

“Bluewashing” the Firm? Voluntary Regulations, Program Design, and Member Compliance with the United Nations Global Compact

Daniel Berliner and Aseem Prakash

Voluntary programs have emerged as important instruments of public policy. We explore whether programs lacking monitoring and enforcement mechanisms can curb participants' shirking with program obligations. Incentive-based approaches to policy see monitoring and enforcement as essential to curb shirking, while norm-based approaches view social mechanisms such as norms and learning as sufficient to serve this purpose. The United Nations Global Compact (UNGC), a prominent international voluntary program, encourages firms to adopt socially responsible policies. Its program design, however, relies primarily on norms and learning to mitigate shirking. Using a panel of roughly 3,000 U.S. firms from 2000 to 2010, and multiple approaches to address endogeneity and selection issues, we examine the effects of Compact membership on members' human rights and environmental performance. We find that members fare worse than nonmembers on costly and fundamental performance dimensions, while showing improvements only in more superficial dimensions. Exploiting the lack of monitoring and enforcement, UNGC members are able to shirk: enjoying goodwill benefits of program membership without making costly changes to their human rights and environmental practices.

KEY WORDS: corporate social responsibility, voluntary programs, United Nations Global Compact, compliance

Introduction

Voluntary regulatory programs are important instruments of public policy. Much scholarly debate has focused on the factors shaping their effectiveness. We explore the role of program design in this context,¹ subjecting two competing policy approaches to empirical enquiry. Norm-based approaches to public policy suggest that program design shapes compliance by facilitating social processes of norm diffusion and learning. Actors may not comply because they lack the knowledge and the capacity to do so, or are not socially embedded in networks that emphasize compliance norms. Hence, program design must incorporate mechanisms for learning and embed actors in appropriate social networks. In contrast, incentive-based approaches view program design as shaping compliance incentives by providing

monitoring and enforcement mechanisms. In this perspective, actors may choose not to comply because compliance is costly. Monitoring and enforcement mechanisms can create countervailing incentives leading to compliance.

We explore this debate on program design and efficacy in the context of the United Nations Global Compact (UNGC), a voluntary regulatory program that seeks to improve environmental, human rights, and labor policies of participating firms. More broadly, voluntary programs have emerged as important policy tools at sectoral, national, and global levels (DeLeon & Rivera, 2009; Vogel, 2005). Firms elect to join these programs to gain benefits including reputation and goodwill (Prakash & Potoski, 2006). Firms may also face pressures from their supply chains, customers, and even regulators to join these programs (Berliner & Prakash, 2013, 2015; Cashore, Auld, & Newsom, 2004; Darnall, Jolley, & Handfield, 2008; Dashwood, 2012). While firms may join voluntary programs for such benefits, it is not clear whether they will invest the resources, or make the behavioral changes necessary to fulfill their program obligations. Such shirking undermines program efficacy and raises legitimate questions about voluntary programs as tools of public policy (Delmas & Cuerel-Burbano, 2011).

The UNGC was launched in 2000. It emphasizes 10 principles pertaining to human rights, labor standards, the environment, and corruption (as dimensions of corporate social responsibility or CSR), all which are rooted in international treaties. As an instrument of global public policy, the UNGC is a global “meta-regime,” translating international treaties signed by governments into firm-level obligations, which firms headquartered or operating anywhere in the globe should respect. To ensure compliance with program obligations, the Compact encourages information sharing and norm diffusion by embedding member firms in networks and communities, instead of providing credible monitoring and enforcement mechanisms. The Compact’s institutional design provides for minimal sanctions—expelling members only if they fail to submit a self-report on their relevant activities *for multiple years in a row*. The Compact, therefore, provides an important case to examine whether a program design relying on norm-based processes instead of incentive-based monitoring and enforcement mechanisms, can curb shirking.

Using a panel of nearly 3,000 U.S. firms over the period 2000–2010, we examine the efficacy of the Global Compact in two areas: human rights (including labor rights) and the environment. We find evidence that UNGC members shirk their program obligations, but that they do so in a strategic way. In relation to nonparticipants, UNGC members perform worse on crucial environmental and human rights dimensions that are costly to change. Yet, in more superficial areas of human rights and environment performance, these members take positive but low-cost steps. This sort of strategic shirking (Delmas & Keller, 2005; DeVaro & Gürtler, 2012) is consistent with the predictions of the incentive-based approach. Global Compact members can take advantage of reputational and goodwill benefits emanating from program membership, especially because the program is sponsored by the UN. Yet its program design creates opportunities for “bluewashing,” wherein members pay lip service to the true goals of CSR instead of undertaking substantive but costly changes in their environmental and human rights performance.

This paper also offers one of the first empirical assessments of the impact of Global Compact membership on human rights and environmental performance that takes into account the complex issues of nonrandom selection into membership, using a combination of differences-in-differences, instrumental variables, and propensity score matching approaches. Previous evaluations of the Global Compact have tended to neglect these factors, thereby risking overly optimistic conclusions as to the efficacy of Global Compact membership.

The paper proceeds as follows. The second section outlines the theory behind voluntary regulatory programs, and the incentive-based and norm-based approaches linking program design to compliance. In the subsequent sections, we discuss the origins and design of the Global Compact, the data and the modeling strategy used in our analysis, our results, and a conclusion.

Institutional Design and Voluntary Regulation

Voluntary programs encourage firms to incur the costs of adopting policies that are aimed at benefiting a diverse range of stakeholders, instead of narrowly focusing on maximizing profits for the benefits of shareholders alone. Arguably, public law and regulation should be designed in ways to encourage firms to pay attention to the needs of a range of stakeholders. While public regulations provide the basic governance framework in most of the world, on their own, some believe that they need to be supplemented with private regulatory efforts. Changing political conditions have made policymakers sensitive to the backlash against public regulation's alleged heavy-handedness. There is also a sense that increasing heterogeneity among firms, the complexity of modern value addition processes, and the poor quality of regulatory infrastructure in developing countries have all made new public regulations harder to enact and enforce (Fiorino, 2006). Practitioners and scholars suggest that instead of treating businesses only as the source of social and environmental ills, policy efforts should look to mobilize their cooperation with positive incentives, supplementing public law with voluntary programs (Börzel & Risse, 2010; Coglianesi & Nash, 2001).

