deportation is nevertheless intimately related to the criminal process. Our law has enmeshed criminal convictions and the penalty of deportation for nearly a century. And, importantly, recent changes in our immigration law have made removal nearly an automatic result for a broad class of noncitizen offenders. Thus, we find it most difficult to divorce the penalty from the conviction in the deportation context.[210]

204. *Id.*, at 1477; *see also* Rosenbloom, *supra* note 201, at 328 ("[*Padilla* holds] that criminal defense attorneys have a duty to provide accurate advice to noncitizen clients regarding the immigration consequences of a guilty plea.").

205. *Padilla,* 130 S. Ct. at 1477.

206. *Id.* at 1478.

207. *Id.* at 1481 (quoting Commonwealth v. Padilla, 253 S.W.3d 482, 483 (Ky. 2008)).

208. *See id.* at 1478.

209. *Id.* at 1478, 1480.

210. *Id.* at 1481. The Court reserved the determination of whether *Padilla* satisfies the necessary "second" element of an ineffectiveness claim—prejudice—for the Kentucky courts. *Id.* at 1483–84.

Padilla conceded that disagreement existed among courts concerning "how to distinguish between direct and collateral consequences."[211] Yet *Padilla* notably chose to avoid applying this distinction in reaching its decision. Instead Justice Stevens explained that "whether that distinction is appropriate is a question we need not consider in this case because of the unique nature of deportation."[212] Thus the Court, perhaps self-consciously recognizing the significance of its reasoning, declined to explicitly abandon the "direct/collateral" distinction.

At a minimum, however, *Padilla* had clearly eroded the formalistic "direct/collateral" distinction within the context of Sixth Amendment effective assistance of counsel claims. The Court elaborated:

> Deportation as a consequence of a criminal conviction is, because of its close connection to the criminal process, uniquely difficult to classify as either a direct or a collateral consequence. The collateral versus direct distinction is thus ill-suited to evaluating a [Sixth Amendment ineffective assistance of counsel] claim concerning the specific risk of deportation.[213]

The Court also observed that "preserving the client's right to remain in the United States may be more important to the client than any potential jail sentence."[214] *Padilla* elaborated that the "severity of deportation—'the equivalent of banishment or exile'—only underscores how critical it is for counsel to inform her noncitizen client that he faces a risk of deportation."[215] *Padilla* emphasized that "the seriousness of deportation as a consequence of a criminal plea, and the concomitant impact of deportation on families living lawfully in this country demand no less."[216]

Nevertheless, in his dissenting opinion, Justice Antonin Scalia asserted that the *Padilla* holding "prevents legislation that could solve the problems addressed by today's opinions in a more precise and targeted fashion."[217] Terming the *Padilla* decision a "sledge hammer," he added, "If the subject had not been constitutionalized, legislation could specify which categories of misadvice about matters ancillary to the prosecution invalidate plea agreements,

211. *Padilla*, 130 S. Ct. at 1481 n.8.
212. *Id.* at 1481.
213. *Id.* at 1482.
214. *Id.* at 1483.
215. *Id.* at 1486.
216. *Id.*
217. *Padilla*, 130 S. Ct. at 1496 (Scalia, J., dissenting).

what collateral consequences counsel must bring to a defendant's attention, and what warnings must be given."[218]

2. Extending *Padilla* to Other Collateral Sanctions in the Sixth Amendment Context

Padilla seemingly sidestepped the "direct/collateral" distinction within the context of Sixth Amendment effective assistance of counsel claims; it instructed courts to focus on whether the consequence is "intimately related to the criminal process" and is "an integral part of the penalty."[219] In doing so, however, *Padilla* "effectively undermined" the formalistic rule that only direct consequences were legally significant for Sixth Amendment purposes and that "collateral" consequences were of no moment.[220]

Within the context of Sixth Amendment ineffective assistance of counsel claims, some courts have expanded the *Padilla* reasoning to encompass collateral sanctions other than deportation.[221] As one observer has asserted:

> These cases all demonstrate that the analysis conducted in *Padilla* can [be], and certainly is, applied to other consequences, especially those that courts have found to be fundamental or related to constitutional rights. These may include the right to vote, serve on a jury, possess a firearm, create and remain with [one's] family, and serve in the military. Likewise, those consequences which result in "drastic measures," including registration as a sex offender, the imposition of legal financial obligations, losing the ability to work, losing the right [to] drive a vehicle, losing stable housing, and losing the ability to seek an education may also fall within *Padilla*'s definition of "integral part" of the penalty.[222]

218. *Id.* at 1496–97.

219. *Id.* at 1476; *see also* Smyth, *supra* note 32, at 800 (noting that the analysis shifted from "collateral" to "integral").

220. Smyth, *supra* note 32, at 798.

221. *Id.* at 809 ("Even a cursory reading of *Padilla* begs an inquiry into its application to other so-called 'collateral consequences.'"). In *Frost v. State,* for example, the Alabama Court of Criminal Appeals held that *Padilla* applied to the failure of an attorney to advise his client that he would not be eligible for parole. *See* CR-09-1037, 2011 WL 2094777, at *1–2 (Ala. Crim. App. May 27, 2011). *But see* Brown v. Goodwin, Civ. No. 09-211, 2010 WL 1930574, at *13 (D.N.J. May 11, 2010) ("[T]he holding of *Padilla* seems not importable—either entirely or, at the very least, not readily importable—into scenarios involving collateral consequences other than deportation.").

222. Stearns, *supra* note 21, at 863–64.

To illustrate, in *Bauder v. Department of Corrections*, the Eleventh Circuit held that an attorney was constitutionally ineffective when he incorrectly advised his client that the client would not be subject to civil commitment past his sentence if he were to plead guilty to aggravated stalking.[223] Citing *Padilla*, *Bauder* reasoned that even if the relevant law is unclear a criminal defense attorney must still advise his client that the "pending criminal charges may carry a risk of adverse [collateral] consequences."[224]

Likewise, in *Taylor v. State*, the Georgia Court of Appeals held that, consistent with *Padilla*, a defense attorney provides constitutionally deficient representation if he fails to advise his client that a guilty plea might require him to register as a sex offender.[225] In reaching this conclusion, *Taylor* observed that *Padilla* "calls into question the application of the direct versus collateral consequences distinction in the context of ineffective assistance claims."[226]

Similarly, in *Commonwealth v. Abraham*, the Pennsylvania Superior Court held that, pursuant to *Padilla*, counsel was obligated to warn a criminal defendant that he would lose his teacher's pension as a consequence of pleading guilty to indecent assault.[227] In reaching this conclusion, the *Abraham* court reasoned that the pension loss is "automatic and inevitable, the stakes are high and the consequences are succinct, clear, and distinct. Because of the automatic nature of forfeiture, the punitive nature of the consequence, and the fact that only criminal behavior triggers forfeiture, the [pension loss] is, like deportation, intimately connected to the criminal process."[228]

223. 619 F.3d 1272, 1274–75 (11th Cir. 2010).

224. *Id.* at 1275. The Eleventh Circuit emphasized that the attorney did not tell his client that he possibly faced civil commitment, that the relevant law was unclear, or that "he simply did not know." *Id.* Instead the *Bauder* attorney told his client "that pleading to the criminal charge would *not* subject Bauder to civil commitment, and this constituted affirmative misadvice." *Id.*

225. 698 S.E.2d. 384 (Ga. Ct. App. 2010).

226. *Id.* at 387. The Pennsylvania Superior Court has likewise observed, "Under *Padilla,* it is unclear if the direct/collateral analysis is still viable. That analysis might still be useful if the nature of the action is not as 'intimately connected' to the criminal process as deportation." Commonwealth v. Abraham, 996 A.2d 1090, 1092 (Pa. Super. Ct. 2010).

227. *Abraham*, 996 A.2d at 1095.

228. *Id.* at 1094–95. Additionally, in *Wilson v. State*, the Alaska Court of Appeals concluded that, consistent with *Padilla,* an attorney provided ineffective assistance of counsel where he incorrectly advised his client that his "no-contest" plea to second-degree assault would not be used against him in a later trial for civil damages. 244 P.3d 535, 536 (Alaska Ct. App. 2010).

As one scholar has contended, "It makes sense that [the *Padilla*] analysis can be applied to other consequences, too. Those consequences that meet the standard set by *Padilla* and are 'intimately related to the criminal process' should be subject to [Sixth Amendment] analysis . . . ."[229] Thus, the *Padilla* analysis "applies to a broad range of penalties traditionally considered 'collateral' and outside the concern of the criminal justice system."[230]

In short, *Padilla* repudiated the "fiction of the collateral consequence doctrine with a simple truth: so-called 'collateral' consequences are anything but collateral. In reality, they are 'enmeshed,' 'intimately related to the criminal charge such that it is 'difficult to divorce the penalty from the conviction.'"[231] Within the Sixth Amendment context, *Padilla* "ripped the foundations from the perennially unsound 'collateral/direct' consequence distinction."[232] At a minimum, *Padilla* "may well limit the courts' ability to disregard some consequences as 'collateral' if a particular consequence can be considered 'truly clear' and an integral part of the punishment."[233]

3. *Padilla* and the § 2254 Custody Dilemma

Nevertheless, some courts have declined to extend the *Padilla* analysis to collateral consequences in the § 2254 custody context. As one court has asserted, "the *Padilla* Court discussed whether it was appropriate to classify deportation as either a 'direct' or 'collateral' consequence of a conviction, but did so only in the context of considering whether advice about the consequences of deportation fell within the ambit of the Sixth Amendment guarantee."[234]

To illustrate, in *Fenton v. Ryan*, the U.S. District Court for the Eastern District of Pennsylvania rejected the argument that *Padilla* "impact[s]" the custody "analysis" under § 2254.[235] The *Fenton* peti-

229. Stearns, *supra* note 20, at 861–62. Because of *Padilla*'s holding and its avoidance of "the broader issue [of other collateral consequences], state and federal courts are grappling with whether to extend *Padilla* beyond the issue of immigration consequences." *Id.* at 856.

230. Smyth, *supra* note 32, at 798–99.

231. *Id.* at 825.

232. *Id.* at 798.

233. Webb v. State, 334 S.W.3d 126, 139 (Mo. 2011) (Wolff, J., concurring).

234. Donahue v. Souders, No. 10-2761, 2011 WL 1838780, at *3 n.7 (E.D. Pa. Apr. 20, 2011).

235. Fenton v. Ryan, No. 11-2303, 2011 WL 3515376, at *2 (E.D. Pa. Aug. 11, 2011). Albeit within the context of a § 2255 petition, the U.S. District Court for the Eastern District of California rejected a similar argument in *United States v. Krboyan*. Nos. CV-F-10-2016 OWW, CR-F-02-5438 OWW, 2010 WL 5477692 at *4–6 (E.D.

tioner, a non-citizen of the United States, pled guilty to a drug offense and served a sentence of nearly two years in prison.[236] Seeking habeas relief, the *Fenton* petitioner alleged that his counsel had failed to inform him of the collateral consequences his guilty plea might have on his immigration status in violation of *Padilla*.[237] Concluding that the *Fenton* petitioner did not satisfy the § 2254 custody requirement, the federal district court reasoned that "removal proceedings and removal itself—much less the possibility of removal proceedings—do not constitute [§ 2254] custody for habeas purposes."[238]

II.
THE PROBLEM: FEDERAL HABEAS REVIEW WHERE THE § 2254 PETITIONER IS NO LONGER IN CUSTODY AT THE TIME OF FILING THE PETITION, BUT ENCOUNTERS COLLATERAL SANCTIONS

Given the ever-increasing frequency of "collateral" sanctions that result from criminal convictions, the impractical formalism of the current bright-line framework concerning "collateral consequences" and § 2254 custody has engendered considerable difficulty. Federal courts have fashioned an incongruous rule according to which, so long as the petition was filed before the sentence expired, "collateral" sanctions are sufficient to prevent a § 2254 petition from becoming moot, but are insufficient to establish custody where the petition was filed after the sentence has expired. Because collateral consequences are insufficient to establish custody, "short sentence" state offenders, unlike more serious offenders, frequently cannot claim the benefit of federal habeas review even if their convictions produce detrimental collateral consequences. At the same time, however, an overly broad definition of § 2254 custody that automatically establishes custody because of an alleged collateral consequence is inimical to the inherently limited purpose of habeas relief as well as the compelling societal interest in finality.

Cal. Dec. 30, 2010). Citing the axiomatic "*Maleng*" rule, the *Krboyan* court noted that the *Padilla* petitioner's sentence, in contrast to that of the *Krboyan* petitioner's, had not yet expired at the time that he filed his petition. *Id.* at *6.

236. *Fenton*, 2011 WL 3515376, at *1.

237. *Id.*

238. *Id.* at *2; *see also, e.g.*, *Donahue*, 2011 WL 3515376 at *3 (rejecting the contention that, under *Padilla,* a sexual offender designation is sufficient to establish § 2254 custody).

A. The Carafas/Maleng Paradox

For the past two decades, courts have applied the "*Carafas/Maleng*" paradox to determine whether custody exists for purposes of § 2254. Importantly, however, the increasing use of collateral consequences as a punitive device seems to militate against the future practicality of this paradox.

The 1968 Supreme Court decision in *Carafas v. LaVallee* held that once federal jurisdiction has attached, the expiration of a state prisoner's sentence and his unconditional release from state prison prior to the final adjudication of federal habeas proceedings failed to render the § 2254 petition moot.[239] *Carafas* reasoned that because of the "disabilities or burdens" that may result from a conviction, a § 2254 petitioner possesses a "substantial stake in the judgment of conviction which survives the satisfaction of the sentence imposed on him."[240] Moreover, *Carafas* explained that a habeas petitioner "should not be . . . required to bear the consequences of [an] assertedly unlawful conviction simply because the path has been so long that he has served his sentence."[241] Yet twenty-one years later, the Supreme Court held in *Maleng v. Cook* that regardless of any collateral consequences he may encounter, a § 2254 petitioner is not "in custody" if the sentence imposed for his underlying conviction has fully expired by the time that he files his habeas petition.[242]

In short, although collateral consequences sufficiently prevent a petition "*filed* when the petitioner was in custody from becoming moot thereafter, [these consequences] do not by themselves satisfy the requirement that the petitioner be 'in custody' when the petition is first filed—even if those consequences have actually materialized."[243] As one scholar has incisively observed, "A critical exception to [the generally] liberal stance in custody jurisprudence concerns the claims of petitioners who, having completed their sentences, remain subject to 'collateral consequences,' the gamut of legal disabilities attaching to convictions that can extend well beyond the period of confinement or conditional release."[244]

239. Carafas v. LaVallee, 391 U.S. 234, 236, 238–39 (1968).

240. *Id.* at 237 (quoting Fiswick v. United States, 329 U.S. 211, 222 (1946)).

241. *Id.* at 240.

242. 490 U.S. 488, 491 (1989); *see also* Resendiz v. Kovensky, 416 F.3d 952, 956 (9th Cir. 2005) (noting that once a sentence imposed for a conviction has "completely expired," collateral results of conviction fail to establish "custody" under § 2254).

243. HERTZ & LIEBMAN, *supra* note 98, at 433–34 n.12.

244. Logan, *supra* note 23, at 154.

Besides the obvious inconsistency between the *Carafas* rule and the *Maleng* rule, the two decades that have elapsed since *Maleng* have witnessed the steady increase in both the frequency and severity of civil disabilities that result from the fact of conviction.[245] At the state level, collateral sanctions result from misdemeanor as well as felony convictions.[246] Furthermore, the increasing use of sentence enhancements and "necessary-predicate-based" offenses has rendered the custody requirement far more complicated than it had been in previous decades.[247] As one scholar has noted, "Contemporary America, . . . faced with ever-mounting prisoner populations, and the enormous associated costs, is pursuing novel methods of social control that press the envelope of the jurisprudence of custody in unprecedented ways."[248]

B. The "Less Serious" Offenders Bear Consequences of Habeas Framework

Ostensibly, the custody requirement does not formally preclude most minor offenders from seeking § 2254 relief.[249] Defendants convicted of misdemeanors can receive sentences involving incarceration, "and some may still be detained after state avenues for testing their claims have been exhausted."[250]

Yet as a practical matter, minor offenders have little opportunity to challenge their state convictions in a federal forum. Before a federal court may review a § 2254 claim, the state petitioner must exhaust available remedies in state court.[251] A § 2254 petitioner must provide state courts with an opportunity to review his federal constitutional claims before he litigates those same claims in a federal habeas petition.[252] In other words, a state petitioner must properly present his federal constitutional claims through one

245. *See* Pinard, *Collateral Consequences*, *supra* note 16, at 634–36.

246. *See, e.g.*, Pinard, *Collateral Consequences*, *supra* note 15, at 635; Pinard, *Broadening the Holistic Mindset*, *supra* note 35, at 1073, 1078.

247. Ordower, *supra* note 28, at 1838. This "new penology" attempts to protect society through surveillance and "management of potential criminal harm on the basis of perceived risk." Logan, *supra* note 23, at 177.

248. Logan, *supra* note 23, at 149.

249. Larry W. Yackle, *State Convicts and Federal Courts: Reopening the Habeas Corpus Debate*, 91 CORNELL L. REV. 541, 562 (2006).

250. *Id.*

251. 28 U.S.C. § 2254(b)(1) (2006); O'Sullivan v. Boerckel, 526 U.S. 838, 842 (1998); *see also* Yackle, *supra* note 69, at 1001 (noting that § 2254 petitioners "need time" to "marshal their federal claims, exhaust state remedies, and frame their contentions for federal adjudication").

252. *Boerckel*, 526 U.S. at 842.

"complete round of the State's established appellate review process."[253] If the petitioner fails to comply with this "exhaustion" requirement, the petitioner generally may not obtain § 2254 review.[254]

The crucial point for determining whether a § 2254 petitioner is "in custody" is the time at which he files his petition. If the would-be § 2254 petitioner files his petition *after* his release from "custody," the federal courts lack jurisdiction and are barred from evaluating the merits of his claim.[255] Therefore, by the time a "short sentence" petitioner exhausts state remedies and is ready to file a § 2254 claim, his sentence may have already expired—thereby rendering him not "in custody."[256] As a result, the practical effect of the current framework is that "short sentence" § 2254 petitioners will no longer be "in custody" by the time they have finally exhausted state remedies and will therefore be unable to obtain § 2254 relief.

To illustrate, in Pennsylvania, a first-time offender convicted of stealing $1,000 will likely face a sentence ranging from restorative sanctions to several months of imprisonment.[257] By contrast, a repeat offender who commits a robbery in which he inflicted serious bodily harm on the victim may receive a sentence of at least twelve years in prison.[258] Under the current legal framework, this recidivist offender will likely have ample time to challenge his conviction in federal district court pursuant to § 2254. By contrast, by the time that the first-time offender "exhausts" his remedies in Pennsylvania courts, his sentence will in all likelihood have expired. Nevertheless, this comparatively more sympathetic offender may suffer collateral consequences including disqualification from obtaining certain occupational licenses. In some sense it is counter-intuitive that the prisoner serving the longer sentence as a result of (1) the severity of his offense and (2) his prior criminal record will have more of an opportunity to vindicate his constitutional rights in a

253. Woodford v. Ngo, 548 U.S. 81, 92 (2006); Jones v. Jones, 163 F.3d 285, 296 (5th Cir. 1998).

254. *Ngo*, 548 U.S. at 92; *Jones*, 163 F.3d at 296.

255. Maleng v. Cook, 490 U.S. 488, 490 (1989).

256. A "short sentence" petitioner is one whose sentence or whose "custody," as that term is currently defined, expires before he has had exhausted state remedies. Given the timing of his "release from custody" and the duration of state court review, a "short sentence" petitioner will be unable to satisfy the threshold jurisdictional requirement of § 2254.

257. 204 PA. CODE § 303.16.

258. *Id.*

federal forum than the first-time offender convicted of a relatively minor crime ever will.

The Eighth Circuit decision in *Harvey v. South Dakota* evinces the difficulties that uniquely attend "short sentence" state convictions.[259] Convicted of grand larceny in 1968, the *Harvey* petitioner was released from prison in March 1971.[260] After exhausting his state post-conviction remedies, the *Harvey* petitioner finally filed an otherwise timely § 2254 claim in 1974—three years after his release from prison.[261] In an attempt to satisfy the threshold jurisdictional requirement of custody, the *Harvey* petitioner asserted that the "disabilities" resulting from his conviction established custody within the meaning of § 2254.[262]

In concluding that the *Harvey* petitioner failed to satisfy the § 2254 custody requirement, the Eighth Circuit reasoned that collateral consequences may keep a case from becoming moot, but they cannot "suffice to give federal courts jurisdiction."[263] The Eighth Circuit also rejected the *Harvey* petitioner's contention that he had not filed his federal habeas petition while he was still "in custody" because of "state delay in processing his attempts to obtain post-conviction relief."[264]

Consequently, the current legal framework defining § 2254 custody problematically "trims the field down to convicts sentenced to substantial terms (or death)."[265] As one scholar has explained:

> Because habeas corpus today is often used by death-row inmates, more often than not the criminal's legal unveiling as a constitutional person serves instead to reveal an unworthy holder of civil rights. Important procedural questions regarding jury, police, and prosecutorial bias fade into the background of the judicial investigation and are replaced by a legal calculus based on the criminal's propensity for violence, his criminal record, and vivid descriptions of his deeds and the nature of his crime. Narratives of violence replace the law's dispassionate inquiry into the merits of punishment.[266]

259. 526 F.2d 740, 841 (8th Cir. 1975).