Voluntary programs are among the new policy instruments aimed at creating a new social compact, a new "embedded liberalism" (Ruggie, 1982) that embeds capitalism with redistribution via private means. Such institutions create incentives for firms to produce public goods beyond the requirements of applicable laws. In return, firms get a credible, low-cost way to signal their stewardship. The assumption is that stakeholders will compensate firms for such "beyond compliance" stewardship by bestowing benefits including goodwill, regulatory relief, higher market shares, customer loyalty, and higher product prices (Lundgren, 2003). In this way, voluntary programs compensate participants with an excludable "club good," in that only firms participating in the program can leverage excludable branding or signaling benefits (Borck & Coglianesi, 2009; Prakash & Potoski, 2006). Program branding allows external stakeholders to sort participants from nonparticipants (Spence, 1973) and target their behavior accordingly (King, 2007).

In effect, voluntary programs can create a new market for corporate virtue (Vogel, 2005), providing firms with a *de facto* “social license to operate” (Gunningham, Kagan, & Thornton, 2004).

Despite their policy promise, there remains considerable skepticism about voluntary programs. Critics believe that regulation of firms to harmonize market pressures and societal demands is a task best left to the state. Given the structural imperatives of profit, such critics do not trust firms to place the concerns of other stakeholders on par with those of their shareholders. At the programmatic level, given the low entry barriers in establishing voluntary programs, there is considerable heterogeneity among such initiatives regarding the types and stringency of obligations they impose on their participants, and mechanisms they incorporate to ensure that those obligations are fulfilled. Thus, it is important to examine voluntary program efficacy empirically, a task we undertake in this paper for one prominent global program.

Empirical research to date has shown the efficacy of voluntary programs to be uneven (Darnall & Kim, 2012; DeLeon & Rivera, 2009; Morgenstern & Pizer, 2007).² Take the case of voluntary environmental programs. Firms participating in the chemical industry’s Responsible Care program (King & Lenox, 2000) and the U.S. Department of Energy’s Climate Wise program (Welch, Mazur, & Bretschneider, 2000) appear to have done little to protect the environment beyond what they would have done had they not joined those programs. Conversely, firms that joined the Environmental Protection Agency’s (EPA) 35/50 voluntary program reduced their emissions of toxic pollutants more than those that did not (Khanna & Damon, 1999). While Potoski and Prakash (2005) find that ISO 14001 adoption improved U.S. facilities’ compliance with the Clean Air Act, and Dasgupta, Hettige, and Wheeler (2000) find that ISO 14001 adoption improved Mexican facilities’ self-reported compliance with public law, Dahlström, Howes, Leinster, and Skea (2003) report that ISO 14001 and EMAS certification did not improve regulatory performance of British facilities. Similarly, while Russo (2002) finds that joining ISO 14001 reduced firms’ pollution emissions, Andrews et al. (2003) provides some evidence that ISO 14001 did not affect firms’ environmental performance.

Given this mixed track record, how might we explain the varying efficacy of voluntary programs? We identify two broad theoretical approaches applicable to program efficacy. Incentive-based approaches expect that, on average, members participating in these programs have incentives to shirk on their program obligations; that is, they might seek to enjoy the reputational and goodwill benefits of program membership while taking no costly actions to act upon their program obligations. Anticipating such free riding, they suggest altering participants’ incentive structures by explicitly incorporating program design mechanisms to detect and punish non-compliance (Ostrom, 1990).

In contrast, norm-based approaches emphasize the ability of social mechanisms such as norms, persuasion, and learning to induce members to comply with program obligations. For them, members’ compliance is shaped by network embeddedness and their actions are informed by the logic of appropriateness. March and Olsen (2006) put it as follows: “Rules are followed because they are seen as natural,

rightful, expected, and legitimate." While incentive-based approach attributes non-compliance to willful shirking, the norm-based approach attributes noncompliance to a lack of socialization into appropriate norms, and ignorance of best practices.³

Indeed, there is some work on this debate in the context of voluntary environmental programs. King and Lenox (2000) and more recently, Gamper-Rabindran and Finger (2013), attribute the inefficacy of the Chemical Industry's Responsible Care CSR program to its design, which did not provide adequate monitoring and sanctioning. Participants shirked their program responsibilities despite normative, mimetic, and coercive pressures from within the Responsible Care and chemical manufacturing communities. In 2005, Responsible Care modified its institutional design and introduced third-party monitoring to curb shirking. In a recent paper, Vidovic, Khanna, and Delgado (2013) report that once the monitoring mechanisms were introduced, participants' behaviors changed significantly: Participants began to pollute less than nonparticipants. The importance of third-party monitoring in curbing shirking is also highlighted by Potoski and Prakash (2005) who report that ISO 14001, a voluntary program that imposes obligations that are similar to the ones imposed by Responsible Care but provides for enforcement via third-party audits, leads to improvements in participating firms' environmental performance.

We investigate how the relative lack of monitoring and sanctioning in the context of the UNGC might influence participants' human rights and environmental performance. Advocates of the Compact explicitly tout its norm-based, as opposed to enforcement-based, mechanisms. On this count, it becomes an important case in which to test the efficacy of norm-based approaches to public policy. If embeddedness in learning and social networks can curb shirking, the lack of monitoring and enforcement mechanisms should not significantly influence the Compact's efficacy. However, a finding that program efficacy is compromised should raise new questions about the Compact's program design.

Institutional Design of the UNGC

The UNGC launched in July 2000, after being first announced by Kofi Annan at the World Economic Forum in 1999. The Compact was founded with nine principles in the issue areas of human rights, labor standards, and the environment with anti-corruption added as a tenth principle in 2004.⁴ The principles are based on four UN documents that enjoy broad international support: the Universal Declaration of Human Rights, Declaration on Fundamental Principles and Rights at Work, Rio Declaration on Environment and Development, and United Nations Convention Against Corruption. In each of these issue areas, the Compact attempts to address Pigouvian negative externalities, but with imprecise targets rather than mandates.

Instead of providing stringent mechanisms to monitor and enforce compliance with program obligations, the UNGC explicitly relies on norm-based principles of institutional design to achieve the same purpose. John Ruggie, the intellectual force behind the Compact, and Georg Kell, current Executive Director of the Global Compact Office, noted that the Global Compact is "meant to serve as a framework of

reference and dialogue to stimulate best practices and to bring about convergence in corporate practices around universally shared values" (Ruggie & Kell, 1999, p. 5). Ruggie (2002, pp. 31–32) explained that:

The major criticism of the GC by the anti-globalisation front has been for what it is not: a regulatory arrangement, specifically a legally binding code of conduct with explicit performance criteria and independent monitoring of company compliance. So how does the GC propose to induce corporate change? Its core is a learning forum. Companies submit case studies of what they have done to translate their commitment to the GC principles into concrete corporate practices. This occasions a dialogue among GC participants from all sectors: the UN, business, labour and civil society organisations. The aim of this dialogue is to reach broader, consensus-based definitions of what constitutes good practices than any of the parties could achieve alone. Those definitions, together with illustrative case studies, are then publicised in an online information bank, which will become a standard reference source on corporate social responsibility. The hope and expectation is that good practices will help to drive out bad ones through the power of dialogue, transparency, advocacy and competition.

Georg Kell has emphasized similar points. In a 2010 interview with the Inter Press Service, he was asked "Has the Global Compact blacklisted or expelled any companies or corporations for violations of ethics or accused of malpractices?" He responded that:

The Global Compact is a platform for dialogue, learning and partnership. Participation in the Global Compact does not imply perfection. It simply means that an organisation is willing to align with U.N. principles and engage in activities that advance U.N. goals. As such, GC does not make judgments. We know this has at times caused misunderstandings or invited criticism, but we have always kept the GC's entry barrier intentionally low, so that those that face serious challenges can join the conversation, learn from others and improve. And we stand by this approach, as long as we can see that there is a sincere commitment to transparency and public accountability.⁵

Some authors explicitly argue that the Global Compact offers evidence *in favor* of the power of norm-based mechanisms to lead to compliance (although they offer only minimal empirical evidence for these claims). Rasche, Waddock, and McIntosh (2013, pp. 10–11), for example, argue that "the initiative's deliberative capacity also enhances compliance with its underlying rules," that its "focus on learning and arguing also affects compliance by gradually 'socializing' actors into new rules," and that "such decentralized deliberations strengthen the willingness and capacity of actors to voluntarily comply with rules," concluding that "although the Global Compact is not a regulatory tool in the narrow sense, it can still unfold significant effects on business corporate responsibility practices." Mwangi, Rieth, and Schmitz (2013) note that "the Compact offers a compelling test case for the significance of

discourses, learning, and capacity-building as pathways to compliance" (pp. 203–4). They argue that "the existence of active regional and local GC networks is a crucial ingredient for setting in motion specific mechanisms, such as peer learning and capacity-building, that can contribute effectively to improved performance of individual member companies" (Mwangi et al., 2013, p. 204). Indeed, the United States has such an active local network, so their approach would expect Compact membership among U.S. firms to contribute to improved performance.

However, civil society critics have questioned claims that the Compact shapes members' human rights and environmental performance, or that its program design can create incentives toward this outcome.⁶ Such critics often accuse the Global Compact of "bluewashing," whereby member firms figuratively drape themselves in the blue UN flag in order to burnish their reputations and distract stakeholders from their poor environmental or human rights records.⁷ For example, a 2000 critique noted that "there will be no mechanism to make adherence to the Compact's principles binding in any way. That is how the International Chamber of Commerce wants it."⁸ In 2007, Daniel Mittler of Greenpeace argued that "instead of organizing expensive summit meetings, the UN must ultimately set internationally-binding CSR standards for corporate behaviour and see to their adherence. The world does not need more declarations of intent from corporations, but real actions that can be measured and monitored."⁹ These criticisms tend to focus on issues of the Compact's program design, claiming that weak monitoring and sanctioning mechanisms allow firms to gain reputational benefits while taking few meaningful actions and suffering little cost from shirking.

For example, the Compact does a poor job of sorting socially responsible firms from laggards, as reflected in its modest membership requirements. The Compact only requires new members to prepare a Letter of Commitment from their CEO expressing support for the 10 principles. Members are also required to submit annually a Communications on Progress, but the Compact does not verify the claims made in these reports, either by itself or through third-party monitors. Instead, it relies on members of the public or issue activists to highlight cases of poor performance or disingenuous reporting. The Compact provides for a complaint system for these stakeholders to report "systematic and egregious abuses" by Compact participants. However, this system lacks transparency and has been criticized as ineffective (EthicalCorp.com, 2008). Notably, the Global Compact Web site does not include any such complaints, or the results of complaint procedures, in its participant listings (Sethi & Schepers, 2014).

The Communication on Progress (COP) is a disclosure of progress made in implementing the Compact's 10 principles, and is shared with stakeholders. The only specific elements required in the COP are a statement of continued support for the Compact, a description of practical actions taken on at least two issue areas (changes in 2009 required that all four issue areas must be addressed by five years of participation), and any measurement of outcomes. The extremely lax deadlines for filing these voluntary self-reports allow firms to fail to submit for a full four years (two years after 2009 policy changes) after joining, before finally being sanctioned by delisting from the Compact.

Additionally, the quality of substantive information provided in the COP reports is poor. The Compact's 2008 Annual Review noted the results of a survey of 40 large companies indicating that "not all Global Compact principles are covered with the same level of detail," that "there is a wide disparity with regard to information available per principle," and that "reported information is not comprehensive, with COPs focusing more on commitments and management systems than on materiality, performance and achievements" (UN Global Compact, 2009). Given that the COP is the only long-term requirement firms must meet to maintain membership, the lax timelines and poor substantive quality of these self-reports reflect poorly on program credibility.

Just as the Global Compact leadership's arguments are grounded in norm-based approaches, criticisms from civil society groups or other critics are implicitly, if not explicitly, grounded in political economy approaches. Empirically testing between these different points of view is, therefore, of great importance both theoretically and practically.