260. *Id.*

261. *Id.*

262. *Id.*

263. *Id.*

264. *Id.*

265. Yackle, *supra* note 258, at 562.

266. FEDERMAN, *supra* note 69, at 185.

Thus, the reality of who may obtain the "Great Privilege" appears inimical to the fundamental purpose of § 2254 review. As Chief Justice Earl Warren once perceptively noted:

> Many deep and abiding constitutional problems are encountered primarily at a level of "low visibility" in the criminal process—in the context of prosecutions for "minor" offenses which carry only short sentences. We do not believe that the Constitution contemplates that people deprived of constitutional rights at this level should be left utterly remediless and defenseless against the repetitions of unconstitutional conduct.[267]

Federal review of state convictions has constituted an integral component of this nation's federalist structure. The Supreme Court has instructed that "conventional notions of finality in criminal litigation cannot be permitted to defeat the manifest federal policy that federal constitutional rights of personal liberty shall not be denied without the fullest opportunity for plenary federal judicial review."[268] Federal habeas review pursuant to § 2254 "insures that, even though the Supremacy Clause of the Constitution already requires state courts to give criminal defendants every protection of the Bill of Rights and federal law, those defendants are also entitled to insist that a federal court review the state court proceedings."[269] Ostensibly, § 2254 review merely permits those convicted of state offenses to "contest the validity of their detention in independent, civil proceedings in [a] federal forum."[270] More fundamentally, however, § 2254 review provides a powerful mechanism for the "federal relitigation of federal questions" following state court adjudication.[271] As one scholar has keenly observed, § 2254 review constitutes far more than a mere "procedural vehicle for the protection of physical liberty."[272] Section 2254 review is instead intrinsically premised on the notion that "state criminal defendants are entitled

267. Sibron v. New York, 392 U.S. 40, 52–53 (1968) (footnote omitted).

268. Fay v. Noia, 372 U.S. 391, 424 (1963). This nation's "system of dual-level scrutiny of state incarceration" has demonstrated the "practical importance of assuring that vigorous federal review actually occurs." FREEDMAN, *supra* note 59, at 153; *see also id.* at 6 (positing that § 2254 review is "not an affront to federalism, but rather implements the theme of checks and balances that pervades our Constitutional structure").

269. FREEDMAN, *supra* note 59, at 1.

270. Yackle, *supra* note 69, at 992–93.

271. *Id.* at 992.

272. *Id.* at 997.

to litigate their federal claims in a federal forum other than the Supreme Court."[273]

Furthermore, unlike those convicted of a federal crime, state convicts cannot seek federal review of the conviction pursuant to the *coram nobis* writ.[274] The *coram nobis* writ is available to vacate a federal conviction after the sentence has been served and the defendant is no longer in custody.[275] In justifying the writ the Supreme Court has noted, "Although a term has been served, the results of the conviction may persist. Subsequent convictions may carry heavier penalties, [and] civil rights may be affected."[276] The writ helps to "bring an end to what may be substantial civil disabilities attached to criminal convictions."[277] A petitioner seeking the *coram nobis* writ must demonstrate (1) that the allegedly invalid conviction produced continuing civil disabilities (collateral consequences) and (2) "that the error is the type of defect that would have justified relief during the term of imprisonment."[278] Crucially, however, federal court review pursuant to the *coram nobis* writ is unavailable to defendants not convicted in federal court.[279]

Therefore, state convicts whose sentences expire before they exhaust state remedies lack a federal forum to challenge their con-

273. *Id.* Yackle elaborates that relitigation pursuant to § 2254 review "is appropriate not because petitioners' interest in physical liberty justifies an exemption from ordinary preclusion rules, but because criminal defendants in state court are not permitted to remove their cases to federal court when they have federal claims to raise in their defense. Because there is no opportunity for removal, it is essential that postconviction habeas be available to ensure the choice of a federal forum—at some point." *Id.*

274. 28 U.S.C. § 1651(a) (2006). State court "*coram nobis* remedies" also exist. Yackle, *supra* note 69, at 1003.

275. *See* United States v. Morgan, 346 U.S. 502, 512 (1954). The writ "has been exclusively used by petitioners who have not yet commenced serving their sentence or have completed service of their sentence." Neyor v. INS, 155 Fed. Supp. 2d 126, 136 (D.N.J. 2001). Limited in its availability, the writ is appropriate "only under circumstances compelling such action to achieve justice." *Morgan,* 346 U.S. at 511.

276. *Morgan,* 346 U.S. at 512–13.

277. United States v. Keane, 852 F.2d 199, 203 (7th Cir. 1988).

278. *Id.* at 203; *see also Morgan,* 346 U.S. at 512–13 ("Although the term has been served, the results of the conviction may persist. . . . [R]espondent is entitled to an opportunity to attempt to show that this conviction was invalid.").

279. *Neyor,* 155 F. Supp. 2d at 136. To illustrate, in *Ogunwomoju v. United States* the petitioner, a Nigerian citizen, was convicted of state drug offenses. 512 F.3d 69, 70 (2d Cir. 2008). Facing deportation, he attempted to challenge the legality of the state court judgment through the *coram nobis* writ. *Id.* at 75. Rejecting the *Ogunwomoju* petitioner's argument, the Second Circuit held that federal courts lack jurisdiction to grant the writ with respect to state convictions because courts have traditionally used it "to correct errors within their own jurisdiction." *Id.*

victions—regardless of any collateral consequences that they may encounter or the underlying merit of their constitutional claims. Several factors necessitate a critical reevaluation of the present § 2254 custody framework, including (1) the continued importance of the federal forum that § 2254 review provides, (2) the recent expansion of collateral sanctions resulting from state convictions, and (3) the departure from the traditional "direct/collateral" distinction that *Padilla* effectuated.

C. The Imprudence of an Overly Broad Definition of "Custody"

As the scholar Larry W. Yackle has aptly noted, "[A]s a docket control mechanism, the 'custody' doctrine has come under enormous pressure to give way in the interest of providing meaningful relitigation opportunities to [federal habeas] applicants."[280] Yet the custody requirement persists as a vital "gatekeeping" device for determining which state convicts should benefit from § 2254 review.[281] Consequently, arguments in favor of jettisoning or indiscriminately liberalizing the custody requirement are, at a minimum, tenuous.[282]

An overly broad interpretation of "custody" contravenes the intrinsic purpose of § 2254 review.[283] The custody requirement is "no mere artificial prerequisite to a habeas action, designed to restrict access to those most in need of judicial attention. It is part and parcel of what habeas corpus is, what it means, or, at least, what it has been and meant traditionally."[284]

As the Supreme Court noted in *Hensley*, "Since habeas corpus is an extraordinary remedy whose operation is to a large extent uninhibited by traditional rules of finality and federalism, its use has been limited to cases of special urgency, leaving more conventional remedies for cases in which the restraints on liberty are neither severe nor immediate."[285] Consistent with its extraordinary nature, habeas corpus should not be "casually available."[286] Thus in order for the § 2254 custody requirement to meaningfully function

280. Yackle, *supra* note 69, at 1001.

281. *Id*. at 999.

282. *But see id.* at 1009–10 (asserting that custody serves a "symbolic function").

283. *See Developments*, *supra* note 90, at 1076.

284. Yackle, *supra* note 69, at 999.

285. Hensley v. Mun. Court, San Jose Milpitas Judicial Dist., 411 U.S. 345, 351 (1973).

286. Lefkowitz v. Fair, 816 F.2d 17, 19 (1st Cir. 1987).

as a gatekeeper to the "Great Privilege," the "definition of the term must be constrained by its natural legal boundaries."[287]

To that end, some courts have understandably cautioned that allowing collateral consequences to establish custody "would render the 'in custody' requirement of [§ 2254] a nullity, and extend the habeas corpus remedy to all persons convicted of crimes. It would ignore the [notion] that the writ of habeas corpus is available only to remedy 'severe restraints on individual liberty.' "[288] Furthermore, individual victims, the government, and society as a whole all have a significant interest in the finality of convictions.[289] Jurisdictions other than the one imposing the sentence acquire "an interest [in finality] as well, as they may then use that conviction for their own recidivist sentencing purposes, relying on the presumption that regularly attaches to final judgments."[290] Establishing custody on the basis of some minor collateral harm inappropriately marginalizes the compelling need for finality.

Moreover, as the Supreme Court noted in *Coss,* concerns about the "ease of administration of challenges to expired state convictions" weigh against an overly inclusive definition of the § 2254 custody requirement.[291] Federal courts reviewing § 2254 claims

> must consult state court records and transcripts to ensure that challenged convictions were obtained in a manner consistent with constitutional demands. As time passes, and certainly once a state sentence has been served to completion, the likelihood that trial records will be retained by the local courts and will be accessible for review diminishes substantially.[292]

Additionally, federal habeas review pursuant to § 2254 places a significant burden on federal judicial resources.[293] Permitting a state convict to establish custody merely because of some purported collateral harm resulting from his conviction ignores the realities of an already overcrowded federal docket.[294]

---

287. *Id.*

288. Furey v. Hyland, 395 F. Supp. 1356, 1360 (D.N.J. 1975); *see also Lefkowitz*, 816 F.2d at 20 (similarly expressing concern about liberalizing the "custody" requirement); Harvey v. South Dakota, 526 F.2d 840, 841 (8th Cir. 1975) (stating that existence of custody based solely on "collateral consequences" would render Congress's "custody" requirement meaningless).

289. Lackawanna Cnty. Dist. Attorney v. Coss, 532 U.S. 394, 402 (2011).

290. *Id.* at 403.

291. *Id.*

292. *Id.*

293. FEDERMAN, *supra* note 69, at 159.

294. Yackle, *supra* note 59, at 562 (noting that having no "threshold" for review of § 2254 claims would be "unrealistic").

The watershed *Padilla* decision, viewed in light of the increasing prevalence and severity of collateral consequences, suggests that some consequences may be sufficient to establish § 2254 custody. Concomitantly, however, any proposed revision to the fundamental habeas jurisdictional requirement of custody should not automatically establish custody upon the mere existence of a collateral consequence.[295] A statutory solution is needed that will allow courts to determine which collateral consequences qualify as restraints sufficient to warrant habeas review under § 2254.

III.
THE STATUTORY SOLUTION

As the Supreme Court instructed in *Peyton v. Rowe*, the "development of the writ of habeas corpus did not end in 1789."[296] Throughout the more than two centuries during which the habeas corpus writ has existed in America, Congress has played an important role in remedying problems concerning the scope and availability of the Great Writ—and thus adapting it to compelling national needs. Under Article III of the U.S. Constitution, Congress is charged with defining the jurisdiction of the federal courts. Because custody is fundamentally a jurisdictional issue, an effective solution to the collateral sanctions and § 2254 custody dilemma will require Congressional intervention.

Congressional intervention will also have the benefit of ensuring a degree of uniformity in federal case law that confronts the relationship between § 2254 custody and collateral sanctions. Such uniformity is particularly crucial given the immense number and variety of collateral consequences that may result from state convictions.

Any legislative solution must draw upon traditional judicial definitions of § 2254 custody. The *Padilla* decision discussed but avoided a formalistic reliance on the "direct/collateral" distinction; a statutory solution to the § 2254 custody issue should incorporate the *Padilla* Court's functional analysis. A flexible standard will allow federal courts to meaningfully determine whether a given collateral consequence sufficiently establishes custody and will prevent an undesirably broad expansion of the § 2254 "custody" definition.

295. For this reason, previous scholarship that has posited that § 2254 "custody" may exist merely because of the presence of a given restraint, such as a registration requirement, is unpersuasive.

296. Peyton v. Rowe, 391 U.S. 54, 66 (1968).

A. The Need for Congressional Intervention

A petitioner seeking relief pursuant to § 2254 must demonstrate that he is in custody pursuant to the judgment of a state court.[297] As the Supreme Court noted in *Jones v. Cunningham*, this "in custody" requirement constitutes a necessary jurisdictional element of § 2254 relief.[298] Because of the inherently jurisdictional nature of the custody requirement, congressional action is required to provide a substantive solution to the collateral sanction and custody dilemma.

Under Article III of the U.S. Constitution, Congress possesses cardinal authority for regulating the jurisdiction of the federal judiciary.[299] Article III instructs that the "judicial Power of the United States, shall be vested in one supreme Court, and in such inferior Courts as the Congress may from time to time ordain and establish."[300]

At the same time, however, the scope of federal jurisdiction and the authority to define that jurisdiction have dynamically developed "through a dialogic process of congressional enactment and judicial response."[301] Thus both Congress and the Supreme Court exercise consequential roles in determining the jurisdiction of the federal judiciary. As one scholar has explained, where an issue concerning the proper scope of federal jurisdiction exists,

> the Court and Congress simply express their opinions in the manner unique to each branch. The Court decides cases and writes opinions that establish doctrines governing the exercise of jurisdiction. Congress considers and passes legislation that governs the exercise of jurisdiction. Through these vehicles, the contours of federal jurisdiction are resolved.[302]

297. 28 U.S.C. § 2254(a) (2006); *see also, e.g.*, McCormick v. Kline, 572 F.3d 841, 847 (10th Cir. 2009) (noting this requirement).

298. *See, e.g.*, Jones v. Cunningham, 371 U.S. 236, 238 (1963) (noting that the "custody" requirement is "jurisdictional"); Erlandson v. Northglen Mun. Court, 528 F.3d 785, 788 (10th Cir. 2008) (terming the "in custody" requirement "jurisdictional").

299. Barry Friedman, *A Different Dialogue: The Supreme Court, Congress and Federal Jurisdiction,* 85 NW. U. L. REV. 1, 1–2 (1990) ("[N]o one has challenged the central assumption that Congress bears primary responsibility for defining federal court jurisdiction.").

300. U.S. CONST. art. III, § 1.

301. Friedman, *supra* note 299, at 2. Friedman elaborates that through "this dialogic process between Congress and the Court . . . the content of federal jurisdiction ultimately is determined." *Id.* at 10.

302. *Id.* at 53 (footnote omitted).

Factors relevant to determining the proper scope of federal jurisdiction include the "protection of federal rights and interests, . . . comity and federalism, caseloads and the extent of judicial resources, and the need for uniformity."[303] Furthermore, within the habeas context, the collaborative relationship between Congress and the Supreme Court promotes the equally imperative interests of finality and individual liberty. Habeas review under § 2254 illustrates "how the Court will take a jurisdictional grant and transform it to serve judicial purposes, all the while engaging Congress in a fairly cooperative discussion concerning the breadth of that exercise of federal judicial authority."[304] In short, the development of jurisdiction under § 2254 has been "a cooperative, and largely cordial, enterprise [between Congress and the Supreme Court], but with the Court demonstrably taking the leading role."[305]

The historical development of the federal habeas writ as a means to review state convictions constitutes a powerful manifestation of this collaborative relationship. In 1867, Congress granted the lower federal courts jurisdiction to review the claims of state convicts challenging the federal constitutionality of their convictions.[306] However, this post-Civil War provision was largely ineffectual; not until the 1950s did the Supreme Court substantially expand the power of federal courts to review habeas petitions challenging state convictions.[307] Nevertheless, despite its weaknesses, the 1867 act "codified the budding relationship that prisoners would have with the national judiciary, to the states, under the Fourteenth Amendment."[308] Through this Reconstruction-era piece of legislation, Congress dramatically altered habeas jurisdiction, enabling the "Great Writ" to eventually emerge as a bulwark for federal constitutional rights during the twentieth-century.[309]

303. *Id.* at 52.

304. *Id.* at 10. Friedman also adds that § 2254 review is "one of the most controversial exercises of federal judicial power." *Id.* at 11.

305. *Id.* at 13.

306. FEDERMAN, *supra* note 69, at 26; Friedman, *supra* note 299, at 11.

307. FEDERMAN, *supra* note 69, at 26; Friedman, *supra* note 299, at 13 (noting that regardless of Congress' precise "intent" in 1867, "it was the Court, approximately eighty years later, that brought habeas review to full bloom"). In 1868, Congress temporarily suspended this habeas jurisdiction, restoring it in 1885. FEDERMAN, *supra* note 69, at x.

308. FEDERMAN, *supra* note 69, at 26. "The act . . . marks . . . [the] passage of power from state court to Congress (as an overseer of federal court jurisdiction) . . . ." *Id.*

309. *Id.* During the apex of the Warren Court's expansion of civil rights and federal power, one contemporary observer noted that it was a "somewhat extraordinary process by which the 1867 statute, after reposing almost quiescent for

Meaningful Congressional action has thus helped adapt the jurisdiction of the federal judiciary to compelling societal needs. The historical evolution of § 2254 review has saliently demonstrated this flexibility.[310] When amending the § 2254 grant of jurisdiction, "Congress either has revised the statute to reflect judicial decisions, or has sanctioned judicial interpretations of the statute in the reenacting legislative history."[311]

The 2010 Supreme Court decision in *Padilla* will likely revolutionize how courts view collateral consequences. Although *Padilla* implicated only Sixth Amendment ineffective assistance of counsel claims, the paradigm shift it announced may have implications that extend to the § 2254 context. Post-*Padilla*, federal courts will likely be faced with the issue of whether collateral sanctions can sufficiently establish custody under § 2254. Consistent with its historical and constitutional role within the habeas context, Congress must again act to provide the federal judiciary with a meaningful framework to resolve this issue.

As an additional matter, Congressional intervention will have the coincident benefit of ensuring a degree of uniformity in federal case law analyzing the relationship between § 2254 custody and collateral sanctions. Given the immense number and variety of collateral consequences that may result from a conviction in each of the states, such uniformity is particularly desirable.

However, any Congressional solution must not categorically dictate which collateral sanctions establish § 2254 custody and which do not. Instead, it should provide substantive factors for federal courts to consider when determining whether a petitioner has sufficiently established custody based on a given collateral consequence.

To that end, a Congressional solution should avoid a "laundry list" approach to resolving the § 2254 custody and collateral sanc-

---

decades, has proved in more recent years to have been a sleeping giant, capable, when aroused, of . . . astonishing results." Lewis Mayers, *The Habeas Corpus Act of 1867: The Supreme Court as Legal Historian*, 33 U. CHI. L. REV. 31, 31 (1965).

310. FEDERMAN, *supra* note 69, at 26. In 1996, Congress, through AEDPA, enacted "a relatively modest set of reforms" that sought to expedite habeas proceedings and clarify the deference that federal courts should accord to state court interpretations of federal constitutional law. FREEDMAN, *supra* note 59, at 153; *see also* Adam N. Steinman, *Reconceptualizing Federal Habeas Corpus for State Prisoners: How Should AEDPA's Standard of Review Operate After* Williams v. Taylor?, 2001 WIS. L. REV. 1493, 1502–03 (2001) (noting the interaction between Supreme Court decision in *Williams v. Taylor* and AEDPA).

311. Friedman, *supra* note 299, at 12; *see also* FEDERMAN, *supra* note 69, at 21–26 (exploring relationship between Congress and habeas corpus).

tions issue. Criticizing the *Padilla* decision, Justice Scalia asserted in his dissenting opinion that *Padilla* "prevents legislation that could solve the problems addressed by today's opinions in a more precise and targeted fashion."[312] He added that "legislation could specify which categories of misadvice about matters ancillary to the prosecution invalidate plea agreements, what collateral consequences counsel must bring to a defendant's attention, and what warnings must be given."[313]

Congressional intervention certainly has the ability to effectively resolve incongruities in federal law—particularly where these incongruities implicate the jurisdiction of federal courts. Yet the formalistic approach that Justice Scalia envisioned in the ineffective assistance of counsel context is profoundly inapplicable to the § 2254 custody and collateral sanctions dilemma. As a practical matter, it would be nearly impossible for Congress to consider every potential collateral sanction that might accompany a state criminal conviction and categorize each one as either sufficient or insufficient for § 2254 custody. Instead, because of the considerable number and variety of collateral sanctions, a practicable legislative solution must provide courts with sufficient flexibility to determine whether a given sanction establishes § 2254 custody.

An appropriate Congressional framework must therefore consider the existing habeas jurisprudence that the federal judiciary has constructed over the past several decades. A meaningful solution to the § 2254 custody and collateral sanctions dilemma should wisely continue the cooperative discussion between Congress and the federal judiciary that the Great Writ has historically necessitated.

In short, the federal judiciary, in determining the appropriate relationship between § 2254 custody and collateral sanctions post-*Padilla*, will likely need substantive guidance. It is entirely appropriate—and critical—for Congress to provide it.

B. Articulating the Essence of Custody

Over the past fifty years, federal courts have identified several factors that are relevant to determining whether custody under § 2254 exists. Given the "dialogic" nature of jurisdiction, any Congressional solution must acknowledge the well-established conceptions of § 2254 custody that the federal judiciary has crafted. In order to appropriately devise the remedial statutory language, Con-

312. Padilla v. Kentucky, 130 S. Ct. 1473, 1496 (2010) (Scalia, J., dissenting).
313. *Id.* at 1496.

gress must articulate the essence of § 2254 custody and include the following elements.