Some scholars have also sought to evaluate the efficacy of Global Compact membership. Bernhagen and Mitchell (2010) analyzed the world's largest two thousand companies, comparing members and nonmembers. They concluded that Compact members are more likely to have internal human rights policies, and are more likely to be listed on "Innovest's 'Global 100' list of the 'most sustainable' corporations" (Bernhagen & Mitchell, 2010, p. 1181). Mwangi et al. (2013, p. 215) analyzed the 20 largest automotive and utility companies in the world, concluding that Compact members were more likely than nonmembers to file sustainability reports using the Global Reporting Initiative's reporting standards, which they interpret as a move "from an initial rhetorical commitment to the GC toward a more meaningful implementation of reporting about progress." On the other hand, Hamann, Sinha, Kapfudzaruwa, and Schild (2009) analyzed the human rights due diligence practices of the top 100 firms listed on the Johannesburg Stock Exchange, and found no effect for Compact membership. However, in addition to not providing a direct measure of human rights or environmental performance, none of these analyses take into account the potential for nonrandom selection of firms into the Compact to drive their results, potentially leading to overly optimistic conclusions.

Sethi and Schepers (2014, p. 1), instead of empirically evaluating data on compliance, conduct an analysis of the Compact's governance structure, sources of financial support, and procedures for reporting and delisting, concluding that "the UNGC has failed to induce its signatory companies to enhance their efforts and integrate the 10 principles in their policies and operations." We seek to build on these existing studies by using a more direct empirical measure of the actual corporate social performance of firms, and by systematically addressing the selection issues that potentially bias past quantitative work.

Data and Model

Our dependent variable is firms' environmental and human rights performance, using data from MSCI ESG Statistics, formerly known as the KLD Social and Envi-

Table 1. Environmental and Human Rights Strengths and Concerns from the KLD Ratings

Environmental Strengths	Environmental Concerns
Beneficial products and services	Hazardous waste
Pollution prevention	Regulatory problems
Recycling	Ozone depleting chemicals
Clean energy	Substantial emissions
Other strengths	Agricultural chemicals
	Climate change
	Other concerns
Human Rights Strengths	Human Rights Concerns
Indigenous peoples relations strengths	Operations in Burma
Labor rights strength	Labor rights concern
Human rights policies and initiatives	Indigenous peoples relations concerns
Other strengths	Other concerns

ronmental Performance ratings (henceforth KLD ratings). These ratings have been called the “the largest multidimensional CSP [corporate social performance] database available to the public” (Deckop, Merriman, & Gupta, 2006, p. 334) and the “the most exhaustive and widely used measure of CSR” (Vogel, 2005). The ratings have been used in numerous studies pertaining to the CSR of firms on topics including the effects of membership in the organization Business for Social Responsibility (Tashman & Rivera, 2010), the effects of anti-sweatshop campaigns on apparel firms (Bartley & Child, 2011), and the relationships between CSR practices and firm performance (Doh, Howton, Howton, & Siegel, 2010; Waddock & Graves, 1997).

The ratings include measures of firm performance on indicators grouped into several categories, including Community, Diversity, Corporate Governance, Gambling, and Nuclear Power. We, however, are interested in the two categories that concern elements of the Global Compact’s 10 principles: Environment and Human Rights.¹⁰ In each category, the individual indicators are grouped into Strengths and Concerns. Table 1 displays the strengths and concerns indicators, which we include for each the Environment and Human Rights categories. We exclude categories that were dropped from the ratings during our period under study, or that were only initiated at the end of the period. Each individual indicator is coded as a 0 or 1 in each year.

Do firms which join the UNGC “walk” the CSR “talk”? The KLD data provides valuable insight on this issue. We interpret the strengths indicators as reflecting relatively superficial efforts that a firm can undertake cheaply, without substantially changing their primary activities or business models. In contrast, concerns reflect the actual changes in structural issues that bear upon the core activities of a firm. This makes intuitive sense if one closely looks at the demands placed on the firm via these strengths and concerns.¹¹

For example, the environmental strengths categories relate primarily to the existence of programs and policies in place to mitigate pollution, reduce emissions, and increase the use of recycled materials and clean energy. Notably, firms can receive points in these categories even if the programs are not particularly effective.

The definition of the “Clean Energy” category even includes the simple “acknowledgement of direct and/or indirect impacts on operations due to climate change.” Meanwhile, the environmental concerns categories more directly address actual levels of pollution and emissions, and the production of environmentally harmful goods such as chemical pesticides. It is much less costly for a firm to create recycling and pollution mitigation programs, which may or may not be effective, than it is to actually address problems of hazardous waste or substantial emissions, listed among the environmental concerns in the KLD ratings. Confirming this interpretation, Chatterji, Levine, and Toffel (2009) find that in the case of KLD’s environmental ratings, the “concerns” indicators are significantly associated with past and future pollution levels and regulatory compliance levels, whereas the “strength” indicators show no such association with objective measures of environmental performance.

Similarly, in the area of labor rights, the “strengths” indicator is defined as “the company has outstanding transparency on overseas sourcing disclosure and monitoring, or has particularly good union relations outside the U.S., or has undertaken labor rights–related initiatives that KLD considers outstanding or innovative,” whereas the “concerns” indicator is defined as “the company’s operations have had major recent controversies primarily related to labor standards in its supply chain.” Once again, it is far less costly to enact monitoring polities and join transnational initiatives than it is to actually address persistent labor rights abuses that take place at the bottom of global supply chains (Locke, 2013).

In sum, the “concerns” indicators reflect costly actions firms have to undertake to improve performance, whereas the “strengths” indicators reflect more superficial, less costly activities. It is far simpler, and less costly, for a firm to take actions that raise its “strength” indicators than to take actions that lower its “concerns” indicators. Given that these categories reflect different types of actions, we treat them separately rather than combining them into an aggregated measure. Indeed, in a study of latent constructs underlying the KLD ratings, Mattingly and Berman (2006) concluded the strengths and concerns were conceptually distinct, and that it was inappropriate to combine them in such a manner.

Thus, we use two separate dependent variables. For each firm in each year, we record the sum of the strengths indicators, and the sum of the concerns indicators, as our two primary dependent variables.¹² Since we do employ some categories that were introduced in the middle of the period under observation, following Tashman and Rivera (2010), we standardize the Strengths and Concerns variables for each year, calculating z-scores using the annual means and standard deviations across all firms for each measure. In some models, we instead treat the two variables as counts, using the raw measures instead of the standardized versions.