1. The Nature of the Restraints

The Supreme Court in *Jones* instructed that § 2254 custody encompasses those "conditions and restrictions" that "significantly restrain" the petitioner's liberty "to do those things which in this country free men are entitled to do."[314] Thus, although the *Jones* petitioner's parole released him from immediate physical imprisonment, it "impose[d] conditions which significantly confine[d] and restrain[ed] his freedom."[315] Federal courts over the past five decades have reaffirmed that an exclusive emphasis on physical restraint contradicts the contemporary purpose of § 2254 review.[316]

As some federal courts have articulated, custody requires (1) "significant restraints" not "shared by the public generally" and (2) "some type of continuing governmental supervision."[317] The Supreme Court in *Hensley* indicated that the necessary harm establishing custody must be "severe" and "immediate."[318] Furthermore, courts should assess whether a determination that custody exists impedes on a "significant interest of the State."[319] This acknowledgment of state interests is vital given that

> [f]ederal supervision of state criminal prosecutions through [§ 2254] departs from traditional notions of deference owed state administration of federal law; problems of federalism aside, ordinary concepts of finality in the judicial process are displaced by the continuing availability of habeas for review of restrictions imposed by the judgments of federal courts. Therefore, the restraints which have been thought appropriate for review in habeas proceedings are those which impinge with es-

314. Jones v. Cunningham, 371 U.S. 236, 240 (1963). Because *Jones* abandoned a strict adherence to the notion of "physical custody," the "significance of the types of facts in *Jones* should be seen in terms of the severity of the restraints they describe." *Developments*, *supra* note 90, at 1075–76.

315. *Jones,* 371 U.S. at 243.

316. *Developments*, *supra* note 90, at 1076 (noting that "physical custody requirement was rooted" in procedural nature of the writ as "device compelling [government] to bring the prisoner before the court" but that habeas "no longer serves that purpose").

317. Barry v. Bergen Cnty. Prob. Dep't, 128 F.3d 151, 159 (3d Cir. 1997); Lefkowitz v. Fair, 816 F.2d 17, 19 (1st Cir. 1987); *see also* Justices of Bos. Mun. Court v. Lydon, 466 U.S. 294, 300–01 (1984).

318. Hensley v. Mun. Court, San Jose Milpitas Judicial Dist., 411 U.S. 345, 351 (1973); Harvey v. South Dakota, 526 F.2d 840, 841 (8th Cir. 1975).

319. *Hensley,* 411 U.S. at 352.

pecial harshness on personal liberty—those severe enough to warrant relitigation.[320]

2. Nexus

Consistent with the essential function of federal habeas review, it is imperative that a § 2254 petitioner establish a nexus between the custody and the restraint that he seeks to challenge. As the Ninth Circuit recently instructed, the language of § 2254(a) "explicitly requires a nexus between the petitioner's claim and the unlawful nature of the custody."[321]

Importantly, however, some Supreme Court precedents have not required a direct nexus between the custody and the "challenge" to the custody. To illustrate, *Peyton v. Rowe* held that a prisoner satisfies the custody requirement where he is serving the first of two consecutive sentences and challenges in his habeas petition the second conviction or sentence.[322] Consistent with the *Peyton* rule, a habeas petitioner may challenge his state sentences even though he is not presently serving them due to his incarceration in federal prison on federal charges.[323] Likewise, a state prisoner who is serving consecutive state sentences is in custody and may attack the sentence scheduled to run first, even after it has expired, until all sentences have been served—at least as long as they "continue to postpone the date for which he would be eligible for [release]" under the expired sentence.[324] When determining whether a petitioner has satisfied the § 2254 custody requirement, courts must view "consecutive sentences in the aggregate, not as discrete segments."[325]

Yet regardless of the judicial willingness in these contexts to liberally construe the relationship between the restraint and the challenge, nexus remains a crucial component of § 2254 custody.[326] A nexus requirement ensures that the challenged harm is the same harm that the § 2254 petitioner is putatively suffering. Within the context of collateral consequences, a nexus requirement effectively

320. *Developments, supra* note 90, at 1073.

321. Bailey v. Hill, 599 F.3d 976, 980 (9th Cir. 2010).

322. Peyton v. Rowe, 391 U.S. 54, 64–65 (1968).

323. Maleng v. Cook, 490 U.S. 488, 493–94 (1989).

324. Garlotte v. Fordice, 515 U.S. 39, 43 (1995).

325. *Id.* at 47.

326. *See, e.g.*, *Bailey*, 599 F.3d at 977–80 (noting importance of nexus); Washington v. Smith, 564 F.3d 1350, 1350–51 (7th Cir. 2009) (explaining importance of nexus).

avoids an inappropriately broad definition of custody.[327] Therefore, a nexus requirement must remain an essential part of any statutory solution.

3. Harm to Person—Not Merely Property

On a fundamental level, a § 2254 petition must demonstrate an actual restraint on the petitioner's person—and not just a financial interest. A challenge to the validity of a conviction must satisfy the "case or controversy" requirement under Article III of the U.S. Constitution.[328] The petitioner must demonstrate "a concrete injury, caused by the conviction and redressable by invalidation of the conviction."[329] Throughout the litigation, the party seeking relief must have suffered or "be threatened with an actual injury . . . [that is] likely to be redressed by a favorable judicial decision."[330] Additionally, under any proposed framework concerning collateral sanctions, § 2254 petitioners must bear the burden of demonstrating harm. Certainly the notion of harm is inextricably linked to the concept of custody—the petitioner seeks release from a restraint imposed on him by the conviction.[331]

Importantly, courts have consistently held that habeas relief "has traditionally been concerned with liberty rather than property, with freedom more than economics."[332] This conception of custody is not unreasonable given that in Latin, "habeas corpus" translates to "you have the body."[333] Therefore, under any proposed statutory solution, for a collateral sanction to constitute "custody," it must harm the liberty of the person seeking relief, and not merely some economic interest of the person.

327. As a hypothetical illustration, a non-citizen is convicted of an offense in Delaware that, under federal law, does not constitute a basis for deportation. Shortly after this Delaware conviction, this non-citizen is convicted of an offense in New York that, under federal law, constitutes a basis for deportation. The "nexus" requirement mandates that where the non-citizen attempts to establish § 2254 custody on the basis of his imminent deportation, he may challenge only the New York conviction—not the Delaware conviction.

328. U.S. CONST. art. III, § 2.

329. Spencer v. Kenna, 523 U.S. 1, 7 (1998).

330. *Id.*

331. *See, e.g.*, Jones v. Cunningham, 371 U.S. 236, 239–40 (1963).

332. *See, e.g.*, Lefkowitz v. Fair, 816 F.2d 17, 20 (1st Cir. 1987) (noting this compelling principle).

333. BLACK'S LAW DICTIONARY 778 (9th ed. 2009).

4. The Identity of the Respondent

Additionally, in order to satisfy the custody requirement, a § 2254 petitioner must have a determinable respondent. Section 2254 requires that a habeas petitioner must name "the state officer having custody of him or her as the respondent to the petition."[334] As some courts have noted, because "the custodian is the state's agent—and the state is therefore the custodian's principal—the state may waive the lack of personal jurisdiction on the custodian's behalf."[335]

Therefore, the collateral sanction that purportedly creates "custody" must include an element of governmental action. The more attenuated the relationship is between governmental enforcement of the sanction and the harm that the § 2254 petitioner alleges, the less likely it is that the sanction will establish "custody." As the Supreme Court has perceptively noted, habeas corpus relief "does not act upon the prisoner who seeks relief, but upon the person who holds him in what is alleged to be unlawful custody."[336] *Jones* provides some guidance in this context, instructing that where the petitioner was on parole, the respondents should be the individual members of the parole board, not the superintendent of the penitentiary system.

C. The Padilla *Paradigm Shift Dispenses with Formalism*

As a result of *Padilla*, collateral sanctions attained legal relevancy as a basis for Sixth Amendment ineffective assistance of counsel claims. Although acknowledging the "direct" and "collateral" distinction, *Padilla* rejected a formalistic insistence that collateral sanctions were, for purposes of post-conviction review, legally insignificant. Although *Padilla* is a Sixth Amendment "assistance of counsel" case, it nevertheless informs how the federal judiciary and Congress should view collateral sanctions within the analogous context of § 2254 custody.

Before 2010, at least one federal circuit court had justified the rule that collateral sanctions do not establish custody for purposes of § 2254 review by invoking the pre-*Padilla* rule that attorneys, in order to provide constitutionally sufficient counsel, did not have to

334. *See, e.g.*, Smith v. Idaho, 392 F.3d 350, 355 (9th Cir. 2004) (quoting Stanley v. Cal. Supreme Court, 21 F.3d 359, 360 (9th Cir. 1994)).

335. *Id.* at 356.

336. Braden v. 30th Judicial Circuit Court of Ky., 410 U.S. 484, 494–95 (1972). *Jones* provides some guidance in this context, instructing that where the petitioner was on parole, the respondents should be the parole board, not the superintendent of the penitentiary system. 371 U.S. at 241-42.

advise their clients regarding the collateral sanctions that may result from a guilty plea; in the 2005 *Resendiz v. Kovensky* decision, the Ninth Circuit held that the collateral sanction of deportation did not establish § 2254 custody.[337] In reaching this conclusion *Resendiz* relied on the reasoning that deportation is "wholly independent of the court imposing the sentence. . . . Removal is not part of the sentence."[338] The Ninth Circuit added, "Extending that holding, we have similarly concluded that, because immigration consequences remain collateral, the failure of counsel to advise his client of the potential immigration consequences of a conviction does not violate the Sixth Amendment right to effective assistance of counsel."[339]

Yet *Padilla* instructed that the "collateral versus direct distinction is . . . ill-suited to evaluating a [Sixth Amendment ineffective assistance of counsel] claim concerning the specific risk of deportation."[340] Within the Sixth Amendment ineffective assistance of counsel context, therefore, *Padilla* appeared to abandon the prior formalistic assumption that collateral sanctions, unlike direct sanctions, were of negligible legal consequence.[341]

Padilla reached its holding by considering several factors. *Padilla* analyzed whether the collateral sanction was "intimately related to the criminal process" and was an "integral part" of the penalty.[342] Furthermore, *Padilla* emphasized the seriousness of deportation as well as its impact on the defendant's family.[343] *Padilla* additionally reasoned that deportation resulting from a conviction may be more important to the defendant than any potential sentence.[344]

A comparison of *Padilla* with the Third Circuit's 2003 decision in *Drakes v. INS* evinces the dramatic transformation that *Padilla* effectuated. *Drakes* termed deportation a "civil action" and asserted that even "if removal involves a greater potential injury to a petitioner than an enhanced sentence, such an injury does not outweigh the interests of finality and ease of administration."[345]

337. Resendiz v. Kovensky, 416 F.3d 952, 956–57 (9th Cir. 2005).

338. *Id.* at 957 (emphasis omitted) (quoting United States v. Amador-Leal, 276 F.3d 511, 516 (9th Cir. 2002)).

339. *Id.*

340. Padilla v. Kentucky, 130 S. Ct. 1473, 1482 (2010).

341. Smyth, *supra* note 32, at 798 (asserting that *Padilla* rejected "formalism").

342. *Padilla,* 130 S. Ct. at 1476.

343. *Id.* at 1481, 1486.

344. *Id.* at 1483.

345. Drakes v. INS, 330 F.3d 600, 605 (3d Cir. 2003).

Although also acknowledging that removal proceedings are "civil in nature," *Padilla* reasoned that removal is "nevertheless intimately related to the criminal process."[346]

Some observers have persuasively suggested that *Padilla* necessitates a "new, more realistic terminology and legal analysis."[347] Consistent with this view, factors such as the "severity" or "enmeshed nature" of the consequence might replace the traditional use of the direct and collateral distinction.[348] Additionally, the "severity" of the "enmeshed penalty" to the defendant should be analyzed "relative to the offense and its traditional criminal penalties."[349]

D. The Proposed Statutory Solution

Drawing upon well-established § 2254 jurisprudence, as well as upon the *Padilla* language, Congress should add to § 2254 the following definition of "custody":

> (a)(1) To satisfy the custody requirement, a petitioner whose term of imprisonment has expired bears the burden of establishing by a preponderance of the evidence that (i) he is subject to significant restraints not shared by the public generally and (ii) governmental action created these restraints.
>
> (a)(2) In assessing whether a petitioner has satisfied the requirements of subsection (a)(1), courts should consider: (i) the permanency of the restraints; (ii) the degree to which the restraints are "intimately related to the criminal process"; (iii) the relative severity of the restraint to the petitioner; (iv) the degree to which the severity of the restraint exceeds the severity of the sentence itself; and (v) the lack of relief from the restraint through means other than those provided under this section.
>
> (a)(3) The petitioner bears the burden of establishing a nexus between this "custody" and the relief that he seeks.
>
> (a)(4) The respondent shall be the governmental entity responsible for enforcing the restraints on the petitioner.
>
> (a)(5) The determination of whether petitioner satisfies the "custody" requirement will be made based on the restraints that petitioner claims at the time that he files his petition pursuant to this section.

346. *Padilla,* 130 S. Ct. at 1481.

347. Smyth, *supra* note 32, at 802.

348. *Id.* at 802, 823–24.

349. *Id.* at 823–24 (noting that "measure of relative severity" assesses whether the "enmeshed penalty overshadows the traditional criminal penalty").

Given the guidance that the Supreme Court provided in *Padilla,* bright-line rules that categorically establish or reject custody will often produce unjust results. The complete rejection of collateral consequences as a basis for establishing § 2254 custody deprives some state convicts of § 2254 relief merely because they completed state court review after their sentences had already expired, regardless of any severe collateral consequences that they encounter. At the same time, however, if courts indiscriminately deem "collateral consequences" sufficient to establish § 2254 custody, state convicts will be able to obtain federal habeas review of their convictions merely because of some alleged "collateral harm." Whereas the first result ignores the contemporary reality of collateral consequences, the second result eviscerates the threshold requirement of custody.

Thus any legislative solution to the § 2254 custody issue should eschew the formalism that had characterized the *Carafas/Maleng* paradox. As the Supreme Court has instructed:

> [W]e have consistently rejected interpretations of the habeas corpus statute that would suffocate the writ in stifling formalism or hobble its effectiveness with the manacles of arcane and scholastic procedural requirements. The demand for speed, flexibility, and simplicity is clearly evident in our decisions concerning the exhaustion doctrine, the criteria for relitigation of factual questions, the prematurity doctrine, the choice of forum, and the procedural requirements of a habeas corpus hearing. That same theme has indelibly marked our construction of the statute's custody requirement.[350]

Instead, the proposed amendment to § 2254 offers a functional framework that provides federal judges with greater flexibility and discretion.[351] Perhaps more importantly, this statutory framework not only accommodates the sometimes-conflicting values of federal judicial review and state autonomy, but also effectively strikes the critical balance between the societal interests in finality and individual liberty.[352] Furthermore, this statutory frame-

350. Hensley v. Mun. Court, San Jose Milpitas Judicial Dist., 411 U.S. 345, 350 (1973) (citations omitted) ("[I]nterpretation of the Great Writ must retain the ability to cut through barriers of form and procedural mazes.").

351. *See, e.g.*, Allison H. Eid, *Federalism and Formalism*, 11 WM. & MARY BILL RTS. J. 1191, 1197–98 (2003) (contrasting formalism with functionalism). Within the context "of setting forth boundaries for future policymaking, formalism favors rules because they give clear guidance, whereas functionalism favors standards because they allow for flexibility." *Id.* at 1197.

352. As one scholar has summarized, "Efforts to conceptualize the federal habeas writ [necessitate balancing] the federal government's interest in enforcing

work maintains the inherently limited power of federal habeas review while concomitantly acknowledging the realities of collateral sanctions as a penal mechanism.

In addition, the proposed statutory solution addresses the permanency of the purported restraint as well as whether means other than § 2254 habeas review could ameliorate it. Given that habeas is properly an extraordinary writ, federal courts should not be needlessly burdened with resolving issues that will either "self-resolve," or that should be resolved by entities other than Article III courts.

For example, a driver's license suspension would almost invariably fail to establish custody, given the temporary nature of the restraint.[353] Likewise, the inability to possess a firearm would be insufficient to establish custody because of the restraint's relative lack of severity. By contrast, deportation, given the permanency of the restraint, may establish custody when the petitioner demonstrates that the deportation exceeds the severity of the sentence itself. Consistent with the *Padilla* factors, the petitioner would bear the burden of demonstrating the relative severity of the restraint on him and his family.[354]

Although the Pennsylvania Superior Court has based a conclusion of ineffective assistance of counsel on the loss of a teacher's pension, this collateral consequence would also likely fail to establish custody under this proposed statutory framework.[355] A pension implicates only a defendant's financial interests and does not constitute a restraint on his liberty within the meaning of § 2254. In short, the predominantly economic nature of a restraint militates against a determination that the restraint establishes custody. Given the substantial extent to which previous § 2254 custody jurispru-

federal rights with the state government's interest in the finality of its convictions. They also focus on the rights of the individual, analyzing the effect of federal habeas on the ability of individuals to assert the rights provided them under federal law." Steinman, *supra* note 310, at 1494.

353. *See* Westberry v. Keith, 434 F.2d 623, 624–25 (5th Cir. 1970); Lillios v. New Hampshire, 788 F.2d 60, 61 (1st Cir. 1986).

354. Under this proposed statutory framework, a non-citizen convicted of multiple robberies and sentenced to fifteen years of incarceration will be unlikely to demonstrate that the severity of his deportation outweighs the severity of his sentence. However, a non-citizen convicted of a minor drug offense and sentenced to two years or probation may be able to demonstrate that the severity of deportation outweighs the severity of his sentence. The presence of this "less serious" offender's family in the United States would be relevant to establishing the severity of his deportation.

355. Commonwealth v. Abraham, 996 A.2d 1090, 1095 (Pa. Super. Ct. 2010).

dence influences this proposed statutory framework, the harm must be personal and not merely financial.[356]

As a result, monetary losses that are a collateral consequence of a sentence will be less likely to satisfy the custody requirement. To illustrate, a person convicted of a drug offense may lose his or her right to reside in public housing or obtain subsidies. Pursuant to the proposed statutory framework, this collateral sanction should be insufficient to establish custody under § 2254 when the sanction adversely affects primarily the offender.

Yet a single mother who cannot obtain essential public housing or benefits in order to regain custody of her children because of a single felony conviction may present a different situation. As some observers have noted, the current public assistance framework establishes in some cases a "lifetime ban on cash assistance and food stamps for individuals with felony drug convictions. Currently, there is no good cause or hardship exemption for parents who resume caretaking responsibilities for their children upon reentry [from prison]."[357] In this case, because of the severity of the restraint and the mother's resulting inability to care for her family, this would-be § 2254 petitioner might satisfy the custody requirement.

As a final consideration, under this proposed statutory framework, petitioners must still adhere to the exhaustion requirement as well as the one-year statute of limitations.[358] Thus a prospective § 2254 petitioner would have to file his petition within one year of the conclusion of state post-conviction review. Consequently, petitioners could not improperly use purported collateral harms to subvert the fundamental time and procedural restrictions that the federal habeas statute currently requires. This statutory proposal thereby preserves essential limits on the availability of § 2254 review.

356. *See supra* notes 182–198.

357. Marne L. Lenox, Note, *Neutralizing the Gendered Collateral Consequences of the War on Drugs,* 86 N.Y.U. L. REV. 280, 297 (2011); *see also id.* at 299 (noting that loss of benefit frustrates "a parent's ability to resume caretaking responsibilities upon reentry, an element particularly critical to ex-offender parents of children in foster care"); Michael Pinard & Anthony C. Thompson, *Offender Reentry and the Collateral Consequences of Criminal Convictions: An Introduction,* 30 N.Y.U. REV. L. & SOC. CHANGE 585, 600 (2006) (noting that welfare laws reduce ex-offenders' access "to benefits that might provide transitional support as they seek employment").

358. *See supra* Part II.B for a brief discussion of these requirements.

CONCLUSION

Although collateral sanctions prevent a § 2254 petition filed when the petitioner was in custody from becoming moot, for the past two decades, federal courts have categorically held that these same sanctions fail to establish custody under § 2254 where the petitioner's sentence has already expired. The 2010 *Padilla* decision, however, marked a critical shift in the legal significance that collateral sanctions possess for purposes of Sixth Amendment ineffective assistance of counsel claims. Federal courts may decide to analogously extend the *Padilla* analysis to the issue of whether some collateral consequences are sufficient to establish § 2254 custody.

Because custody is inherently a jurisdictional issue, Congress must ultimately provide a solution to this dilemma. A statutory reform to § 2254 should draw upon both the *Padilla* analysis as well as the nearly five decades of custody jurisprudence that the federal courts have carefully developed. The legislative reform that this Article proposes effectively avoids an inappropriately broad definition of § 2254 custody that would contravene the fundamental purpose of federal habeas review and disregard the compelling societal interest in finality. Instead, the proposed statutory reform maintains that those seeking to claim relief under the "Great Privilege"—as Chief Justice Marshall aptly termed it—must demonstrate harm sufficient to establish custody.

In rejecting the proposition that collateral consequences can establish custody under § 2254, the First Circuit noted, "There are no magic mirrors."[359] Ultimately, the statutory framework proposed in this Article provides an ordinary mirror that reflects not only the realities of contemporary criminal prosecutions but also *Padilla*'s profound effect on how courts will view collateral sanctions. This mirror reflects the deference that federal courts must exercise when reviewing state convictions; yet this mirror concomitantly reflects the necessity for federal habeas review of state convictions. Likewise, this mirror reflects Congress' and the federal courts' collaborative role in defining the jurisdictional limits of § 2254 review. Finally, and perhaps most fundamentally, this mirror reflects with clarity and precision the equally compelling interests of society, the government, and the individual that the "Great Writ" has embodied throughout this nation's history.

359. Lefkowitz v. Fair, 816 F.2d 17, 20 (1st Cir. 1987).

MORAL HAZARD DURING THE SAVINGS AND LOAN CRISIS AND THE FINANCIAL CRISIS OF 2008–09: IMPLICATIONS FOR REFORM AND THE REGULATION OF SYSTEMIC RISK THROUGH DISINCENTIVE STRUCTURES TO MANAGE FIRM SIZE AND INTERCONNECTEDNESS

*ELISA S. KAO**

In the aftermath of the most recent financial crisis, policymakers and regulators have faced the formidable challenge of designing regulatory reforms that adequately address the problems of moral hazard and "too big to fail" in banking. While these concepts have remained persistent stumbling blocks in bank regulatory policy, the modern landscape of the financial services industry has introduced new challenges in the areas of systemic risk and the concentrations of financial power held by a few dominant firms.

This Note compares and contrasts the causes of and responses to the subprime mortgage crisis with those of the Savings and Loan Crisis of the 1980s. It concludes that the provisions under the Dodd-Frank Wall Street Reform and Consumer Protection Act intended to address "too big to fail" actually represent the same type of traditional responses to banking crises as seen before. By subjecting systemically significant firms to more stringent operating requirements and closer monitoring, current reforms merely treat complex financial conglomerates as traditional depository institutions whose incentives can be adequately managed through prudential regulation customarily used to counteract the moral hazard of deposit insurance. To deal with the moral hazard of systemic risk, however, reform efforts should place greater emphasis on anticipatory regulation in a framework of stronger disincentives to discourage growth towards "too big to fail" status.

This Note also supports the recommendation of commentators who advocate for an industry-funded emergency liquidity pool that would incorporate actuarially fair premiums to price the cost of systemic risk, to be paid by the largest and most complex institutions themselves. This represents a combination of ex ante and ex post approaches to inhibit systemic risk at its source as well as manage the consequences of a future systemic crisis, the

* Development Editor, *New York University Annual Survey of American Law*, 2010-2011; J.D., New York University School of Law, 2011; B.A., University of Washington, 2006. I would like to thank the staff members of the *Annual Survey*.

costs of which would otherwise be borne again by the public and the economy as a whole.

INTRODUCTION

"Regulation is, however, an evolving art, change is a necessary concomitant of progress, and changed circumstances require a receptivity to revised approaches."[1]

In the wake of a catastrophic financial meltdown and the subsequent passage of a nearly one-thousand-page-long regulatory reform bill by Congress,[2] federal banking agencies are now tasked with implementing comprehensive rules and regulations that address both the causes and consequences of the subprime mortgage crisis—a sizeable undertaking considering the global scale and systemic impact of the latest asset bubble. "Too big to fail" ("TBTF"), moral hazard, and systemic risk are challenges at the forefront of current reform efforts, yet none are particularly novel in the context of government bailouts that leave taxpayers footing the bill. While the financial sector has undergone rapid changes, many of the same themes of past crises remain, evolving with the ingenuity and innovation of Wall Street. With technological advances in the financial markets occurring at a feverish pace while government oversight lags, fashioning effective reforms for old problems in new settings and applying them to an increasingly complex financial services industry is immensely challenging.

This Note will examine the problem of moral hazard and its manifestation during the subprime mortgage crisis ("Crisis of 2008–09") as compared to that of the Savings and Loan ("S&L") Crisis of the 1980s. It argues that, with regard to systemic risk-related reforms, the Dodd-Frank Wall Street Reform and Consumer Protection Act ("Dodd-Frank Act", or "Dodd-Frank") is an inadequately tailored response to the TBTF problems of the Crisis of 2008–09. Despite new regulatory structures and stricter prudential requirements exacted upon large and complex financial institutions ("LCFIs"), Dodd-Frank accomplishes little in the way of shifting the tremendous costs of future systemic risk events to the firms posing such risks.

1. In re Rate Redesign for Electric Corps., 15 P.U.R.4th 434, 451 (1976) (N.Y. Pub. Serv. Comm'n Aug. 10, 1976) (inquiring into the merits of revising electric rate schedules according to marginal cost).

2. Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, 124 Stat. 1376, 1376 (2010) [hereinafter Dodd-Frank Act] (An Act "[t]o promote the financial stability of the United States by improving accountability and transparency in the financial system, to end 'too big to fail,' to protect the American taxpayer by ending bailouts, to protect consumers from abusive financial services practices, and for other purposes").

While both the S&L and the subprime-mortgage crises involved regulatory lapses and costly miscalculations by public and private sector actors, critical differences between the crises underscore the need for specific reforms addressing the type of moral hazard that arises from systemically significant LCFIs given the costly aftermath of TBTF policies in a globally integrated financial marketplace. Instead of addressing this sort of moral hazard, however, the Dodd-Frank Act's provisions concerning TBTF and regulatory failures resemble the traditional responses of post-crisis regulatory reform and mirror the legislative response to the S&L Crisis. That response may have been suitable at the time for a crisis resulting from lax regulatory oversight and the moral hazard of deposit insurance, but Dodd-Frank's establishment of additional oversight mechanisms and prudential requirements for LCFIs merely papers over the type of moral hazard currently at issue. As a result, it fails to address the true sources of systemic risk: the incentives of LCFIs, both depository and non-depository, to grow TBTF without internalizing the costs of the risks they pose.

Part I of this Note provides a broad overview of the concepts of moral hazard, "too big to fail," and systemic risk. Part II examines the background and regulatory environments preceding and in reaction to both the S&L Crisis and the Crisis of 2008–09 and uses the comparison to highlight the evolution of risk-taking in the expanding financial sector. This comparison is made in the context of two contrasting forms of moral hazard: the explicit form, as demonstrated by the longstanding government guarantee of federal deposit insurance, and the implicit form, as demonstrated by the unspoken guarantee of a government safety net benefiting systemically important institutions. Part III argues that, unlike the explicit moral hazard problem posed by government-insured depository institutions, the moral hazard arising from the implicit guarantee enjoyed by both depository and non-depository TBTF institutions cannot be adequately addressed by increased prudential regulation alone. Part III then concludes that the characteristics of the moral hazard problem during the Crisis of 2008–09, as distinguished from those observed in past crises, provide further justification for stronger disincentives to curb firm growth and manage interconnectedness—methods previously suggested by other commentators, but rejected by Congress. Congress's specific responses to the TBTF problem are unlikely to have a major impact on reducing the threat to financial stability during times of systemic crisis despite the enactment of a host of wide-ranging reforms, including the elimination of a captured agency, establishment of a new systemic risk

regulator, and the reduction of gaps in oversight over non-depository financial institutions. Because these reforms will have a negligible impact on financial firms' actual costs of harboring systemic risk, such measures do little to counteract the current incentive structure that merely encourages LCFIs to grow and consolidate toward TBTF status.

I.
MORAL HAZARD, TOO BIG TO FAIL, AND SYSTEMIC RISK

A. Moral Hazard

"Moral hazard" refers to the problem that arises when a party is incentivized to engage in excessive risk-taking because it is not required to bear the full cost of its potential losses.[3] As a result of this insulation from the risk of loss, an actor will engage in riskier behavior in an effort to enhance its economic returns.[4]

Two distinct types of moral hazard have plagued bank regulatory policy. The first originates from the implementation of government-backed deposit insurance following the Great Depression.[5] Banks receive consumer deposits with the knowledge that the funds are insured by the Federal Deposit Insurance Corporation ("FDIC") up to certain levels, leading insured institutions to take on greater risks in their lending and investment activities than they would without the government guarantee.[6] This moral hazard problem is thus explicit: bankers, investors, and consumers know that deposits held at FDIC-insured institutions are guaranteed up to $250,000 per deposit account, as mandated by statute.[7] The total amount of FDIC insurance coverage is capped by a specific dollar amount in accordance with such limits, and the insurance fund is managed with such limits in mind.[8]

3. *See, e.g.*, Karl S. Okamoto, *After the Bailout: Regulating Systemic Moral Hazard*, 57 UCLA L. REV. 183, 204–05 (2009).

4. *See id.*

5. The phenomenon of moral hazard, and its reduction, has been a focal point of commentary on bank regulatory policy since the implementation of government-backed deposit insurance in 1933. *See* RICHARD SCOTT CARNELL, JONATHAN R. MACEY & GEOFFREY P. MILLER, THE LAW OF BANKING AND FINANCIAL INSTITUTIONS 17–19, 328–30 (4th ed. 2009).

6. *Id.* at 326–33.

7. 12 U.S.C. § 1821(a)(1)(E) (2010).

8. However, the government has the ability to increase deposit insurance coverage and increase its exposure to the risk of loss. In 2008, the deposit insurance limit was temporarily increased from $100,000 to $250,000 per deposit account, and the Dodd-Frank Act made the new insurance limit permanent. *See* Press Re-

Despite its explicit nature, this moral hazard is tolerated as a lesser evil as compared to the liquidity crises that would result from bank runs.[9] Moreover, it is mitigated by a comprehensive system of prudential regulation and government oversight that better aligns ownership and management interests with regulators and the insurance fund.[10] FDIC-insured institutions are subject both to increased scrutiny via frequent regulatory examinations, or "safety and soundness exams,"[11] and to a wide assortment of prudential operating requirements, including higher capital adequacy requirements than other industrial firms.[12] The susceptibility of the banking industry to widespread panics that jeopardize the overall stability of financial markets remains a key rationalization for government-backed deposit insurance and greater regulatory scrutiny.[13]

In contrast to the explicit moral hazard that arises when depository institutions are funded by government-backed deposits, the moral hazard that arises from TBTF policies is implicit. Implicit moral hazard is the result of generally unspoken but established market expectations of government intervention during a systemic financial crisis.[14] Instead of making express promises of a safety net, the government will consistently deny any policy to rescue troubled

lease, FDIC, Basic FDIC Insurance Coverage Permanently Increased to $250,000 Per Depositor (Jul. 21, 2010), http://www.fdic.gov/news/news/press/2010/pr10161.html.

9. *See, e.g.*, Richard Scott Carnell, *A Partial Antidote to Perverse Incentives: The FDIC Improvement Act of 1991*, 12 ANN. REV. BANKING L. 317, 319–21 (1993).

10. *See* Vincent P. Polizatto, *Prudential Regulation and Banking Supervision: Building an Institutional Framework for Banks* 2 n.4 (World Bank, Working Paper Series No. 340, 1990) ("Prudential regulation refers to the set of laws, rules, and regulations which is designed to minimize the risks banks assume and to ensure the safety and soundness of both individual institutions and the system as a whole. Examples include lending limits, minimum capital adequacy guidelines, liquidity ratios, etc.").

11. CARNELL ET AL., *supra* note 5, at 291–92. R

12. *See id.* at 252–53.

13. The domino effects of liquidity crises among depository institutions that result from mass deposit withdrawals by panicked consumers reflect the traditional conception of the consequences of systemic risk in banking. *See* Walker F. Todd & James B. Thompson, *An Insider's View of the Political Economy of the Too Big to Fail Doctrine* 11–13 (Federal Reserve Bank of Cleveland, Working Paper No. 9017, 1990).

14. *See* Marcelo Dabós, *Too Big to Fail in the Banking Industry: A Survey*, *in* TOO BIG TO FAIL: POLICIES AND PRACTICES IN GOVERNMENT BAILOUTS 141, 141–42 (Benton E. Gup ed., 2004); *see also* David Moss, *An Ounce of Prevention: The Power of Public Risk Management in Stabilizing the Financial System* 10 (Harvard Bus. Sch., Working Paper No. 09-087, 2009), *available at* http://www.hbs.edu/research/pdf/09-087.pdf.

firms[15] and will thus operate under a policy of "constructive ambiguity."[16] A notable characteristic of TBTF policies is the refusal by authorities to publicly confirm the existence of any de facto government guarantee that would prevent the insolvencies of systemically significant LCFIs.[17] As a result, not only is the existence of a TBTF guarantee under a veil of constructive ambiguity, but the amount of public funding that would be made available to private firms to prevent a systemic collapse also remains unknown.[18] As this Note will later argue,[19] reducing the TBTF problem and improving the regulation of systemically important institutions requires combating the implicit moral hazard at its source by making the costs of imposing systemic risk explicit.

B. "Too Big to Fail"

"Too big to fail" describes a government's policy of awarding discretionary support to a firm's uninsured creditors out of concern that allowing the firm's failure would have a disastrous impact on the financial system as a whole.[20] As such, TBTF firms are more likely to take excessive risks due to their confidence of government intervention in the event of near-insolvency.[21]

15. Benton E. Gup, *What Does Too Big to Fail Mean?*, *in* TOO BIG TO FAIL: POLICIES AND PRACTICES IN GOVERNMENT BAILOUTS, *supra* note 14, at 29, 30 (on governments' denials of the existence of TBTF policies). R

16. *See* Dabós, *supra* note 14, at 141–42. R

17. "Constructive ambiguity" refers to the "policy of using ambiguous statements to signal intent while retaining policy flexibility." James B. Thomson, *On Systemically Important Institutions and Progress Systemic Mitigation* 8–9 (Fed. Reserve Bank of Cleveland, Policy Discussion Paper No. 27, 2009). In the bank regulatory context, the phrase refers to the practice of limiting public knowledge of which firms the government considers TBTF, in order to create uncertainty over the availability of a government safety net and improve market discipline. *Id.* In contrast, a policy of supervisory transparency would involve public disclosure of the government's list of systemically important financial firms. *Id.* at 9.

18. *See* Moss, *supra* note 14, at 10 (describing implicit federal guarantees of systemically significant financial institutions as open-ended); *see also id.* at 12 (contrasting implicit guarantees with explicit guarantees that are clear, well-defined, and delimited). R

19. *See infra* Part III.C.

20. GARY H. STERN & RON J. FELDMAN, TOO BIG TO FAIL: THE HAZARDS OF BANK BAILOUTS 1 (2004).

21. *Id.* at 2.

> To the extent that creditors of TBTF banks expect government protection, they reduce their vigilance in monitoring and responding to these banks' activities. When creditors exert less of this type of market discipline, the banks may take excessive risks. TBTF banks will make loans and other bets that seem

During the recent subprime mortgage crisis, seventeen financial conglomerates accounted for at least half of the $1.1 trillion in global losses cited by the world's financial institutions.[22] In response, central banks and governments in the United States, the United Kingdom, and Europe provided nearly nine trillion dollars of support in various forms, including emergency liquidity pools, capital injections, asset purchase programs, and financial guarantees, all in an effort to save the global markets from systemic collapse.[23] The challenges presented by the insolvencies of firms such as AIG, Lehman Brothers, and Bear Stearns have underscored the need to address the risk posed by entities that are TBTF.

The risk of moral hazard increases when governments consistently intervene to support distressed financial institutions, thus solidifying expectations of such intervention in times of financial upheaval.[24] As a result of the incentive structure that these expectations establish, financial firms are encouraged to grow and combine such that they may take advantage of the lower cost of capital of firms reaching TBTF status.[25] This lower cost of capital reflects the investor expectation that the government would never allow such entities to fail in the event of near-insolvency.[26] When a sufficient number of firms receive discretionary support, the market internalizes an expectation of future government-orchestrated rescues, resulting in excessive risk-taking and economic waste.[27] This rational

quite foolish in retrospect. . . . This undesirable behavior is frequently referred to as the 'moral hazard' of TBTF protection. . . .

22. *See* Arthur E. Wilmarth, Jr., *The Dark Side of Universal Banking: Financial Conglomerates and the Origins of the Subprime Financial Crisis*, 41 CONN. L. REV. 963, 968 (2009) [hereinafter Wilmarth, *Universal Banking*].

23. *See id.*

24. Gup, *supra* note 15, at 43. R

25. *See* Arthur E. Wilmarth, Jr., *The Transformation of the U.S. Financial Services Industry, 1975–2000: Competition, Consolidation, and Increased Risks*, 2002 U. ILL. L. REV. 215, 215 (2002) [hereinafter Wilmarth, *U.S. Financial Services*].

26. The cost of funding for government-sponsored entities and mortgage giants Fannie Mae and Freddie Mac serves as an example of the lower costs of debt with an implicit government guarantee. Viral V. Acharya et al., *The Government-Sponsored Enterprises, in* REGULATING WALL STREET: THE DODD-FRANK ACT AND THE NEW ARCHITECTURE OF GLOBAL FINANCE 429, 434–35 (Viral V. Acharya et al. eds., 2011) ("The liabilities of Fannie Mae and Freddie Mac also give some idea of the importance of this implicit government guarantee. The GSE debt is typically issued at interest rates that are somewhere between AAA-rated corporate and U.S. Treasury obligations . . . .").

27. *See* STERN & FELDMAN, *supra* note 20, at 23 ("Expectations of TBTF coverage are costly because they lead to a wasting of resources and a reduction in the welfare of the citizenry."); *see also id.* at 24 ("The costs of lost output can dwarf the transfers from financial losses. While the fiscal flows of the savings and loan

calculation by firms weakens managerial incentives to act prudently and cabin institution-specific risks, resulting in greater leverage and riskier investment activities to enhance profitability.[28]

While the term TBTF is not limited to the banking industry, the origins of the term's popularity originated in congressional hearings in 1984 following the federal government's bailout of Continental Illinois Bank,[29] when regulators feared the bank's failure would lead to a systemic financial crisis.[30] Since the Continental Illinois failure, the term "too big to fail" has been used widely throughout banking literature and is most commonly associated with the government's assistance to banks that are actually "too big to liquidate," as opposed to TBTF.[31] Indeed, size cannot be the only reason for TBTF because the concept is more directly related to the risk of contagion.[32] Thus, the phrase is somewhat misleading to the extent it also describes firms that are not necessarily big, but may still receive discretionary support due to their "interconnectedness" with the rest of the market. While interconnected institutions are often significant in terms of total asset size, a large financial institution that is relatively disconnected from the financial markets in terms of transactions with counterparties does not present the same TBTF concerns as a smaller institution with a high concentration in particular investment areas giving rise to more counterparty relationships.[33] As the recent crisis demonstrated, "too interconnected to fail" has become a functional equivalent of TBTF, widening its ap-

bailout . . . equaled $150 billion, lost output . . . largely attributed to moral hazard and poor resource allocation—was on the order of $500 billion.").

28. *See id.* at 23.

29. Gup, *supra* note 15, at 30: R

> Bank regulators feared that Continental's problems might spread to more than 1,000 other banks that had deposits and/or federal funds there and they too might fail if Continental failed. Accordingly, Comptroller of the Currency Todd Conover went before the U.S. Congress in 1984 to declare that Continental and 10 other of the nation's largest banks were 'too big to fail.'

30. *Id.* at 30–31 (quoting Congressman Stewart McKinney as saying in 1984: "Mr. Chairman, let us not bandy words. We have a new kind of bank. It is called too big to fail. TBTF, and it is a wonderful bank").

31. *Id.* at 31.

32. *See* Dabós, *supra* note 14, at 141–43. "Contagion" is demonstrated when a R "cascading series" of bank failures or liquidity events results from the failure of one institution and a sequence of interbank counterparty relationships. Jeffrey N. Gordon & Christopher Muller, *Confronting Financial Crisis: Dodd-Frank's Dangers and the Case for a Systemic Emergency Insurance Fund*, 28 YALE J. REG. 151, 160 (2011).

33. *See* Dabós, supra note 14, at 142–43; *see also* Jean Burson, *A Framework for* R *Systemically Important Institutions*, FOREFRONT, Winter 2009–2010, at 14, 16, *available at* http://www.clevelandfed.org/forefront/2009/12/pdf/ff_winter_2009-2010_00.pdf.

plicability outside the realm of traditional depository institutions and into the "shadow banking system" of nonbank financial firms. As the term TBTF is used in the remainder of this Note, it is also meant to evoke the sense of the term in its more precise form of "too interconnected to fail."[34]

C. Systemic Risk

Systemic risk—also referred to as counterparty risk—arises when one financial institution's actions have the potential to adversely impact the operations or solvency of another institution as a result of transactions the firms have entered into with each other.[35] For example, when Bank A reduces its interbank lending to Bank B, this creates problems for Bank B's ongoing operations. The fact that Bank B's risk of insolvency is adversely impacted by the activities of Bank A, even if Bank B was otherwise prudently managing the asset quality of its own portfolio, highlights the susceptibility of financial institutions involved in a wide array of counterparty transactions to systemic risk.

Economic distress magnifies the problem of systemic risk. Regulators faced with the potential insolvency of Continental Illinois Bank feared that the risk of widespread failures among the bank's counterparties could result in devastation across the financial markets.[36] At the time, John LaWare, a former governor of the Federal Reserve Board, offered the following dramatic description of systemic risk:

> [A] nightmare condition that is unfair to everybody. The only analogy that I can think of for the failure of a major international institution of great size is a meltdown of a nuclear generating plant like Chernobyl. The ramifications of that kind of failure are so broad and happen with such lightning speed that you cannot after the fact control them.[37]

34. Gordon & Muller, *supra* note 32, at 160. ("It was commonly stated that Bear Stearns was not 'too big to fail,' the general moral hazard objection to government rescues, but 'too interconnected to fail.'").