The sample of firms covered by the KLD rating has increased over time. As of 2000, it covered firms in the S&P 500 Index and firms in the Domini 400 Social Index. In 2001, it was extended to cover the thousand largest publicly traded firms in the United States. In 2002, firms in the Large Cap Social Index were added. In 2003, coverage expanded to include an additional two thousand publicly traded firms and firms listed in the Broad Market Social Index.¹³ To maximize our data coverage, we first used all firms that we were able to match between the KLD data and additional

variables from the Thomson Reuters Datastream database.¹⁴ To ensure sufficient coverage of individual firms over time, we omitted any firms observed for three or fewer years. Our sample includes the remaining 2,950 firms. Since some firms are acquired or go bankrupt, we have an unbalanced panel of 21,738 firm-year observations from 2000 to 2010. In some additional models, however, we also show results using only firms present in the panel from the year 2000, or using samples matched on observable covariates.

Our key independent variable is UNGC Membership, an indicator of years in which a firm was a member of the Compact. We do not count a firm as a member in the year in which it joined, to avoid attributing to member firms environmental and human rights performance that took place prior to joining—that is, our membership variable is lagged by one year. Of the 2,950 firms in our sample, 42 firms are Global Compact members, having joined by the end of 2009. While just over one hundred U.S. firms in total had joined the Compact by that date, many of these were not publicly listed and so not eligible for inclusion in the KLD ratings.¹⁵ Following the CSR literature, we also include control variables for the size of the firm, measured as logged total sales, and the firm's profits in each year. In some models that do not include firm fixed effects, we instead include dummy variables for industrial sectors, at the level of Standard Industrial Classification (SIC) divisions.¹⁶

Modeling Approaches

Any modeling strategy to determine the effects of membership in a voluntary initiative such as the Global Compact must confront the fact that membership itself is nonrandom. Indeed, selection into the program is likely to be driven by omitted variables, if not past environmental and human rights performance themselves. It is precisely this potential selection problem that past empirical work on Global Compact efficacy has failed to address. Simply comparing the performance of member firms with nonmember firms, at a single point in time, risks mistaking a selection process for a causal effect.

We use three different methods to address selection issues: (i) a differences-in-differences design, (ii) instrumental variables, and (iii) propensity score matching. Differences-in-differences models compare units before and after they received a treatment with other units that received no treatment but were measured at comparable time periods. Such methods have been used in research at both the country level (Djankov, McLiesh, & Shleifer, 2007; Slaughter, 2001) and the firm level (Bertrand & Mullainathan, 2003; MacGarvie, 2006). In our case, the differences-in-differences approach compares the changes in ratings of firms that joined the Compact before and after joining, with the changes in ratings of other firms that did not join, over the same time period. Such a design has many beneficial qualities. By including firm fixed effects, it ensures that any firm-specific omitted variables, such as varying baseline propensities to join the Compact, do not bias the model estimates. By including year fixed effects, it ensures that across-the-board changes in corporate responsibility over time are not mistaken for effects of the Global Compact. For the

model to identify an effect of Compact membership, any change in average ratings over time for member firms must be greater than the changes undergone by non-member firms over the same time period.

However, such a model may not fully take into account selection effects, if firms that join the Global Compact have independent tendencies toward different trends over time in their social performance than firms that do not join. We, therefore, supplement the differences-in-differences design with two additional approaches: instrumental variables and propensity score matching. We first use two instruments to assess the effects of only the exogenous element of Global Compact membership. To be a valid instrument, a variable must be associated with the potentially endogenous independent variable, but must remain orthogonal to the error term of the final outcome equation. That is, a potential instrument must predict Global Compact membership among U.S. firms, but must not influence their corporate social performance through any other causal mechanism.

We identify two instruments that fulfill these conditions both theoretically and empirically. The first instrument is the proportion of Global Compact member firms in each firm's sector in the previous year, within the observed sample of U.S. publicly traded firms.¹⁷ The second instrument is the total number of Global Compact member firms in each firm's sector in the previous year, in a selection of seven other countries: Canada, China, France, Germany, Japan, Mexico, and the United Kingdom.¹⁸ The intuition behind both instruments is that firms have a tendency to mimic the behavior of salient peer groups, especially with regard to decisions to join voluntary initiatives. This draws on the insights of institutional isomorphism (DiMaggio & Powell, 1983), and the focus on emulation in the policy diffusion literature (Berliner, 2013, 2014; Greenhill, 2010; Simmons, Dobbin, & Garrett, 2006). Mimicry of adoption in other countries has been identified as an important factor by research on firm decisions to join the Global Compact (Berliner & Prakash, 2012; Lim & Tsutsui, 2012; Perez-Batres, Miller, & Pisani, 2011; Perkins & Neumayer, 2010) as well as other voluntary programs such as ISO 9000 (Guler, Guillén, & Macpherson, 2002; Prakash & Potoski, 2007). When more firms in the same sector have joined the Compact, both in the United States and in other countries that are economically or geographically proximate to the United States, a given firm is likely to face greater mimetic pressure to do the same, independent of the financial cost-benefit calculus that might also play into its decision whether or not to join.¹⁹ And since the two instrumental variables pertain only to the actions of firms *other than* the reference firm, are lagged temporally and/or spatially, and do not measure the actual corporate social performance of any firm, there is no plausible causal mechanism by which they might influence the reference firm's actual corporate social performance, except for influencing its likelihood of joining the Compact.

Indeed, the validity of the instruments is also supported empirically, as the models that include them pass tests for both instrument strength and exogeneity. The Cragg-Donald F statistic of the first-stage model is far above the most restrictive critical value of a Stock-Yogo weak identification test, and the *p*-values of the Hansen J statistics of the second-stage models both fail to reject the null hypothesis that the over-identification restrictions are valid.

The second additional approach we use is to supplement the differences-in-differences design with propensity score matching. Previous studies have also used matching as a way to address potentially endogenous firm behaviors (Girma, Greenaway, & Kneller, 2004; Levine & Toffel, 2010). This method matches each firm that joined the Global Compact with another firm that did not join, yet is as similar as possible on observed covariates. While propensity score matching is no panacea for potential endogeneity issues, it provides greater assurance that results of the differences-in-differences approach are not the result of comparing firms that are very different from each other in terms of firm size, profits, and industrial sector.