35. Steven L. Schwarcz, *Systemic Risk*, 97 GEO. L.J. 193, 204 (2008). Professor Schwarcz defines systemic risk as

> the risk that (i) an economic shock such as market or institutional triggers (through a panic or otherwise) either (X) the failure of a chain of markets or institutions or (Y) a chain of significant losses to financial institutions, (ii) resulting in increases in the cost of capital or decreases in its availability, often evidenced by substantial financial-market price volatility.

36. Todd & Thompson, *supra* note 13, at 5. R

37. *Economic Implications of the "Too Big to Fail" Policy: Hearing Before the Subcomm. on Econ. Stabilization of the H. Comm. on Banking, Fin. and Urban Affairs*, 102d

LaWare might as well have been discussing the systemic impact of the Lehman Brothers Chapter 11 bankruptcy filing in 2008.[38]

While systemic risk has traditionally been associated with deposit runs resulting from bank panics and cascading bank failures, a similar domino effect occurs in the context of nonbank financial firms when there is a "run" on investor confidence. Where systemic risk gives rise to uncertainty about a firm's ability to meet multiple obligations, counterparties respond by engaging in a mass exodus from existing investment relationships, and creditors respond by freezing existing lines of credit.[39] As with bank runs by consumers in the depository context, a firm's sudden inability to meet obligations leads to widespread liquidity shortages that give rise to panic among financial firms and contagion throughout the industry.[40] Such contagions arose in the 2008 cases of Bear Stearns, Lehman Brothers, and AIG, each of which acted as a major intermediary in the unregulated over-the-counter ("OTC") derivatives market.[41] Further, each firm was heavily concentrated in collateralized debt obligations and other structured finance products and had balance sheet structures largely supported by short-term liabilities.[42] Consequently, instead of bank panics leading to insufficient liquidity to fund consumer deposit withdrawals, these nonbank institutions experienced liquidity shortages when suddenly faced with the collapse of a certain investment type—e.g., assets backed by overvalued subprime mortgages. Investors reacted by pulling out, while creditors quickly cut off access to short-term funding sources by freezing credit lines.[43]

Systemic risk presents a host of regulatory challenges, particularly with respect to nonbank financial institutions. Deposit insurance, while a notable source of moral hazard in the traditional banking context, is meant to diminish the occurrence of bank pan-

Cong. 34 (1991) (statement of John LaWare, Governor, Board of Governors of the Federal Reserve System).

38. *See* ANDREW ROSS SORKIN, TOO BIG TO FAIL 351–61 (2009); *see also* Okamoto, *supra* note 3, at 196. R

39. *See* Okamoto, *supra* note 3, at 196–98 (describing Bear Stearns, Lehman Brothers, and AIG in the context of the credit default swaps market). R

40. *See id.* at 200.

41. *Id.* at 198. Over-the-counter derivatives are privately traded agreements that allow parties to negotiate contract terms according to their specific needs, in contrast to exchange-traded derivatives that have standardized terms and are traded through organized exchanges. ALAN N. RECHTSCHAFFEN, CAPITAL MARKETS, DERIVATIVES AND THE LAW 162–63 (2009).

42. *See* Okamoto, *supra* note 3, at 200–03. R

43. *See id.* at 196–98, 203.

ics and thus the risk of contagion and liquidity shortages that disrupt the stability of the banking system.[44] However, no such explicit guarantee of emergency funding exists for nonbank financial firms to prevent runs on investor or creditor confidence. Further, existing safety and soundness regulations apply only to depository institutions. Non-depository nonbank financial firms, financial holding companies, and their nonbank subsidiaries are not subject to the same stringent operating requirements, monitoring by regulators, and intrusive examination processes imposed on traditional banks. While the lines have blurred among the various types of financial firms,[45] the continued growth and consolidation of these firms has also produced larger and more complex financial conglomerates.[46] Accordingly, regulatory reforms with regard to systemic risk must be comprehensive in addressing the problems of moral hazard and TBTF in the context of both depository and nondepository financial institutions.

II.
COMPARING THE DEREGULATORY ENVIRONMENTS OF THE SAVINGS AND LOAN CRISIS AND THE SUBPRIME MORTGAGE CRISIS

While a great deal can be said about the many causes of the S&L Crisis[47] and the Crisis of 2008–09,[48] this Part focuses specifically on the deregulatory environments preceding each crisis and the manifestations of moral hazard in each case. Part A discusses the causes of the S&L Crisis and highlights the role of deposit insurance as a source of an explicit moral hazard problem. This, in combination with a deregulatory environment that expanded permissible activities for thrifts,[49] dramatically altered incentives

44. CARNELL ET AL., *supra* note 5, at 309–10. R

45. *See* Wilmarth, *Universal Banking*, *supra* note 22, at 975–981.

46. *See* Wilmarth, *U.S. Financial Services*, *supra* note 25, at 251–54.

47. *See, e.g.*, *Administration's Plan to Resolve the Savings & Loan Crisis: Hearing Before the H. Comm. on Banking, Fin. and Urban Affairs*, 101st Cong., 71 (1989) (statement of Nicholas F. Brady, Sec'y of the Dep't of the Treasury).

48. *See* Ben S. Bernanke, *Opening Remarks*, 2008 FED. RES. BANK KANSAS CITY ECON. POL'Y SYMP. 1–3 (2009) [hereinafter Bernanke, *Opening Remarks*], *available at* http://www.kc.frb.org/publicat/sympos/2008/Bernanke.03.12.09.pdf; *see also* Ben S. Bernanke, Chairman, Fed. Reserve Bd. of Governors, Address at the Council on Foreign Relations (Mar. 10, 2009) [hereinafter Bernanke, Address at the Council on Foreign Relations], *available at* http://www.federalreserve.gov/newsevents/speech/bernanke20090310a.htm.

49. Thrift institutions, or "thrifts," are financial institutions that "primarily accept[ ] savings account deposits and invest[ ] most of the proceeds in mortgages."

and thus elevated the risk profiles of these institutions without a corresponding increase in regulatory supervision or prudential requirements. Part A continues by contrasting the deregulatory environment of the S&L Crisis with that of the Crisis of 2008–09, drawing distinctions between the regulatory failures at issue and giving particular attention to the contrast between the regulatory forbearance that characterized the 1980s and the regulatory arbitrage preceding the recent crisis. Legal developments in the 1990s paved the way for the "shadow banking system" and the regulatory arbitrage strategies that led to the buildup of systemic risk and growing interconnectedness among major financial firms.[50] Part B describes the post-crisis responses to the explicit and implicit moral hazard problems at issue, as set against the overarching themes of financial liberalization and shortsighted regulatory design. Part B then emphasizes how a thinly veiled TBTF safety net is particularly problematic for attempts to reform the regulatory framework.

A. Mismanaging Deregulation

1. S&L Crisis: A proactive Congress and regulatory forbearance

Deregulatory measures. Discussion of the myriad causes of the S&L Crisis generally includes the deregulatory environment established by Congress during the 1980s. In hindsight, the legislative attempts to improve competition among commercial banks and thrift institutions preceding the crisis involved an expeditious series of miscalculations—a failure of public policy that emerged from a blind commitment to the principle that competition and market discipline would always prevail.[51] During the 1980s, a troubled S&L industry faced substantial interest-rate mismatches on its balance sheets and growing insolvencies resulting from a business model that suffered when thrifts making long-term, fixed rate consumer

A savings and loan association is an example of a thrift. *Definition of BHCs and Banking Terms*, FEDERAL FINANCIAL INSTITUTIONS EXAMINATION COUNCIL, http://www.ffiec.gov/nicpubweb/Content/HELP/Institution%20Type%20Description.htm (last visited Feb. 28, 2012).

50. Patricia A. McCoy, Andrey D. Pavlov & Susan M. Wachter, *Systemic Risk Through Securitization: The Result of Deregulation and Regulatory Failure*, 41 CONN. L. REV. 1327, 1329 (2009).

51. *See* FEDERAL DEPOSIT INSURANCE CORPORATION, 1 HISTORY OF THE EIGHTIES—LESSONS FOR THE FUTURE 172, 187, *available at* http://www.fdic.gov/bank/historical/history/167_188.pdf (describing the government's response to the early crisis as "a patchwork of misguided policies that set the stage for massive taxpayer losses to come").

loans were funded primarily through short-term deposits.[52] Amid a rising interest-rate environment, the cost of funding—i.e., interest expenses—swiftly outpaced interest income earned on fixed-rate mortgages. Faced with regulatory restrictions governing the asset-liability structures of thrift institutions, including limitations on allowable investments and maximum ceilings on deposit rates, many S&Ls became technically insolvent.[53]

Congress also enacted two principal deregulatory initiatives in this period that would later exacerbate the severity and cost of the S&L Crisis. The Garn-St. Germain Depository Institutions Act of 1982 (Garn-St. Germain Act)[54] expanded permissible investments for S&Ls and has been blamed for promoting a policy of permitting excessive risk-taking without subjecting risk-takers to the cost of such risk, thus encouraging capital forbearance.[55] Additionally, the Depository Institutions Deregulation and Monetary Control Act of 1980 increased deposit insurance levels from $40,000 per account to $100,000, which dramatically increased the total cost of the S&L Crisis to the public.[56]

The Garn-St. Germain Act deregulated the S&L industry by eliminating deposit rate ceilings and loosening restrictions on allowable business activities for S&Ls.[57] After deposit interest rate ceilings were removed to alleviate problems of disintermediation from deposit accounts to money market mutual funds, an influx of

52. JEFFREY CARMICHAEL & MICHAEL POMERLEANO, WORLD BANK, THE DEVELOPMENT AND REGULATION OF NON-BANK FINANCIAL INSTITUTIONS 186 (2002).

53. *Id.*

54. *See* Garn-St. Germain Depository Institutions Act of 1982, Pub. L. No. 97-320, 96 Stat. 1469, 1469 (1982) ("An Act to revitalize the housing industry by strengthening the financial stability of home mortgage lending institutions and ensuring the availability of home mortgage loans.").

55. *See* Franklin E. Zimring & Gordon Hawkins, *Crime, Justice, and the Savings and Loan Crisis*, 18 CRIME & JUST. 247, 266 (1993).

56. FEDERAL DEPOSIT INSURANCE CORPORATION, *supra* note 51, at 176. The staggering increase in coverage was made after the Senate had initially approved a new limit of $50,000 per qualified deposit account. While the House first left the figure alone, a final effort to change the insurance limit to $100,000 per account gained rapid industry support as rules regarding deposit interest rates were also loosened—a combination of measures that would dramatically increase the eventual cost of the financial crisis to follow. *See* CARNELL ET. AL, *supra* note 5, at 316; *see also* Christine M. Bradley, *A Historical Perspective on Deposit Insurance Coverage*, 13 FDIC BANKING REV., no. 2, 2000, at 1, 17, *available at* http://www.fdic.gov/bank/analytical/banking/2000dec/brv13n2_1.pdf (citing congressional records indicating Congress's reasoning that "[a]n increase from $40,000 to $100,000 will not only meet inflationary needs but lend a hand in stabilizing deposit flows among depository institutions and noninsured intermediaries"). R

57. CARMICHAEL & POMERLEANO, *supra* note 52.

insured consumer deposits flowed into thrifts.[58] This influx escalated thrifts' perverse incentives to engage in excessive risk-taking and exacerbated moral hazard. As a result of Congress' effort to help thrifts to "grow out" of their problems via deregulation, S&Ls broadened their activities into areas outside their traditional realms of home mortgages and consumer lending.[59] This led to greater profitability, but it also significantly increased the overall risk profile of the S&L industry.[60] To compete with commercial banks, S&Ls engaged in riskier acquisition, development and construction loans, and other unfamiliar investment areas, even when they were technically insolvent.[61] The industry also began littering balance sheets with low-grade assets such that capital levels were not commensurate with the risk profiles of the institutions.[62] Significantly, the Garn-St. Germain Act also reduced required regulatory capital levels, thereby allowing S&Ls to shrink their equity cushions while simultaneously expanding on the risks they could take.[63]

In addition to the general undercapitalization of S&Ls and the lack of sufficient capital adequacy regulations,[64] thrifts were plagued by many other factors that would prove to be problematic. These factors included: intense competition among thrifts and banks for local deposits, coupled with rising interest rates that in-

58. *See* FEDERAL DEPOSIT INSURANCE CORPORATION, *supra* note 51, at 176. Disintermediation is the outflow of deposits from financial institutions into other investments offering higher interest rates. *Id.*

59. *See* Karen Harris, Note, *Anticipatory Regulation for the Management of Banking Crises*, 38 COLUM. J.L. & SOC. PROBS. 251, 261–62 (2005); *see also* FEDERAL DEPOSIT INSURANCE CORPORATION, *supra* note 51, at 179.

> Another major change resulting from deregulation was that, beginning in 1982, S&L investment portfolios rapidly shifted away from traditional home mortgage financing into new activities. This shift was made possible by the influx of deposits and also by sales of existing mortgage loans. By 1986, only 56 percent of total assets at savings and loan associations were in mortgage loans, compared with 78 percent in 1981 . . . .

60. *See* FEDERAL DEPOSIT INSURANCE CORPORATION, *supra* note 51, at 180. Between 1982 and 1985, total thrift assets invested in commercial mortgages and land loans increased from 7.4% to 12.1%, a total increase of $78.6 billion. *Id.* at 184.

61. *See* FEDERAL DEPOSIT INSURANCE CORPORATION, *supra* note 51, at 176, 181; *see also* Harris, *supra* note 59.

62. CARMICHAEL & POMERLEANO, *supra* note 52.

63. *See* CONG. BUDGET OFFICE, J932-41, THE COST OF FORBEARANCE DURING THE THRIFT CRISIS 2 (1991).

64. NORMAN STRUCK & FRED CASE, WHERE DEREGULATION WENT WRONG: A LOOK AT THE CAUSES BEHIND SAVINGS AND LOAN FAILURES IN THE 1980S 14–16 (1988), *reprinted in* Arthur W. Leibold, Jr., *15 Major Causes of Losses that Hurt the Savings and Loan Business in the 1980s*, *in* THE SAVINGS AND LOAN CRISIS: LESSONS FROM A REGULATORY FAILURE 58 (James R. Barth et al. eds., 2004).

creased the cost of funds; elimination of rules limiting both direct lending and loan participations with other banks in non-local markets; overall management inexperience and incompetence, exacerbated by a lack of board oversight; and, in many instances, outright fraud and insider abuse.[65] An industry once defined by the conservative thirty-year fixed home mortgage and similar garden-variety consumer loans became characterized by a culture of "high rolling" and speculative ventures.[66]

These contributing factors—the sudden expansion of permissible investments by S&Ls, a dramatic increase in the amount of government-backed liabilities, capital forbearance, and management failures—combined with significant risk-taking in new business lines by inexperienced thrift managers, undermined the S&L industry.[67] Dramatically increasing deposit insurance limits while simultaneously reducing net worth requirements worsened the moral hazard problem of funding riskier activities with insured funds without a corresponding increase in supervision or deposit insurance reforms to combat the perverse incentives.[68]

Regulatory forbearance. These initiatives by Congress were consistent with most other political, legislative, and regulatory decisions made in the spirit of deregulation during the early 1980s.[69] In fact, the government's initial reactions to the burgeoning crisis allowed the negative effects of deregulation to snowball in an environment already conducive to policies of regulatory forbearance.[70] Those charged with supervising S&Ls either failed to act, acted too slowly,

65. *Id.*

66. *See generally* MARTIN E. LOWY, HIGH ROLLERS: INSIDE THE SAVINGS AND LOAN DEBACLE (1991) (documenting the events leading up to, and the personalities behind, the 1980s S&L crisis).

67. *See* FEDERAL DEPOSIT INSURANCE CORPORATION, *supra* note 51, at 180 ("[H]igh-risk development loans and the resultant mortgages on the same properties were most likely the principal cause for thrift failures after 1982."). Unfamiliar business lines included real estate, equity securities, casinos, fast-food franchises, ski resorts, and windmill farms, while new securities investments included junk bonds, arbitrage schemes, and derivative instruments. *See id.*

68. *Id.* at 175.

69. *Id.* at 177.

70. CONG. BUDGET OFFICE, J932-41, THE COST OF FORBEARANCE DURING THE THRIFT CRISIS 2 (1991).

> Forbearance is the discretionary practice of not enforcing an existing rule. In the 1980s, thrift regulators elevated forbearance to a general policy for the entire thrift industry—they did not close institutions when they became insolvent. Regulators did not violate statutes; rather, in altering agency regulations they interpreted those statutes in the most liberal way possible, thereby allowing themselves to avoid closing insolvent institutions.

or took less-than-meaningful enforcement action when confronted with insolvent institutions.[71] While deposit insurance creates moral hazard, the problem can be mitigated by prudent government regulation.[72] When, however, regulatory forbearance is consistently exercised as a matter of policy, as in the 1980s, managers face few roadblocks to deter them from imprudent banking practices, including risky lending, speculative investing, and forays into high-risk ventures without adequate experience, monitoring, and capital levels to support such activities.

A lax regulatory environment was also conducive to widespread fraud and insider abuse, as S&L managers were incentivized to engage in imprudent, often reckless, and even criminal business practices.[73] Some commentators, noting the overall prevalence of criminal activity by insiders during the S&L crisis, have argued that purposeful fraud was one of the primary causes of the losses incurred during the crisis.[74] While its significance as a contributing factor to the collapse is debated, the fact that fraudulent activity on the part of insiders occurred in an estimated 70% of failed S&L associations[75] has led some to describe the 1980s as "a decade of commercial lawlessness,"[76] and the crisis as a "theft from the taxpayer" that resulted in "the worst public scandal in American history."[77]

In sum, deregulatory measures resulting from misguided congressional action—increased deposit insurance, elimination of deposit interest-rate controls, and permissive operating requirements—combined with policies of forbearance due to regulatory inaction, resulted in the thrift industry's collapse. The significant expansion of government guarantees through deposit insurance worsened moral hazard and increased the cost of rescuing insolvent institutions.[78] In addition, the benefits of deregula-

71. *See id.*

72. *See* CARNELL ET AL., *supra* note 5, at 329 ("The FDIC also expects those banks' regulators to impose controls stringent and pervasive enough to constrain moral hazard."); *see also* Moss, *supra* note 14, at 4. R R

73. *See* STRUNK & CASE, *supra* note 64, at 15 (suggesting that fraud and insider abuse caused approximately 20% of failures between 1985 and 1988).

74. Zimring & Hawkins, *supra* note 55, at 264.

75. Harris, *supra* note 59, at 266.

76. Zimring & Hawkins, *supra* note 55, at 265 (quoting Michael M. Thomas's description of the crisis as a "mosaic of disaster . . . complex in the extreme, mixing simple thuggery with subtle feats of financial and legal prestidigitation").

77. *Id.* at 265.

78. *Id.* at 267; FEDERAL DEPOSIT INSURANCE CORPORATION, *supra* note 51, at 176; *see also Administration's Plan to Resolve the Savings and & Loan Crisis: Hearing Before the H. Comm. on Banking, Fin. and Urban Affairs*, 101st Cong., 55 (1989) (state-

tion—healthy competition, improved profitability, and economic growth—were overshadowed by the costs of excessive risk-taking and the resulting insolvencies.[79] By ignoring the complexities of the dynamics among competing financial institutions, a pertinent factor in regulatory design, Congress failed to implement a proportionate increase in regulatory scrutiny to manage additional risks as they materialized. Where no corresponding expansion of prudential regulation was implemented to counteract the moral hazard effects of an expanded government safety net, such as more stringent regulatory capital or asset quality requirements, financial liberalization and unrestrained risk-taking paved the way for financial crisis. Regulatory forbearance and worsening moral hazard problems further distorted the marketplace such that the celebrated benefits of free markets were lost almost as soon as they were sought.[80] Unsurprisingly, moral hazard was quickly identified as a key cause of the excessive risk-taking by thrift management during the S&L Crisis.[81]

2. Crisis of 2008–09: Congressional inaction and regulatory arbitrage

The legislative actions associated with the expansion of the thrift industry in the 1980s exhibited a free market fervor that ultimately increased societal costs, and a similar philosophy reappeared in the years preceding the Crisis of 2008–09. The 1990s saw a combination of legislative responses to the S&L Crisis and the expansion of permissible activities of financial institutions and their holding companies, representing a deregulatory environment of a markedly different nature. Opportunities for regulatory arbitrage emerged from the formation of regulatory gaps and the elimination

ment of Nicholas F. Brady, Sec'y of the Dep't of the Treasury) ("So the main problem here, you put your finger on, but it is worth repeating: it is idiocy to allow institutions to go out and get Federally insured deposits and let them do whatever they want once they have got that insurance. That's what we are trying to stop.").

79. Arthur E. Wilmarth, Jr., *Does Financial Liberalization Increase the Likelihood of a Systemic Banking Crisis? Evidence from the Past Three Decades and the Great Depression*, *in* TOO BIG TO FAIL: POLICIES AND PRACTICES IN GOVERNMENT BAILOUTS, *supra* note 14, at 77, 80–82 [hereinafter Wilmarth, *Financial Liberalization*]. R

80. *See* FEDERAL DEPOSIT INSURANCE CORPORATION, *supra* note 51, at 181 ("Although in a free-market economy competition is normally considered healthy, regulatory forbearance in the thrift industry and moral hazard created marketplace distortions that penalized well-run financial institutions.").