Finally, an additional potential concern arises from our use of ratings compiled by CSR analysts as the dependent variable. It is possible that the results are biased by measurement problems, if member firms are more likely than others to receive the scrutiny (either from KLD analysts themselves or from the various data sources they rely on) that leads to them receiving nonzero ratings on the various indicators (either strengths or concerns). Indeed, 90 percent of firms in our sample received scores of zero on all of the strengths we consider, while 83 percent received scores of zero on all of the concerns we consider. As such, we employ a zero-inflated Poisson model, which assumes that these zero observations are of two types, arising from two separate data-generating processes: regular zeroes, and "structural" or "inflated" zeroes. Here, the regular zeroes are those firms that received adequate scrutiny but nonetheless were scored zero, while the inflated zeroes were insufficiently scrutinized. By simultaneously estimating a binary model of such "inflation" and a count model of the final outcomes, the zero-inflated Poisson model can estimate whether Compact members are less likely than other firms to receive insufficient scrutiny (i.e., more likely to receive adequate scrutiny), as well as estimating the effects of membership once any potential inflation process is taken into account. However, due to the complexity of this model, it is only possible to include sectoral fixed effects, not firm fixed effects, and so the results will not take into account potential omitted variables at the firm level.

Results

The basic results of a differences-in-differences approach to the data are presented in Figure 1. This figure does not present model results, but rather simply a means of visualizing the intuition of the differences-in-differences approach. For both Strengths and Concerns (using unstandardized data), this figure shows the average ratings among firms that joined the Global Compact over a period from three years before joining to three years after. These are compared with aggregated average ratings of firms that did not join the Global Compact over the same periods. The figure highlights several differences between firms that joined and did not join. First of all, even before joining the Global Compact, firms that would go on to join tended to be different from firms that would not go on to join, characterized both by greater strengths and greater concerns. The fact that firms with greater concerns were

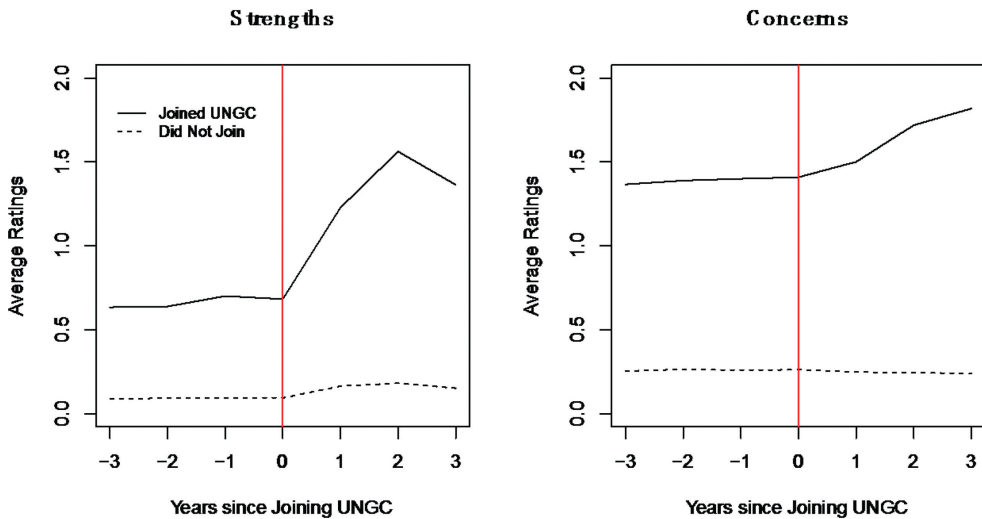


Figure 1. Illustration of Differences-in-Differences Approach.

The nondashed lines show average ratings among firms that joined the Global Compact, over a period from three years before joining to three years after. The dashed lines reflect aggregated average ratings of firms that did not join the Global Compact over the same periods.

more likely to join the Compact than other firms implies that joining is a strategic act intended to ameliorate reputational problems arising from criticism over performance in these areas. However, after joining the Global Compact, firms tended to undergo changes in their ratings that cannot simply be attributed to across-the-board trends experienced by all firms. Both Strengths ratings and Concerns ratings increased for member firms, compared both with their ratings prior to joining and to the ratings of nonmember firms over the same period.

Table 2 presents the results of our empirical models. Since we employ many different combinations of model, sample, dependent variable, different fixed effects, all oriented around the results of one specific variable, we present all the results in a single table. Each row presents the coefficients and associated standard errors for UNGC membership in two separate models, one with the dependent variable for Strengths and the other with the dependent variable for Concerns, but otherwise with identical modeling choices. All models also contain control variables for Firm Size and Firm Profits, but the results for these are not of primary interest and will be made available in Table S1 online. Unless otherwise noted, all the models labeled as linear use the standardized dependent variables, while the Poisson and zero-inflated models use unstandardized count dependent variables.

Row 3 in this table presents the results of differences-in-differences models of each outcome, while the previous two rows show comparison models including year fixed effects only, and both year and sector fixed effects. In all cases, the effects of Compact membership are positive and significant for both Strengths *and* Concerns. However, the inclusion of firm fixed effects in the differences-in-differences model, which control for any omitted variables at the firm level, yields coefficients smaller

Table 2. Main Results

Model	Fixed Effects	Strengths: UNGC Coef.	Concerns: UNGC Coef.	N Obs.	N Firms
<i>Linear results</i>					
1. Linear	Year	2.425** (0.460)	1.451** (0.485)	21738	2950
2. Linear	Year, sector	2.309** (0.442)	1.341** (0.458)	21738	2950
3. Linear	Year, firm	1.452** (0.267)	0.713** (0.174)	21738	2950
4. Instrumental variables ^{ab}	Year, firm	5.083** (1.564)	3.530** (1.292)	21736	2948
<i>Count results</i>					
5. Poisson	Year	0.981** (0.159)	0.579** (0.215)	21738	2950
6. Poisson	Year, sector	0.684** (0.163)	0.389** (0.168)	21738	2950
7. Poisson ^c	Year, firm	0.143 (0.177)	0.169* (0.0879)	6717/6223	729/725
8. Zero-inflated Poisson ^d	Year, sector	0.559** (0.134)	0.310* (0.167)	21738	2950
<i>Alternate samples</i>					
9. Linear (firms present in 2000)	Year, firm	1.465** (0.342)	0.716** (0.225)	5130	522
10. Linear (matched firms)	Year, firm	0.958** (0.273)	0.492** (0.182)	2174	246
11. Linear (matched by sector)	Year, firm	0.888** (0.269)	0.363** (0.183)	2183	246
<i>Alternate DVs</i>					
12. Linear (unstandardized DV)	Year, firm	1.132** (0.132)	0.530** (0.130)	21738	2950
13. Linear (Env. categories)	Year, firm	1.214** (0.270)	0.463** (0.174)	21738	2950
14. Linear (H.R. categories)	Year, firm	1.522** (0.646)	1.023** (0.278)	21738	2950

Note: All models control for Firm Size and Profits (see Online Appendix for full results). Each row presents the coefficient for UN Global Compact membership from two different models, one with strengths as the DV and the other with concerns as the DV.

* $p < 0.10$, ** $p < 0.05$; Standard errors clustered by firm in parentheses.

^aInstruments: The proportion of firms in the sample from the same sector (SIC 2 digit level) in the previous year, and the number of firms in the same sector (SIC division level) in the previous year in Canada, China, France, Germany, Japan, Mexico, and the United Kingdom. First-stage equations also includes Firm Size, Profits, and year dummies.

^bCragg-Donald F statistics: Strengths: 139.6; Concerns: 139.6. Hansen J statistics (p -values): Strengths: 0.76; Concerns: 0.857.

^cReduced, and different sample sizes due to omission of firms with zero values on the DV for full panel.

^dUNGC coefficients in Inflation equations: Strengths: -28.71^* (1.371); Concerns: -0.670 (0.887).

in magnitude than those in the previous comparison models. Row 4 presents the results of instrumental variables models, using membership among firms in the same sector in the United States and in seven other countries as instruments. In these models, isolating the exogenous component of Compact membership while still using a differences-in-differences design, the effects on environmental and human rights strengths and concerns are shown to both be of even larger magnitude than in the previous sets of models.

The next sets of models treat the dependent variables as counts, relying on Poisson and zero-inflated Poisson models. While the models with year fixed effects only, and with year and sector fixed effects all show positive and significant effects of Compact membership on both outcomes, the results in Row 7 show different results. The results of the differences-in-differences Poisson models show that while Compact membership has a positive and significant effect on Concerns, there is no significant effect on Strengths. These are even more negative results from the point of view of the Compact's efficacy than the results of the previous models.

Row 8 presents the results of the second-stage equations of zero-inflated Poisson models. While it is not computationally feasible to include firm fixed effects in these models, the inclusion of sectoral fixed effects at least controls for any unobserved omitted variables at the sector level. The results of the inflation equations yield coefficients for Global Compact membership that are negative and statistically significant for Strengths, but not statistically significant for Concerns.²⁰ That is, Compact members are *less* likely to receive *insufficient* scrutiny of their environmental and human rights strengths than other firms, or in other words are more likely to receive adequate scrutiny. However, there is no difference in scrutiny of environmental and human rights concerns between Compact members and nonmembers. Once taking these first-stage equations into account, however, the results show positive and statistically significant coefficients for Compact membership for both Strengths and Concerns. That is, even once taking into account potential measurement problems relating to greater scrutiny for members, the primary results remain.

Rows 9, 10, and 11 in Table 2 present the results of models using alternative samples of firms. The models in Row 9 include only those firms present in the sample in the year 2000, ensuring that the results are not driven by the changing sample size as the KLD ratings expanded their coverage. Rows 10 and 11 use two different strategies based on propensity score matching. As noted previously, differences-in-differences procedures may not fully take into account all potential forms of selection bias. Specifically, if firms that select into joining the Global Compact have tendencies to differentially trend over time in the outcome variables in relation to firms that do not join, the estimates from the differences-in-differences models could still be artifacts of selection. While the models using instrumental variables offer one way to counter this possibility, we also employ an alternative based on propensity score matching.

Propensity score matching procedures match each observation that received a "treatment" (joining the Global Compact in our case) with another observation (or multiple observations) that did not, yet is as similar as possible as the treated obser-

vation on the other independent variables. In an ideal situation, the two observations differ in one dimension only: one is treated and the other is not. In our case, by matching each firm that joined the Compact with another similar firm that did not join, we can minimize the possibility that firms that join will tend to have systematically different trends over time in their performance. However, as with all matching procedures, we are limited to matching on observables. We use nearest-neighbor matching to match firms on several salient predictors of CSR performance, namely firm size, profits, and industrial sector. Given that each firm is observed for multiple years, we match firms that join with firms that never join, rather than matching individual firm-years. Thus, we match firms based on their sales and profits in the first year of each firms' panel.

We employ two alternate matching procedures, the first based solely on firm size and profits, and the second taking into account firms' sector. In the latter approach, we match firms within sectors, effectively stratifying the matching procedure by sector, using SIC divisions. Both matching procedures employ a ratio of five, matching each firm that joined the Compact with five others. Since there are so many firms available for the matching procedure, matching with five other firms instead of just one avoids the possibility that idiosyncrasies of individual particular matches may drive the results. Each matched dataset contains 246 firms. The results for differences-in-differences models using these matched datasets still show positive and statistically significant effects of Global Compact membership on both Strengths and Concerns.

Finally, Rows 12, 13, and 14 in Table 2 present results using alternative dependent variables in linear differences-in-differences models. Row 12 presents models using the unstandardized versions of the Strengths and Concerns variables. Row 13 presents models using only the environmental strengths and environmental concerns, while Row 14 uses only human rights strengths and human rights concerns. In all sets of models, the results still show positive and significant effects of Compact membership on both outcome variables.

In sum, across all of these different sets of results, the conclusions for the efficacy of UNGC membership are negative. At best, firms that joined tended to have both more corporate responsibility positives and more negatives to begin with, and the effect of Compact membership appears to have been to make both of these characteristics more extreme—more of the good along with more of the bad. This finding was robust to numerous different sets of choices regarding the modeling approach, addressing both selection and measurement problems. At worst, in the differences-in-differences Poisson model, the results show that membership still leads to greater concerns without leading to any greater strengths at all.