81. John C. Coffee, Jr., *What Caused Enron? A Capsule Social and Economic History of the 1990s*, 89 CORNELL L. REV. 269, 278 (2004) ("After the S&L crisis, investigators quickly identified a classic 'moral hazard' problem. Because the government guaranteed banks' financial obligations to depositors, these depositors had little reason to monitor management, and accordingly bank promoters were able to leverage their firms excessively.").

of barriers between banking and nonbanking activities, leaving mounting levels of systemic risk largely unmonitored and wholly unregulated.

Legal developments through the 1990s. Following the S&L Crisis, a series of legislative enactments changed the operating environment of the banking industry. Congress passed the Financial Institutions Reform, Recovery, and Enforcement Act of 1989 ("FIRREA")[82] and provided $50 billion to close the banks that had failed and to prevent additional losses in hopes of restoring public confidence in the thrift industry.[83] Congress also abolished the Federal Savings & Loan Insurance Corporation ("FSLIC") and the deposit insurance fund for thrift institutions, and handed the responsibility of insuring S&L deposits to the FDIC.[84] In addition, FIRREA created two new agencies—the Federal Housing Finance Board and the Office of Thrift Supervision—to replace the Federal Home Loan Bank Board.[85]

In 1991, Congress passed the Federal Deposit Insurance Corporation Improvement Act of 1991 ("FDICIA"),[86] which greatly expanded the FDIC's powers and allowed it to borrow from the Treasury, mandated risk-based deposit insurance assessments for banks, and established new capital requirements and regulatory standards.[87] Under FDICIA's Prompt Corrective Action ("PCA") system, safety and soundness examiners assess each bank's capital adequacy using guidelines that mandate specific restrictions as a bank's regulatory capital ratios dip below certain thresholds: "well-capitalized," "adequately capitalized," "deficient," and "critically deficient."[88] The goal of the PCA regime is to provide a mandatory system for the resolution of failed banks in order to avoid the regulatory forbearance problems that plagued the thrift crisis.[89]

82. Financial Institutions Reform, Recovery, and Enforcement Act of 1989, Pub. L. No. 101-73, 103 Stat. 183 [hereinafter FIRREA].

83. *See* Timothy Curry & Lynn Shibut, *The Cost of the Savings and Loan Crisis: Truth and Consequences*, 13 FDIC BANKING REV., no. 2, 2000, at 26, 28, *available at* http://www.fdic.gov/bank/analytical/banking/2000dec/brv13n2_2.pdf.

84. 12 U.S.C. § 1821a(a)(1) to (2) (2006).

85. FIRREA §§ 307, 308, 401 ("FSLIC and Federal Home Loan Bank Board Abolished").

86. Federal Deposit Insurance Corporation Improvement Act of 1991, Pub. L. No. 102-242, 105 Stat. 2236 [hereinafter FDICIA].

87. *See* CARNELL ET AL., *supra* note 5, at 30; *see also* Christopher J. Pike & James B. Thompson, *FDICIA's Prompt Corrective Action Provisions*, FED. RES. BANK CLEV. 1 (Sept. 1, 1992), www.clevelandfed.org/research/Commentary/1992/0901.pdf. R

88. 12 U.S.C. § 1831o(b)(1) (2006).

89. *See* CARNELL ET AL., *supra* note 5, at 30; *see also* Pike & Thompson, *supra* note 87, at 3 (suggesting that mandating prompt intervention by regulators R

In addition, the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 ended the prohibition of interstate banking, allowing bank holding companies with adequate capital levels to acquire banks across state lines.[90] Most notably, the last major banking law enacted in the 1990s, the Gramm-Leach-Bliley Act of 1999 ("GLB Act"), eliminated the strict separation among investment banks, commercial banks, securities firms, and insurance companies that had existed since the Glass-Steagall Act of 1933 ("Glass-Steagall").[91] The GLB Act established "financial holding companies," or bank holding companies meeting certain criteria that are allowed to engage in a wide variety of business activities.[92] The Act thus permitted the consolidation of different types of financial firms that would otherwise have been precluded by Glass-Steagall, and a new kind of nontraditional financial conglomerate was born.[93] For example, Citicorp, a bank holding company, merged with Travelers, a financial firm that owned insurance subsidiaries, and Salomon Smith Barney, a major securities firm, to form Citigroup—the first "universal bank" in the United States since 1933.[94] In sum, legal developments throughout the 1990s included using billions of government dollars to prevent further losses from the S&L Crisis, restructuring the regulatory framework and creating new banking agencies, reducing regulatory discretion in an effort to prevent regulatory forbearance, and expanding bank powers such that a new era of "nonbank" banking could be born.

Regulatory arbitrage. If the regulatory environment of the S&L industry in the 1980s was characterized by regulatory forbearance, the 1990s and early 2000s were marked by regulatory arbitrage. Such arbitrage occurs when firms take advantage of gaps in regulatory oversight by exploiting business areas not subject to government supervision, taking excessive risks, and becoming highly

reduces the political pressures that otherwise give regulatory agencies perverse incentives to engage in regulatory forbearance).

90. *See* Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994, Pub. L. No. 103-328, 108 Stat. 2338.

91. Gramm-Leach-Bliley Act, Pub. L. No. 106-102, § 101(a), 113 Stat. 1338, 1341 (1999) (repealing Section 20 of the Banking Act of 1933 (12 U.S.C. § 377)).

92. *See* CARNELL ET AL. *supra* note 5, at 465. R

93. *See id.*; *see also* Wilmarth, *Universal Banking*, *supra* note 22, at 972–73.

94. The Federal Reserve Board approved the merger in 1998 via an exemption in the Bank Holding Company Act and allowed Citigroup to provide universal banking services for a period of up to five years. The approval eventually led Congress to repeal Glass-Steagall and amend the Bank Holding Company Act such that the financial conglomerate could exist on a permanent basis; in November of 1999, Congress enacted the Gramm-Leach-Bliley Act. *Id.*

concentrated unregulated products.[95] These "legal or supervisory gaps through which organizations or individuals can act contrary to the purposes of the regulation"[96] can lead to the failure to regulate an area entirely—an arguably worse outcome than the mere laxity in government oversight observed in the 1980s.

Regulatory arbitrage is particularly dangerous for several reasons. First, as the term "shadow banking" might suggest, such strategies occur "in the dark" and fall outside of government scrutiny. Thus, the problems arising from such strategies, or the strategies themselves, are not readily apparent as a source of potential distress until it is too late.[97] Difficulties in monitoring an area of investing also lead to difficulties in measurement—a problem that bore out painfully during the crisis. Second, as Congress's response to the regulatory forbearance of the 1980s demonstrated, government laxity and regulatory capture can be dealt with in a fairly direct manner: by eliminating and rebuilding, or by abolishing ineffective agencies and redistributing responsibilities to new agencies and regulators.[98] In addition, the implementation of PCA guidelines under FDICIA reflects Congress's ability to simply curb the discretion given to regulators in fulfilling their duties.[99] The "prompt res-

95. Bernanke, Address at the Council on Foreign Relations, *supra* note 48.

96. Harris, *supra* note 59, at 254 n.15.

97. *See* Viral V. Acharya et al., *The Dodd-Frank Wall Street Reform and Consumer Protection Act*, *in* REGULATING WALL STREET: THE DODD-FRANK ACT AND THE NEW ARCHITECTURE OF GLOBAL FINANCE, *supra* note 26, at 2–3 (describing the development of a "parallel (shadow) banking system" that reflected regulatory arbitrage opportunities left unmonitored as a result of "regulatory naïveté . . . , the ideology of the times, and a cognitive failure by everyone to appreciate fully the unintended consequences of existing regulation and to develop the tools to deal with them").

98. In passing FIRREA in 1989, Congress eradicated the FSLIC and redistributed responsibilities of insuring and regulating the thrift industry to the FDIC and to two new agencies, respectively. *See* FIRREA, §§ 307, 308. Twenty years later, one of those agencies—the Office of Thrift Supervision—eventually met a similar fate and was dismantled via the Dodd-Frank Act after criticisms of regulatory capture. *See* Dodd-Frank Act, § 313, 124 Stat. at 1523; *see also* Office of Thrift Supervision Integration, Dodd-Frank Act Implementation, 76 Fed. Reg. 43549 (July 21, 2011) (to be codified at 12 C.F.R. pt. 4, 5, 7, 8, 28, & 34) (rule transferring authority from the OTS to the Office of the Comptroller of Currency); Shriram Harid, *Report on Financial Crisis Singles Out the Office of Thrift Supervision*, REGBLOG (Apr. 27, 2011), http://www.law.upenn.edu/blogs/regblog/2011/04/senate-subcommittee-report-on-financial-crisis-singles-out-the-office-of-thrift-supervision.html.

99. Pike & Thompson, *supra* note 87, at 3 ("FDICIA strips regulators of much of their supervisory discretion over significantly undercapitalized . . . depositories . . . ."). As described by Senator Donald W. Riegle, Jr., Chairman of the Senate Committee on Banking, Housing, and Urban Affairs:

> The prompt corrective action provisions . . . . say, in effect: "Regulators, you should act earlier and more aggressively when a bank or thrift begins to get

olution" demanded by Congress through the PCA regime represented a simple and direct mandate to banking agencies to close down insolvent institutions. In contrast, reforms to repair the damage from regulatory arbitrage are relatively more complicated than simply responding to agency shortcomings by reducing discretion or shutting down weak regulators. The challenge of closing regulatory gaps involves more complex issues, not the least of which is identifying the sources of arbitrage opportunities, which are as numerous and dynamic as Wall Street is innovative, and which are difficult to manage as a result of lack of monitoring and measurement.[100] In fact, most systemically significant firms that effectively failed during the crisis had largely escaped capital requirements by employing regulatory arbitrage strategies involving funding loans through off-balance sheet vehicles and purchasing AAA-rated securities of dubious quality.[101] Despite leveraging bets on risky loan portfolios, LCFIs were not required to have any "skin in the game."[102] Regulatory gaps foster opportunism among financial innovators by allowing for areas of investment activity evading government supervision and, perhaps, comprehension.

The "shadow banking" industry. The shadow banking system exists under what many perceive to be a type of regulatory arbitrage.[103] "Shadow banks" are non-depository financial institutions that are not overseen by state or federal agencies or subject to the

into trouble. Get in there, correct the problems, and turn the place around, if you can. And if you cannot, sell the place, or close it down, before it becomes a loss to the deposit insurance system and a liability to the American people." 138 Cong. Rec. 19,533 (1992).

100. During the recent crisis, regulatory arbitrage arose in more ways than one, including the use of off-balance sheet investment vehicles, which allowed LCFIs to exploit regulatory loopholes under the Basel Accords. Viral V. Acharya et al., *Capital, Contingent Capital, and Liquidity Requirements*, *in* REGULATING WALL STREET: THE DODD-FRANK ACT AND THE NEW ARCHITECTURE OF GLOBAL FINANCE, *supra* note 26, at 143, 148 [hereinafter Acharya et al., *Capital*].

101. Matthew Richardson et al., *Securitization Reform*, *in* REGULATING WALL STREET: THE DODD-FRANK ACT AND THE NEW ARCHITECTURE OF GLOBAL FINANCE, *supra* note 26, at 469, 473, 476. For more on the role of issuers, securitization, and the credit ratings agencies in the crisis, see Edward I. Altman et al., *Regulation of Rating Agencies*, *in* REGULATING WALL STREET: THE DODD-FRANK ACT AND THE NEW ARCHITECTURE OF GLOBAL FINANCE, *supra* note 26, at 443.

102. Many LCFIs purchased AAA-rated tranches along with underpriced credit protection on those products through insurers, such as AIG, that allowed firms to take on additional risks without having to augment capital reserves. Acharya et al., *Capital*, *supra* note 100, at 147–50 (describing the several steps of regulatory arbitrage employed by LCFIs to get around Basel regulatory capital rules).

103. Bernanke, Address at the Council on Foreign Relations, *supra* note 48.

same rigorous safety and soundness examinations as depository institutions.[104] They include institutional investors such as hedge funds and pension funds, investment banks, and nonbank subsidiaries of depository institutions.[105] Such institutions can take advantage of gaps in the regulatory framework by exploiting profitable business lines that are under the radar of regulatory scrutiny and can subsequently engage in concentrated risk-taking in those areas to maximize profitability.[106]

While shadow banks do not have access to deposit insurance and are thus not subject to the same explicit moral hazard as traditional banks, such nonbank firms are still subject to implicit moral hazard to the extent an institution believes it is TBTF. Moral hazard is fueled when firms become highly concentrated in unregulated financial products through transactions with a multitude of counterparties spread throughout the financial sector, thus becoming "too interconnected to fail." While forbearance policies of the 1980s expose inaction on the part of the regulators in complying with their mandate, regulatory arbitrage reflects the decisions of the firms themselves to concentrate activities in areas in which regulators stand by, wholly without a mandate.

Due to the rapid pace of financial innovation and the limitations of regulatory agencies in keeping up with changes in complex derivatives markets and associated trading strategies, regulatory arbitrage presents an even greater cause for concern than policies of regulatory forbearance. This is particularly true in the context of privately traded OTC contracts not subject to standardization, where counterparty risks are not reduced via exchange trading.[107] Such contracts had damaging consequences during the Crisis of 2008–09 insofar as credit default swaps and other derivative instruments increased the risks posed by interconnectedness by linking the survival of major financial firms to the performance of unregulated contractual obligations among many firms.[108] Where firms availed themselves of regulatory arbitrage strategies to take on excessive risk, the result during the crisis included the virtual death

104. Mike Konczal, *Shadow Banking: What It Is, How it Broke, and How to Fix It*, THE ATLANTIC (Jul. 13, 2009, 1:08 PM), http://www.theatlantic.com/business/archive/2009/07/shadow-banking-what-it-is-how-it-broke-and-how-to-fix-it/21038/.

105. *Id.*

106. *Id.*

107. RECHTSCHAFFEN, *supra* note 41, at 163. ("Exchange-traded derivatives are designed to virtually eliminate counterparty risk. An organized exchange addresses the counterparty credit risk inherent in bilateral contracting by standardizing derivatives contracts to create a liquid market in the contracts themselves.").

108. McCoy, *supra* note 50, at 1358.

knell of financial markets—uncertainty. Investors' pronounced distaste for risk left unmeasured (or poorly measured by the institutions themselves through quantitative models), and therefore unaccounted for, is generally reflected by dramatic, stomach-turning reactions of the markets themselves.[109]

In contrast to the inexperienced and sometimes fraudulent thrift managers jumping into the waters of a newly deregulated environment,[110] the past two decades have been characterized more by the innovation of savvy financial engineers and a complex derivatives market that, with the blessing of Congress, escaped the grasp of regulators.[111] Instead of the proactive Congress of the 1980s seeking the fruits of healthy competition via deregulation, the years preceding the Crisis of 2008–09 were marked not only by an embrace of free markets by policymakers, but also by a failure to regulate entire areas of financial activity, despite warning.[112] Warnings were dismissed as alarmist and unfaithful to the market-discipline approach advocated by leading economic policymakers, including Federal Reserve Chairman Alan Greenspan, thus giving rise to regulatory gaps.[113]

Whereas simple regulatory forbearance and banker incompetence were critical factors leading to the S&L Crisis, sophistication and the skillful, but ultimately disastrous layering of risk played a much greater role in the subprime mortgage crisis.[114] As such, inventive financiers had no need to rely on insider abuse and outright fraud when greater profitability could be achieved through per-

109. *See* Jenny Anderson & Ben White, *Wall Street's Fears on Lehman Bros. Batter Markets*, N.Y. TIMES, Sept. 10, 2008, at A1, *available at* http://www.nytimes.com/2008/09/10/business/10place.html.

110. *See* Harris, *supra* note 59, at 266.

111. For a detailed look into the unsuccessful attempt to regulate OTC derivatives in the late 1990s by Brooksley Born, former CFTC Chairman and current member of the Financial Crisis Inquiry Commission, in the face of strong opposition from Congress and economy policymakers (including Alan Greenspan, Chris Cox, Robert Rubin, and Larry Summers), see *Frontline: The Warning* (PBS television broadcast Oct. 20, 2009), *available at* http://www.pbs.org/wgbh/pages/frontline/warning/view/; *see also The Financial Derivatives Supervisory Improvement Act of 1998 and the Financial Contract Netting Improvement Act: Hearing on H.R. 4062 and H.R. 4239 Before the H. Comm. on Banking and Fin. Servs.*, 105th Cong. 83–86 (1998) (statement of Brooksley Born, Chairperson, Commodity Futures Trading Commission).

112. *See Frontline: The Warning*, *supra* note 111; *see also* Manuel Roig-Franzia, *Credit Crisis Cassandra: Brooksley Born's Unheeded Warning Is a Rueful Echo 10 Years On*, WASH. POST, May 26, 2009, *available at* http://www.washingtonpost.com/wp-dyn/content/story/2009/05/25/ST2009052502127.html?sid=ST2009052502127.

113. *See* Roig-Franzia, *supra* note 112.

114. *See* Okamoto, *supra* note 3, at 200–03. R

fectly legitimate and legal means of avoiding regulatory oversight in an increasingly opaque financial marketplace.

B. Post-Crisis Responses to the Moral Hazard at Issue

Each crisis presents unique challenges, and therefore each reform effort must provide tailored responses. As noted by Federal Reserve Board Governor Kevin Warsh, "If you've seen one financial crisis, you've seen one financial crisis."[115] As a result, effective reform is particularly difficult. The S&L Crisis and the Crisis of 2008–09 both reveal a similar pattern of financial liberalization without consideration of long-term consequences. Beyond this, the distinctions that can be drawn between the two crises should inform the response to the Crisis of 2008–09.

1. Managing the explicit guarantee of government-backed deposit insurance versus the implicit promise of government bailouts

The post-crisis response to the shortcomings of regulators during the S&L Crisis, while difficult, did not require the structural revisions necessary to correct for the regulatory deficiencies that led to the Crisis of 2008–09. As discussed in Part II, regulatory forbearance was revealed as a lack of enforcement of existing regulations and as a failure to meet the public's expectation that insolvent institutions would be closed.[116] As such, problems with the execution of supervisory objectives could be properly addressed by simply cabining the discretion given to the banking agencies. For example, by enacting mandatory PCA guidelines that demand increasingly severe enforcement actions based on regulatory capital levels.[117]

Excessive risk-taking due to the moral hazard of deposit insurance is generally managed by placing more stringent requirements on institutions' managers via prudential regulation.[118] Thus, such reforms are more easily carried out when the regulatory framework already in place is consistent with the goals of reform: improving regulator accountability and ensuring that those regulators enforce more stringent prudential requirements as part of their existing monitoring and supervision functions, all in an effort to reduce excessive risk-taking by individual firms.

115. Janice Revell, *6 Signs of an Economic Rebound*, CNN MONEY (May 13, 2008), http://money.cnn.com/2008/05/09/pf/rebound_predictors.moneymag/index.htm.

116. *See supra* Part II.A.1.

117. *See* 12 U.S.C. § 1831o(a)(2) (2006); *see also* CARNELL ET AL., *supra* note 5, at 279–92. R

118. *See* CARNELL ET AL., *supra* note 5, at 252–53. R

In contrast, the moral hazard problem created by being TBTF cannot be managed by merely increasing the stringency of operating requirements. Despite the fact that liabilities of nonbank financial firms were not backed by deposit insurance, the size and complexity of several institutions resulted in a belief that some firms were too interconnected to fail—the product of an implicit guarantee.

In realizing the advantages of the government's implicit TBTF safety net, systemically significant LCFIs impose negative externalities on society.[119] Because such guarantees are paid via an ad hoc "bailout" mechanism with no planned source of funding, the public bears the burden while bailed-out institutions enjoy the benefits. Thus, the financial firms imposing enough systemic risk on the financial system to take advantage of publicly funded emergency liquidity are not forced to internalize the costs of that risk.[120] This problem has less to do with regulator discretion or the management inexperience observed during the S&L Crisis and instead stems primarily from design failures within the existing regulatory infrastructure. While some structural shortcomings can be partially remedied by closing gaps in oversight and enhancing supervisory responsibilities,[121] merely adding traditional prudential reforms to the current regulatory framework does little to cure the dangers arising from systemic importance, the relationships among financial institutions, and the unmitigated build-up of counterparty risk.

Elimination of supervisory gaps and increased supervision are necessary first steps in regulatory reform and have recently been addressed by provisions to improve transparency in the derivatives markets under Title VII of the Dodd-Frank Act, which includes

119. Viral V. Acharya et al., *Taxing Systemic Risk*, *in* REGULATING WALL STREET: THE DODD-FRANK ACT AND THE NEW ARCHITECTURE OF GLOBAL FINANCE, *supra* note 26, at 121, 122 [hereinafter Acharya et al., *Taxing Systemic Risk*] ("That some financial institutions contribute more than others to the overall capital shortfall in a crisis is a prototypical example of the negative externality of systemic risk in the financial sector. Markets do not price negative externalities, so if unchecked, they get produced in excess.").

120. *See* Viral V. Acharya et al., *Systemic Risk and Deposit Insurance Premiums*, FED. RES. BANK N.Y. ECON. POL'Y REV., Aug. 2010, at 89, 91–92 [hereinafter Acharya et al., *Deposit Insurance*].

121. U.S. DEP'T OF THE TREASURY, FINANCIAL REGULATORY REFORM—A NEW FOUNDATION: REBUILDING FINANCIAL SUPERVISION AND REGULATION 6–7 (2009) (on establishing "comprehensive regulation of financial markets" by bringing the OTC derivatives and asset-backed securities markets into a coordinated regulatory framework).

clearing and margin requirements.[122] However, as Part III will discuss in further detail, while the Dodd-Frank Act closes some loopholes, it does not necessarily relieve the difficulties of marketplace opacity, nor does it adequately address the perverse incentives that encourage financial firms to grow too interconnected to fail.

2. Financial liberalization, financial crisis, and shortsighted regulatory design

The deregulatory environment preceding each crisis resulted from a free-market approach embraced by both Republican and Democratic administrations, which has long rested on principles of laissez-faire capitalism.[123] While deregulation has led to observable benefits,[124] some commentators have suggested that a link exists between deregulation efforts and banking crises, indicating that the most troubling part of financial liberalization may be the resulting tendency to create financial systems more susceptible to systemic risk.[125] Such a system would thus be more vulnerable to systemic crises and, as a result, manifest more TBTF dilemmas faced by the government, similar to the decisions on whether or not to rescue failing firms during the Crisis of 2008–09.

By "amplifying" the stages of the business cycle, financial liberalization creates a difficult tradeoff between the benefits of deregulation, particularly economic growth and expansion, and minimizing the dangers of relying on market discipline alone—i.e., the risk of economic downturns as a result of bursting asset bubbles.[126] Deregulation efforts by Congress, which follow a theme of broadening lending powers and permissible investments in response to a particular industry's aspirations of improved competitiveness, in turn place greater pressure on banks to expand into more risk-laden areas.[127]

Despite the many differences between the causes of the S&L Crisis and the Crisis of 2008–09, the regulatory environments pre-

122. Dodd-Frank Act, Title VII, 124 Stat. at 1641 ("Wall Street Transparency and Accountability Act").

123. *See* Charles G. Leathers and J. Patrick Raines, *Some Historical Perspectives of "Too Big to Fail" Policies*, *in* TOO BIG TO FAIL: POLICIES AND PRACTICES IN GOVERNMENT BAILOUTS, *supra* note 14, at 3. R

124. Key benefits of deregulation include the efficient distribution of resources, economic development, and trade growth. Wilmarth, *Financial Liberalization*, *supra* note 79, at 77.

125. *Id.*

126. *Id.* at 77–78 (on the seven general stages of banking crises associated with deregulation).

127. *See id.*

ceding both crises underscore the dangers of shortsighted deregulatory initiatives that do not adequately account for long-term implications. Deregulation in and of itself is not the sole cause of financial crisis, but deficiencies in regulatory design often go undetected until times of crisis, when the costs are at their highest.[128] Thus, each case highlights the importance of pursuing viable long-term objectives when crafting both deregulatory initiatives and regulatory reforms. As observed with regard to the regulatory changes preceding the S&L Crisis, for such measures to be successful in the long run, they must give adequate consideration to the impact on the competitive dynamic among financial institutions and the role moral hazard plays in the decision-making of both depository and non-depository institutions.

III.
GREATER ANTICIPATORY REGULATION TO MANAGE SYSTEMIC RISK

This Part focuses on structural changes needed to combat the challenges of TBTF and moral hazard following the Crisis of 2008–09. While preventing the consequences of systemic risk has been a centerpiece of the regulatory framework, the regime has proven inadequate in managing the particular type of contagion associated with the systemic risk exhibited during the subprime mortgage crisis. As the separation between commercial and investment banking has blurred and distinctions among financial services have eroded, the concentration of business housed in large "money center" banks has grown.[129] Since the S&L Crisis and the repeal of Glass-Steagall, the landscape of the financial services industry has changed dramatically. Moreover, the expanding presence of non-depository financial institutions has shifted the significance of systemic risk arising from contagions of panic among depositors to that arising from contagions of panic across the firms themselves. Because this risk arises from the relationships among transactional counterparties, as opposed to the overall health of an institution in isolation, traditional prudential regulation methods that impose stricter operating requirements on single institutions cannot suffi-

128. *See* Arthur E. Wilmarth, Jr., *Does Financial Liberalization Increase the Likelihood of a Systemic Banking Crisis? Evidence from the Past Three Decades and the Great Depression*, *in* TOO BIG TO FAIL: POLICIES AND PRACTICES IN GOVERNMENT BAILOUTS, *supra* note 14, at 77, 77–78, 96 (on the boom and bust cycles of deregulation and banking crises, recommending that regulators consider the "long-term economic risks of financial liberalization programs"). R

129. *See* Wilmarth, *Universal Banking*, *supra* note 22, at 975–80.

ciently address the implicit moral hazard that arises from expectations of TBTF protection.

A. Systemic Risk: Contagion and Concentration

The lack of a proper regulatory infrastructure to address systemic risk from both an ex ante and an ex post perspective increases the ultimate cost of financial crises borne by the public.[130] Regulation of systemic risk has historically focused on the prevention of bank failures rather than on the systemic risk itself. Deposit insurance has traditionally been viewed as the optimal way to prevent depositor runs and mitigate contagions of bank panics caused by commercial bank failures.[131] As discussed in Part II, although the explicit guarantee of federal deposit insurance creates moral hazard, its purpose in protecting depositors has ostensibly served a generally accepted and legitimate policy aim.[132] However, this objective of protection does not transfer as well to the risk-taking beneficiaries of an implicit TBTF guarantee.

Furthermore, the recent crisis calls attention to the concentrations of financial power in a few dominant mega-firms—a problem distinct from, and in addition to, contagion, the primary concern of previous crises.[133] Thus, while contagion speaks to the classic case of Continental Illinois Bank in 1984 and its correspondent banking relationships, or the more recent case of Bear Stearns and the potential domino effect through the credit default swaps market,[134] systemic risk arising through concentrations of financial power results from the pure dominance of a firm whose failure has the potential to disrupt the market well beyond its counterparty relationships.[135] When the country's largest and most complex institutions, including Washington Mutual, Wachovia, Lehman Brothers, AIG, Merrill Lynch, and Citigroup approached failure in 2008, their status as financial giants and their potential failures posed great risk to the economy, and as a result, the government applied

130. Harris, *supra* note 59, at 254–55.

131. Schwarcz, *Systemic Risk*, *supra* note 35, at 210–11; MILTON FRIEDMAN & ANNA JACOBSON SCHWARTZ, NAT'L BUREAU OF ECON. RESEARCH, A MONETARY HISTORY OF THE UNITED STATES, 1867–1960 440 (1963).

132. *See* CARNELL ET AL., *supra* note 5, at 309–11. R

133. *See* Todd & Thompson, *supra* note 13, at 12. R

134. Thompson, *supra* note 17, at 3. R

135. *Id.* at 4–5 (describing how a systemically significant institution's failure may have "spillover effects that impede the functioning of broader financial markets and/or the real economy" as a result of the firm's dominance in volume over certain financial services).

resources, virtually limitless in nature, in its efforts to avoid systemic collapse.[136]

Because the dangers of the implicit guarantee are particularly hazardous, the costs of systemic risk should be internalized by the guarantee's beneficiaries so as to avoid economically inefficient cost shifting to taxpayers during a systemic event. Furthermore, a framework to combat TBTF should attempt to counteract an obvious collective action problem: because market participants incur the costs of systemic failures, but do not themselves bear those costs, it is unlikely that institutions will voluntarily curb activities resulting in greater systemic risk, given that they stand to gain from such transactions.[137] Without providing a new systemic risk regulator with a mandate of prevention (in addition to traditional monitoring and supervision), LCFIs will be incentivized to continue increasing concentrations of financial power, wholly without charge for the implicit guarantee of public funding that they enjoy during times of crisis.

While commercial banks have been subject to the most rigorous regulatory requirements and supervision in the financial services industry, financial modernization has all but eliminated the distinctions between banks and other types of financial firms.[138] As a result, the number of financial service providers capable of posing systemic risk has increased.[139] An institution, regardless of its status as depository or non-depository, can now be well-capitalized, have excess liquidity, and practice sound risk management, while at the same time grow in size and complexity, form counterparty relationships, and concentrate its assets in high-risk areas. Thus, systemic risk regulation must cast a wide net and be comprehensive in its reform by addressing all financial firms, regardless of type.

B. Prudential Regulation Alone is Insufficient

More stringent regulation, greater limitations on allowable investments, and closer supervision of operations is a natural response to a banking crisis. The traditional bank regulatory framework of prudential regulation imposed on individual banks

136. *See* Okamoto, *supra* note 3, at 200, 203; *see also* Acharya et al., *Taxing Systemic Risk*, *supra* note 119, at 123; Moss, *supra* note 14, at 5, 8. R R

137. *See* Schwarcz, *supra* note 35, at 206.

138. *See* Wilmarth, *Universal Banking*, *supra* note 22, at 975–81.

139. Rose Marie Kushmeider, *The U.S. Federal Financial Regulatory System: Restructuring Federal Bank Regulation*, 17 FDIC BANKING REV., no. 4, 2005, at 1, 17, *available at* http://www.fdic.gov/bank/analytical/banking/2006jan/article1/article1.pdf.

has rested on the assumption that the moral hazard problem at issue is one of an explicit nature—i.e., the moral hazard of deposit insurance that is alleviated to the extent that comprehensive operating requirements imposed on individual institutions will deter the excessive risk-taking that puts the insurance fund at risk. Up to a point, however, increased prudential regulation is of little marginal benefit in the context of systemic risk because such methods affect only the risk-taking of individual firms and do little to regulate an institution's counterparty risks, or the risk it poses to the economy as a whole.[140] Thus, comprehensive reform must include both anticipatory and mitigating measures to, first, discourage the rapid build-up of systemic risk by changing the incentive structure of TBTF and, second, minimize the systemic impact when a systemically important institution faces insolvency. The latter point operates on the assumption that despite preventive efforts, systemic events will still occur in the future. Without combating systemic risk both ex ante and ex post, neither the consequences nor the causes of the Crisis of 2008–09 will be fully addressed.

The effectiveness of reform lies principally in combating the source of implicit moral hazard. Unfortunately, the incentive to grow TBTF is now stronger as a result of bailouts during the recent crisis,[141] while the systemic impact of the government's refusal to rescue firms such as Lehman Brothers may have only reinforced the notion that systemic effects are severe, thus merely emphasizing the importance of avoiding those effects through bailout or otherwise.[142] Because an implicit guarantee has no set dollar amount, and because its funding source is determined ad hoc only at the time of crisis, nothing in the regulatory system forces a TBTF firm to internalize the cost of the systemic risk it poses prior to a systemic collapse. Therefore, anticipatory regulation is better suited to deal with such implicit guarantees where mere stringency of operating requirements falls short.

The reform effort must also concede that the rate at which the government is able to adapt regulatory processes, even with a vigilant systemic risk regulator in place, is unlikely to keep up with the swift pace of financial innovation. Prudential requirements can only be made so stringent and, in a rapidly changing marketplace, cannot realistically address the supervisory needs of every new risk-shift-

140. *See generally*, Moss, *supra* note 14, at 9 ("[S]ystematically significant institutions should face enhanced prudential regulation . . . .") R

141. *See* Bernanke, *Opening Remarks*, *supra* note 48, at 3.

142. *See* Anderson & White, *supra* note 109; *see also* Okamoto, *supra* note 3, at 200. R

ing financial instrument to hit Wall Street. The limitations of prudential regulation suggest that a broader and more prophylactic approach to managing systemic risk would be more effective to slow the race toward unnecessary interconnectedness and concentrations of financial power. A preventive strategy, in addition to being better suited to dealing with implicit moral hazard, is more likely to strike the right balance between freely thriving innovation and a parade of complex instruments flooding the marketplace before their risks are understood by the parties.[143] Lack of such a balance can result in opacity in the financial system, leading to greater uncertainty during economic downturns—i.e., the panic and contagion that originates from a systemic event.

C. Combating TBTF and Implicit Moral Hazard Through Anticipatory and Mitigating Measures

Reform of the banking system has been remarkably hard to come by for almost a century, despite the fact that the debate over how the highly decentralized system should be restructured has been going on for just as long.[144] This section argues that the reforms intended to address TBTF fall short of resolving the underlying structural problems that incentivize risk-taking and externalizing behaviors among LCFIs, and they thus fail to reduce implicit moral hazard among systemically significant institutions. Instead, effective regulation of systemic risk requires a revision to the regulatory structure that incorporates both preventive and mitigating measures. Previous commentators have noted that ex ante mechanisms for managing systemic risk are needed to improve the regulatory framework.[145] A successful approach must, first, embrace the traditional role of prudential operating requirements by imposing on the largest financial institutions stricter capital adequacy and risk management standards (such as those recently enacted under the Dodd-Frank Act) and, second, incorporate greater anticipatory regulation of systemic risk. Such a strategy would reflect a dual purpose in managing systemic risk by combating the incentives to be-

143. *See Wall Street to Main Street: Is the Credit Crisis Over and What Can the Federal Government Do to Prevent Unnecessary Systemic Risk in the Future?: Hearing Before the Joint Econ. Comm.*, 110th Cong. 11 (2008) (statement of Paul A. Volcker, Former Chairman of the Fed. Reserve Board of Governors) ("[T]he executives of these companies, I think, to put it mildly, have great difficulty in really understanding the amount of risk and complexity involved in their organizations.").

144. *See generally*, Howard A. Hackley, *Our Baffling Banking System*, 52 VA. L. REV. 565 (1966).

145. *See, e.g.*, Acharya et al., *Deposit Insurance*, *supra* note 120, at 97; Moss, *supra* note 14, at 12; Thompson, *supra* note 17, at 1. R

come TBTF and by preserving financial stability when systemic crises do (as they inevitably will) arise.

In addition to greater prudential regulation and filling in regulatory gaps, a set of economic disincentives should be established to deter firms from becoming "too interconnected to fail," carefully balanced against the danger of overdeterrence of economically productive financial innovation. Greater market discipline should be imposed through these disincentive structures to make it less attractive and more costly to become systemically significant. Because lack of liquidity among LCFIs during a systemic crisis threatens economic instability, such reforms should also focus on nonpublic emergency funding sources that might be used to reduce market volatility as an alternative to publicly funded bailouts.[146]

1. Ex ante: disincentives to manage firm size and interconnectedness

Regardless of whether government intervention was justified in 2008, the market is likely to have taken note of the expanded government safety net, thus compounding moral hazard and exacerbating the cycle of excessive risk-taking and costly government assistance.[147] To counteract these perverse incentives created by bailouts, regulatory tools should impose greater market discipline than the existing framework demands.

As discussed above,[148] the systemic risk regulator should pursue the objective of anticipatory regulation to discourage the buildup of systemic risk. In doing so, the regulator must attempt to strike a balance between an overly restrictive environment where financial innovation is stifled, and a regulatory framework that effectively curbs systemic risk that threatens the overall economy.[149] By establishing economic disincentives, such as taxes, insurance premiums, or other types of regularly charged assessments based on levels of systemic risk, regulators can impose greater regulatory burdens on only those institutions for which the benefits of "systemically important" status outweigh the costs and force such institutions to internalize the costs of that status.[150] An explicit set of supervisory policies to guide decision-making by potentially systemically important firms would dramatically reduce the benefits received and the externalities imposed by TBTF institutions by sub-

146. *See infra* Part III.C.2.

147. *See* Bernanke, *Opening Remarks*, *supra* note 48, at 3.

148. *See supra* Part III.C.

149. *See* Burson, *supra* note 33, at 15.

150. *See* Acharya et al., *Taxing Systemic Risk*, *supra* note 119, at 124–26; Moss, *supra* note 14, at 9–10; Acharya et al., *Deposit Insurance*, *supra* note 120, at 92. R

jecting firms that pose greater levels of systemic risk to greater regulatory interference and supervisory attention.[151] Altering the incentive structure that encourages firms to become TBTF is the most efficient and effective way to correct excessive risk-taking behavior by TBTF firms or those seeking TBTF status.[152]

The Dodd-Frank provisions intended to combat TBTF include prudential regulations,[153] but more stringent prudential standards may not be sufficiently costly to a LCFI to discourage firms from growing TBTF. Even if such requirements make being large more expensive in the short-term, existing TBTF firms already benefit from the lower cost of capital resulting from the implicit government safety net.[154] Thus, such requirements are of limited effectiveness in combating the moral hazard of TBTF because they do little to regulate the interconnectedness among firms or the incentives to grow.[155]

Stricter prudential regulations, such as higher capital and liquidity requirements, are included within a broader category of

151. *See* Thompson, *supra* note 17, at 6 (arguing for a system of "progressive systemic mitigation," or an explicit set of regulations and supervisory policies based on categorizing institutions and designed to reduce the advantages of being systemically important). R

152. *See* CARNELL ET AL., *supra* note 5, at 283 ("To control the undesirable behavior resulting from perverse incentives, one can regulate the behavior or change the incentives. . . . The most efficient and effective way to correct undesirable behavior is by mitigating the incentives giving rise to that behavior."). R

153. The Financial Stability Oversight Council

> may make recommendations to the Board of Governors concerning the establishment and refinement of prudential standards and reporting and disclosure requirements applicable to nonbank financial companies supervised by the Board of Governors and large, interconnected bank holding companies, that—
>
> (A) are more stringent than those applicable to other nonbank financial companies and bank holding companies that do not present similar risks to the financial stability of the United States; and
>
> (B) increase in stringency, based on the considerations in subsection (b)(3).

Dodd-Frank Act, § 115, 124 Stat. at 1403. Stricter standards should include "(A) risk-based capital requirements; (B) leverage limit; (C) liquidity requirements; (D) a contingent capital requirement; (E) resolution plan and credit exposure report requirements; (F) enhanced public disclosures; (G) concentration limits; (H) short-terms debt limits; and (I) overall risk management requirements." *Id.*

154. Acharya et al., *Taxing Systemic Risk*, *supra* note 119, at 133–34.

155. The primary tools to combat systemic risk include instituting a new regulatory body in the form of the Financial Stability Oversight Council and, under the Collins amendment, increasing capital adequacy standards for those institutions found to pose significant systemic risk. Dodd-Frank Act, Title II, 124 Stat. at 1442 ("Orderly Liquidation Authority"); *id.*, Title I, 124 Stat. at 1391 ("Financial Stability"); *id.* § 171, 124 Stat. at 1435 (on leverage and risk-based capital requirements).

regulatory measures discouraging the growth and interconnectedness that increase systemic risk.[156] More stringent operating requirements can also be swiftly imposed and easily measured, making it immediately more expensive for firms to incur unnecessary growth in the short-term. However, these provisions merely represent a continuation of standard methods of banking regulation seen time and again—greater government intervention by increasing the number of regulatory bodies, relatively more intrusive oversight and monitoring, and stricter operating requirements under conventional measures of bank safety and soundness. A disincentives framework should go further still to fully address the long-term consequences of implicit moral hazard.

While stronger disincentives represent a foundational shift in regulatory policy, solutions that fail to consider structural impediments to optimal regulation can exacerbate the costs of financial crises by providing perverse incentives during the crises. The establishment of a systemic risk regulator and subsequent identification of TBTF institutions that will receive increased scrutiny and be subject to requirements distinct from those imposed on non-systemically significant firms may actually exacerbate the moral hazard of TBTF. The market will likely take note of such institutions and continue to treat them as TBTF based on their transparent and observable systemically important status. Thus, if such reforms under the Dodd-Frank Act are not also combined with a corresponding revision to the incentives to seek TBTF status, LCFIs are unlikely to alter their behaviors. As observed with regard to the S&L Crisis, failure to adjust regulatory policies when reforms can otherwise worsen moral hazard can lead to unintended consequences much further down the line.[157]

This approach is also superior to other alternatives for preventing firms from becoming TBTF, such as a providing regulators with

156. *Establishing a Framework for Systemic Risk Regulation: Hearing Before the S. Comm. on Banking, Hous., and Urban Affairs*, 111th Cong. 63 (2009) (statement of Sheila C. Bair, Chairman, Federal Deposit Insurance Corporation).

> Financial firms that pose systemic risks should be subject to regulatory and economic incentives that require these institutions to hold larger capital and liquidity buffers to mirror the heightened risk they pose to the financial system. In addition, restrictions on leverage and the imposition of risk-based premiums on institutions and their activities would act as disincentives to growth and complexity that raise systemic concerns.

157. *See supra* Part II.A.1 (discussing a failure in regulatory design where Congress neglected to implement a corresponding prudential regulation to manage additional risks and counteract the moral hazard effects of various deregulatory initiatives).

a "break-up authority" to dismantle LCFIs into smaller institutions.[158] In contrast to placing responsibility in the hands of the regulators to make judgment calls on when LCFIs have outgrown their utility, economic disincentives force the firms themselves to make the ultimate decision of when growing in size and complexity has become too costly to make economic sense. Furthermore, because such determinations will concern the largest and most high profile financial firms, the decisions also risk becoming highly politicized if made by government actors.

Firms are also better suited and better positioned than government actors to determine whether the costs of systemic risk (to the firms themselves) are justified—as opposed to the government's determination that the costs of systemic risk (to the financial system) are not justified. FDIC Chairman Sheila Bair, a proponent of a disincentives approach, has noted that requiring firms to be proactive through self-monitoring should be included as a mechanism to con-

158. A more drastic approach to fighting TBTF is the delegation of "break-up" authority, empowering regulators to identify institutions posing excessive systemic risk and order those firms to downsize. Peter Boone & Simon Johnson, *How Big is Too Big?*, N.Y. TIMES ECONOMIX BLOG (Nov. 26, 2009, 7:18 AM), http://economix.blogs.nytimes.com/2009/11/26/how-big-is-too-big/?pagemode=print. In 2009, Rep. Paul Kanjorski sought to give regulators preemptive authority to break up the largest 50 financial firms. He also proposed a number of criteria to be used by a Financial Services Oversight Council to determine when financial firms should be broken apart, such as if their "size . . . , scope, nature, scale, concentration, interconnectedness, or mix of activities . . . poses a grave threat to the financial stability" of the country. *See* STAFF OF H. FIN. SERV. COMM., 111TH CONG. FINANCIAL STABILITY IMPROVEMENT ACT OF 2009 amend. 39 (Comm. Print 2009) (amendment of Rep. Paul Kanjorski, Member, H. Fin. Serv. Comm.), available at http://www.house.gov/apps/list/speech/financialsvcs_dem/amdt_in_nature_of_substitute_to_hr_2609_10_16_09.pdf; *see also* Press Release, House Fin. Serv. Comm., Kanjorski Releases Amendment to Address Companies That Are "Too Big to Fail" and Prevent Future Bailouts (Nov. 18, 2009), http://democrats.financialservices.house.gov/press/PRArticle.aspx?NewsID=557.

While efforts to establish any type of break-up authority in the U.S. have been unsuccessful thus far, the strategy has gained more traction in Europe. European Commission officials were swift in adopting break-up authority under Competition Law in 2009 after determining that several TBTF firms had adversely impacted the competitiveness of the banking sector. Edward Greene & Katia Kirova, *"Too Big to Fail"—Should Breaking Up Large Financial Institutions Be an Answer?: U.S. and European Approaches*, 16 COLUM. J. EUR. L. ONLINE 19, 20 (2009). As a result, the Commissioner for Competition, Neelie Krose, broke up ING Group NV as a first step in October 2009. *Id.* The Commission also pressured the United Kingdom to downsize its largest banks, resulting in the forced sale of parts of the Royal Bank of Scotland, Lloyds Banking Group, and Northern Rock. *Id.* at 21.

trol risk-taking within a robust and complex financial system.[159] Placing this responsibility on firms themselves forces those best situated to understand the levels and nature of their risk exposure as well as the additional costs to be incurred as a systemically important institution. Through managerial discretion and improved market discipline, the true costs of a firm's activities will be borne not by the public, but by the firm's shareholders.[160] As with all strategic decisions, financial institutions can "organically" determine the appropriate pace and amount of growth to achieve, as regulatory policy will push such firms to reconsider existing business models and to evaluate the tradeoffs.[161] Placing greater reliance on market discipline, but only after the costs of systemic risk are internalized by LCFIs, is also more desirable than depending solely on the regulatory body. That is to say, in establishing the Financial Stability Oversight Council to monitor systemic risk to the financial system, policymakers should not also award regulators too much discretion with which to determine what is "too big to succeed." Unlike the regulator discretion that Congress attempted to curb through the PCA regime following the S&L Crisis, objective guidelines, such as capital ratios, cannot be implemented in the realm of systemic risk regulation reform as a practical matter, given that TBTF is a question of not only size, but of interconnectedness, complexity, and business model. Furthermore, a body composed entirely of agency heads is still subject to risks of regulatory capture.[162] These potential issues offer a reprise of two familiar themes of regulatory failures: regulatory capture and the misuse of discretion by supervisory authorities.

159. *See Establishing a Framework for Systemic Risk Regulation: Hearing Before the S. Comm. on Banking, Hous., and Urban Affairs*, 111th Cong. 62 (2009) (statement of Sheila C. Bair, Chairman, Federal Deposit Insurance Corporation).

160. Acharya et al., *Taxing Systemic Risk*, *supra* note 119, at 126.

161. *Id.* ("These firms will therefore be encouraged to rethink their business models. In particular, they will have to consider reducing their scope, scale, risk exposures, and interconnectedness, thus trading off the returns from such activities against the insurance premiums attached to them.").

162. Under Title I of the Dodd-Frank Act, membership of the Financial Stability Oversight Council includes ten voting members: Secretary of the Treasury (Chair of the Council), Chairman of the Federal Reserve, Comptroller of the Currency, Director of the Consumer Financial Protection Bureau, Chair of the Securities and Exchange Commission, Chair of the Federal Deposit Insurance Corporation, Chair of the Commodity Futures Trading Commission, Director of the Federal Housing Finance Agency, Chair of the National Credit Union Administration Board, and an independent member with insurance expertise, appointed by the President with advice and consent of the Senate. *See* Dodd-Frank Act, Title I, § 111, 124 Stat. at 1392–93.

Where sufficient disincentives are in place to diminish the implicit moral hazard of TBTF policies, regulators can place greater reliance on firms forced to incur the costs of imposing systemic risk to engage in a certain level of self-policing. This achieves a better balance between intrusive regulatory intervention in firm-specific decisions regarding growth and strategy, and complete dependence on free markets alone to produce optimal economic outcomes, even where markets are distorted by problems of moral hazard. With a new oversight mechanism over systemic risk and TBTF firms, regulatory policies should still permit managers to make strategic institutional decisions, so long as those firms bear the costs of imposing additional risk to the overall financial stability of the system.

While some of the Dodd-Frank provisions designed to address systemic risk mirror the responses to the S&L Crisis, the prominent role of implicit moral hazard in the Crisis of 2008–09 warrants more than just stricter operating requirements and enhanced supervision of LCFIs. An ex ante approach that explicitly prices the costs of imposing systemic risk places that burden squarely on the firms themselves. Moreover, it places some faith in a type of market discipline approach to shape business strategy and in the ability of private-sector actors to make appropriate cost-benefit calculations for individual financial institutions. Furthermore, an ex ante approach to inhibiting systemic risk at its source is a more efficient way to address the regulatory lag that results from the government's inability to keep up with changes in the financial markets. A framework of anticipatory regulation channels the ultimate decision regarding a firm's appetite for risk to those who can most efficiently weigh the costs and benefits.

2. Ex post: an industry-funded emergency liquidity pool

Economic disincentives in the form of systemic risk-rated assessments paid by LCFIs should be used to fund an industry-specific liquidity pool for exclusive use during systemic events. Some commentators have previously suggested instituting deposit insurance reforms that link a portion of the cost of FDIC insurance premiums to the level of systemic risk posed by an individual depository institution.[163] However, a uniform risk-based system for all financial firms, regardless of legal status, should be established. While the

163. *See* Acharya et al., *Deposit Insurance*, *supra* note 120, at 92 (arguing that the extent of systemic risk in the financial sector is a key determinant of efficient deposit insurance premiums, and proposing a model for measuring actuarially fair deposit premiums).

calculations may be complex,[164] the concept is simple and can be applied equally well to nonbank financial firms: an actuarially fair assessment levied on systemically significant firms, both depository and non-depository, covering the expected cost to an emergency liquidity provider during times of systemic crisis, should increase in relation to both the risk of the individual firm's failure and the related risk of joint insolvencies. These assessments should be aggregated into an emergency source of liquidity, the use of which would be contingent upon a determination by an independent regulator tasked with monitoring systemic risk, such as the Financial Stability Oversight Council, that a firm's failure would qualify as a systemic event. The funds should be used during such an event and should preclude the use of taxpayer dollars to cover the costs of mitigating the effects of a systemic crisis.[165]

In order to avoid the pool becoming a "bailout fund," and thus merely an additional source of comfort to TBTF firms, a liquidity backstop must also be the means through which the costs of systemic risk are imposed on those who pose such risk. Thus, the pool must be funded solely through the risk-rated contributions, in whatever form they may take, from systemically significant firms. The assessments, in combination with the liquidity pool, would act as both (1) a disincentive, in that firms can evade being levied assessments by avoiding unnecessary growth and high concentrations of investment activity resulting in "over-interconnectedness" insofar as the costs outweigh the benefits for individual firm, and (2) the source of emergency liquidity if an institution threatens to bring down the rest of the economy.

A significant challenge of the reform effort involves improving the government's credibility with regard to pledges against future

164. While outside the scope of this Note, a significant part of this proposal depends upon the accurate measurement of systemic risk, both in determining which firms would be subject to the systemic risk premiums, and in calculating the actual dollar amount of the premiums, whether it exists in the form of an assessment, tax, or fee. For more on the challenges of, and proposed solutions to, measuring systemic risk, see Acharya et al., *Measuring Systemic Risk*, *in* REGULATING WALL STREET: THE DODD-FRANK ACT AND THE NEW ARCHITECTURE OF GLOBAL FINANCE, *supra* note 26, at 87.

165. As part of a mitigation function, the framework for systemic risk regulation should include a mechanism that acts as a government backstop to provide necessary liquidity to "soften the blow" of a systemic event. This would ensure that the least harm is imposed on the public as a consequence of a firm's systemic importance and subsequent insolvency. *See* Schwarcz, *supra* note 35, at 241–42 (concluding that "[a] regulation establishing a liquidity-provider of last resort . . . is the approach to minimizing systemic risk that would have the best chance of success . . . ").

bailouts,[166] while also preparing for the inevitable reality of future systemic events. Thus, any liquidity backstop must be accompanied by detailed guidance that narrows the definition of a systemic event such that use of the liquidity pool is limited to only the most extreme circumstances. This ensures that a guarantee provided by a liquidity provider of last resort would only function when the stability of the financial system is threatened. Furthermore, it allows the government to maintain at least some discretion and a certain element of "constructive ambiguity"[167] over the use of emergency liquidity when faced with a potential systemic event. These measures would require what would then be an explicit guarantee to be transparent in its process (and thus preferable to the government's ad hoc actions during the recent crisis), while its use would necessarily remain discretionary, so as to not exacerbate the moral hazard of TBTF.

This framework concedes that even a robust disincentives framework cannot eliminate all systemic risk, nor deter all firms from becoming TBTF. Future systemic crises are inevitable, though their frequency can be minimized if strong anticipatory regulation is combined with prudential regulation to encourage greater market discipline. TBTF cannot be eliminated altogether, nor should large financial mega-firms be banned from existence. Such firms fulfill an important role in the global economy, a fact Federal Reserve Chairman Ben Bernanke has recognized, saying that a "technologically sophisticated and globalized economy" needs "large, complex, and internationally active financial firms."[168] Proper mechanisms for reducing the adverse impact of systemic events are therefore as important as reducing systemic risk itself.

Systemic risk can be viewed as a negative externality that can be internalized by imposing the costs of systemic events on those firms

166. In response to government intervention during both the S&L Crisis and the Crisis of 2008–09, "no-bailout" pledges unsurprisingly followed. *See Wall Street to Main Street: Is the Credit Crisis Over and What Can the Federal Government Do to Prevent Unnecessary Systemic Risk in the Future?: Hearing Before the Joint Econ. Comm.*, 110th Cong. 39 (2008) (statement of Alex J. Pollock, Resident Fellow, American Enterprise Institute) ("[F]ollowing the 1980s bust, the Secretary of the Treasury said about the reforms of 1989 and the early 1990s, they have the motto of 'Never again.' And those are the mottos of every reform. 'Never again.' Yet, Mr. Chairman, here we are again.").

167. *See infra* Part I.A.1.

168. *See* Kristina Cooke, *Bernanke: Too Big to Fail a "Pernicious" Problem*, REUTERS (Mar. 20, 2010, 10:17 AM), http://www.reuters.com/article/idUSTRE62J0SM20100320.

posing the risk.[169] This Note argues that what is missing from the regulatory framework is a method of transforming the unknown costs of an implicit guarantee of unknown funding during an ad hoc rescue into an explicit guarantee funded by those posing the most risk during times of severe market distress. A liquidity provider of last resort allows the process of handling distressed, systemically significant institutions to be transparent and orderly, instead of opaque, chaotic, and vulnerable to politicization.[170] In addition, it offers a mechanism through which the regulatory structure can be altered to address the need for both ex ante and ex post solutions to the TBTF problem.

D. Addressing Criticisms

Critics would be correct to note that an emergency liquidity pool would be dangerous in its potential to exacerbate moral hazard to the extent that institutions are aware of the available "bailout fund" for exclusive use by TBTF firms. Thus, a crucial part of this proposal is designed to ensure that the parties bearing the risk are the ones that fund such a pool and that there are appropriate avenues of government support that do not risk another taxpayer-funded bailout. If the firms themselves do not fund the liquidity backstop, then it serves no purpose but to encourage TBTF. In addition, the liquidity pool should be used to fund the expenses of resolution of failed nonbank financial firms under Title II of the Dodd-Frank Act ("Orderly Liquidation Authority").

The circumstances in which the fund is used must be narrowly defined such that a threat of mere insolvency of a LCFI does not trigger the fund's use. Instead, a "systemic event" must be defined to preclude use of the liquidity pool when a firm's potential failure does not rise to the level of a systemic threat. Further, the costs associated with being systemically significant must be so high that they simply do not make economic sense for most firms. The vast majority of firms are not large and complex, nor TBTF, and will remain unaffected by the proposed systemic risk regulation regime. Substantially increasing the costs associated with being systemically significant will only deter firms from becoming large and complex

169. *See* Acharya et al., *Deposit Insurance*, *supra* note 120, at 97.

170. *See* Arthur E. Wilmarth, Jr., *The Dodd-Frank Act: A Flawed and Inadequate Response to the Too-Big-to-Fail Problem*, 89 OR. L. REV. 951, 1005–06, 1022 (2011) ("Dodd-Frank's post-funded [Orderly Liquidation Fund] creates a strong incentive for regulators to grant forbearance in order to avoid or postpone the politically unpopular step or borrowing from the Treasury to financing a failed [systemically important financial institution]'s liquidation.").

beyond the point where it makes sense strategically, as opposed to the current system, which imposes no corresponding tradeoff in costs.

Critics may also argue that such a system does nothing to alter the perverse incentives of TBTF policies and merely institutionalizes a bailout regime. Opponents of disincentive structures and emergency liquidity funds include Treasury Secretary Tim Geithner, who instead advocates taxing financial firms only after bailouts occur, due to his concern is that any type of "bailout reserve" system would only worsen moral hazard and provide greater industry confidence in future bailouts.[171] However, this argument is misguided in assuming that the market cannot already determine which firms are TBTF and therefore pose significant systemic risk. In all likelihood, it already has.[172]

Furthermore, regardless of the public's knowledge of an institution's status as systemically important, the systemic risk regulator would retain discretion over (1) whether an institution poses sufficient risk to warrant use of the emergency liquidity pool, and (2) whether the funds will be used to replace short-term credit lines or to fund the institution's orderly wind-down (i.e., failure) under the newly established resolution authority.[173] The liquidity provider of last resort can thus maintain "constructive ambiguity" over such decisions, even if the status of systemically significant institutions remains somewhat clear to an observant market. The government's credibility is also improved to the extent there is appropriate resolution authority over failing financial firms under Title II of the Dodd-Frank Act.

Under this proposal, excessive risk-taking will be curbed in precisely the situations in which it is most important: where an institu-

171. Craig Torres & Alison Vekshin, *Bair, Bernanke Want Tougher Curbs on Biggest Banks (Update 1)*, BLOOMBERG (Jul. 15, 2009), http://www.bloomberg.com/apps/news?pid=20601087&sid=aB4OVrCHNQmE.

172. *See* David Cho, *Banks "Too Big to Fail" Have Grown Even Bigger*, WASH. POST, Aug. 28, 2009, *available at* http://www.washingtonpost.com/wp-dyn/content/article/2009/08/27/AR2009082704193.html (observing that federal bailouts only reinforced the idea that the government will save big banks, and banks have responded in kind by continuing growth and risk taking); *see also* Stevenson Jacobs, *Risk-Taking is Back for Banks 1 Year After Crisis*, ASSOCIATED PRESS, Sept. 13, 2009 (noting that the lack of systematic changes to the industry has been followed by large banks regaining their appetites for risk while investors have taken notice of implicit government backing, resulting in profitable rebounds a year after the crisis for Bank of America, Citigroup, Goldman Sachs, JPMorgan Chase, and Wells Fargo).

173. *See* Dodd-Frank Act, Title II, 124 Stat. at 1442 ("Orderly Liquidation Authority").

tion is on the borderline between "too big" and "small enough" to fail. Under the current system, a firm that believes it might be TBTF has little reason to avoid throwing itself over the top by seeking a greater concentration of financial power. In contrast, under the proposed system, a firm that is uncertain of whether it would receive emergency support from the government should be discouraged from walking this line, and must make a strategic decision to either pay the increased costs of systemic significance or to scale back certain activities that elevate such risks. As long as the costs of TBTF status clearly outweigh the benefits for most firms, those firms will be discouraged from gambling for such status. The costs of such status would not outweigh the benefits for most firms since most firms are neither systemically significant nor TBTF.

Others may argue that such disincentives will stifle the benefits of technology and innovation in the financial markets by curbing the activities of sophisticated financial firms. However, the continued dominance of TBTF institutions actually inhibits competition such that success and innovation are reduced among smaller and less interconnected firms.[174] Past a certain point, the growth of LCFIs into money center mega-banks does little to improve efficiencies, profitability, or service.[175] Furthermore, the cost of TBTF policies results in the kind of inefficient risk-bearing by the public that outweighs the benefits of innovation, particularly to the extent that increased risk-taking occurs before new financial instruments are completely understood.[176] Where new products do not serve the traditional and socially productive function of risk-spreading and instead create greater opacity in the market, the flood into the mar-

174. *See* Carter H. Golembe, *Consolidation and Competition in the Financial Services Industry*, 9 ANN. REV. BANKING L. 451, 454 (1990).

175. *See* Arthur E. Wilmarth, Jr., *Too Good to be True? The Unfulfilled Promises Behind Big Bank Mergers*, 2 STAN. J.L. BUS. & FIN. 1, 87 (1995) ("Big bank mergers have not improved the relative efficiency of profitability of large banks, and they have adversely affected competition as well as the quality and cost of service to consumers and small businesses. Perhaps the most alarming aspect of the consolidation trend is the continued predilection of our largest banks to pursue high-risk business strategies that threaten their solvency and the stability of our financial system.").

176. *See* David Nickerson & Ronnie J. Phillips, *The Federal Home Loan Bank System and the Farm Credit System: Historic Parallels and Implications for Systemic Risk*, *in* TOO BIG TO FAIL: POLICIES AND PRACTICES IN GOVERNMENT BAILOUTS, *supra* note 14, at 107, 107 ("Inefficient public risk bearing can occur whenever directed lending by public credit institutions is guaranteed by the federal government, without risk-adjusted pricing of the put option implicit in such a guarantee or the implementation of equivalent capital requirements."). R

ketplace of such instruments is to the detriment of investors and the marketplace as a whole.[177]

CONCLUSION

A good crisis should never go to waste.[178] The tension created by a patchwork system of supervisory authorities and agency turf wars led commentators to observe early on that reform of the bank regulatory structure is highly unlikely, given that an arcane and illogical framework of regulation has remained generally unchanged since the 1930s.[179] Others have remarked that true regulatory reform is unlikely to occur without an extraordinary event to propel it forward.[180] The Crisis of 2008–09 should qualify as such an event, to spur wholesale change and address the way that moral hazard has persisted and evolved since the thrift crisis before it. When even the most prudent and rigorous supervision cannot keep up with the rapid pace of financial innovation, regulatory mechanisms should be in place to inhibit the interconnectedness recently used to justify taxpayer-funded bailouts of "too big to fail" institutions and to battle the "new" old problem of moral hazard in its systemic incarnation.

177. *See* RECHTSCHAFFEN, *supra* note 41, at 164.

178. "A crisis is a terrible thing to waste" is attributed to Stanford economist Paul Romer in his comments at a November, 2004 meeting of venture capitalists. Jack Rosenthal, *A Terrible Thing to Waste*, N.Y. TIMES, Aug. 2, 2009, at MM12, *available at* http://www.nytimes.com/2009/08/02/magazine/02FOB-onlanguage-t.html. The sentiment was later shared by then-Chief of Staff Rahm Emanuel during a conference with corporate executives in 2008 following the election of President Barack Obama. Gerald F. Seib, *In Crisis, Opportunity for Barack Obama*, WALL ST. J., Nov. 21, 2008, *available at* http://online.wsj.com/article/SB122721278056345271.html.

179. *See generally*, Hackley, *supra* note 144, at 579, 580; Kenneth E. Scott, *The Dual Banking System: A Model of Competition in Regulation*, 30 STAN. L. REV. 1, 40–48 (1977)*; see also* Geoffrey P. Miller, *The Future of the Dual Banking System*, 53 BROOK. L. REV. 1, 1 (1987) (describing how the dual banking system has been a "sacred cow" of American political tradition, but is problematic due to fundamental changes in the banking industry).

180. *See* Kushmeider, *supra* note 139, at 20.