Given that the environmental and human rights strengths included in the KLD ratings reflect low-cost, more superficial actions that firms can take, whereas the concerns included in the ratings are costly for firms to ameliorate, we interpret the concerns ratings as reflecting more genuine efforts at compliance with the 10 principles of the Global Compact. The fact that these concerns did not improve but instead tended to become worse for member firms—a finding consistent across every single model in our analysis—implies that firms are shirking on their obliga-

tions, taking advantage of the reputational benefits of membership while not undertaking serious efforts toward compliance. If norm-based mechanisms are at work, we find evidence for their role only in encouraging more superficial, low-cost actions toward environmental and human rights goals, like creating recycling programs or adopting written human rights policies.

It is important to note that this study only covered large, publicly traded, U.S.-based firms, and thus its inferences cannot be assumed to extend to smaller Global Compact members, or those in other countries. Nonetheless, it is precisely large, publicly traded, U.S.-based firms that tend to play dominant roles in both global supply chains and in the public discourse over CSR issues. No CSR program that intends to be global in scope and impact can afford to leave such members behind.

Conclusion

Collective action is a core issue in the study of public policy. Since Olson's (1965) seminal contribution, scholars have examined mechanisms that curb free riding and induce actors to participate in collective endeavors. In addition to the Olsonian concerns of free riding associated with recruitment into a program, scholars have also examined another pressing issue: whether participants comply with or shirk their responsibilities. Actors may join a program, yet not fulfill their obligations (Downs, Roche, & Barsboom, 1996; Mitchell, 1994; Simmons, 1998). Shirking is particularly important in the context of voluntary programs. Because goodwill and reputational benefits associated with program membership are excludable club goods, they serve as "selective benefits" to recruit members to these efforts (Potoski & Prakash, 2005). However, participants may have incentives to free ride on the efforts of other participants, if they can garner the excludable benefits of program membership without incurring its costs (Delmas & Keller, 2005; King & Lenox, 2000). Thus, how to curb shirking is a crucial issue for policy scholars.

This paper provides evidence that the UNGC, a voluntary program lacking stringent monitoring and enforcement mechanisms, does not adequately motivate its participants to comply with program obligations. Consequently, participants tend to undertake strategic shirking (Delmas & Keller, 2005; DeVaro & Gürtler, 2012): they tend not to adopt costly policies or take costly actions to improve their human rights and environmental performance. Instead, they have tended to adopt symbolic, low-cost steps to convey the impression that they are fulfilling their obligations (Lim & Tsutsui, 2012).

Should we then conclude that the Global Compact reflects a pattern of "bluewashing," whereby member firms figuratively drape themselves in the blue UN flag in order to distract stakeholders from their real, as opposed to cosmetic, poor environmental or human rights records? While the Compact's merits may have been oversold, its program design inadequate, and its performance disappointing, we recommend taking a more nuanced approach to the complex issue of program efficacy. We suggest viewing the Compact as an initial foray toward a potentially more successful global "meta-regime." While clearly inadequate in its design, we

believe it can potentially lay the foundation for a more robust program. Importantly, the Global Compact's leadership has recently shown a willingness to alter the program's rules to be slightly more stringent—shortening the deadlines for firms to submit their voluntary self-reports, and expelling greater numbers of members who failed to take even this most basic action. It remains to be seen the extent to which this small step has mitigated the shirking problem. Nevertheless, it is an important shift in focus, which may be followed by greater steps in the future.

Our paper raises important issues about the link between program design and program efficacy. While stringent monitoring and enforcement mechanisms provide greater assurance that obligations will be honored by program participants, such programs may set the bar too high, eventually leading to low levels of program participation. Further, monitoring and enforcement are expensive, and after some limit, have diminishing returns. A major challenge is to find the "optimal" stringency of such monitoring and enforcement mechanisms. This is an issue that any governance system, whether voluntary or mandatory, needs to confront.

Daniel Berliner is an assistant professor in the Department of Political Science at the University of Minnesota.

Aseem Prakash is the Walker Family Professor for the College of Arts and Sciences in the Department of Political Science at the University of Washington.

Notes

1. There is a vast literature on policy design (DeLeon, 1988; Linder & Peters, 1984; May, 1991; Schneider & Ingram, 1997) as well as social construction of public policy (Fischer, 2003; Schneider & Ingram, 1993; Stone, 1997). We do not directly contribute to this literature. Instead, our focus is on how program design influences program efficacy.
2. In addition to pollution reduction, efficacy can have other dimensions as well. There is some work on how voluntary programs may influence firms' regulatory compliance. Another body of works examines how voluntary programs may have second order effects such as on innovation levels (Lim & Prakash, 2014)
3. Noncompliance might also be attributed to ignorance (Brehm & Hamilton, 1996) and the lack of capacity (Scholz, 1991; Winter & May, 2001).
4. See <http://unglobalcompact.org/AboutTheGC/TheTenPrinciples/index.html>.
5. See <http://www.ipsnews.net/2010/06/qa-bluewashing-has-become-a-very-risky-business>.
6. In other work (Berliner & Prakash, 2012), we have found that conflict between important global stakeholders has shaped the international diffusion of UNGC membership, with countries that are more embedded in networks of international NGOs experiencing slower rates of adoption among local firms.
7. See <http://www.nytimes.com/2000/08/10/opinion/10iht-edglob.2.t.html>, <http://www.economist.com/node/2771522>, <http://globalcompactcritics.blogspot.com/2007/07/ngos-criticize-blue-washing-by-global.html>.
8. See <http://www.nytimes.com/2000/08/10/opinion/10iht-edglob.2.t.html>.
9. See <http://globalcompactcritics.blogspot.com/2007/07/ngos-criticize-blue-washing-by-global.html>.
10. The Human Rights category also contains indicators relevant to labor rights issues.
11. The definitions of each indicator we include can be found in Appendix S1 online.
12. In a robustness check, we present separate results for Environmental Strengths, Environmental Concerns, Human Rights Strengths, and Human Rights Concerns.

13. In 2008, coverage expanded again to ensure inclusion of the three thousand largest U.S. publicly traded companies by market capitalization. However, as we exclude firms that were only observed for three or fewer years, none of these enter our sample.
14. We matched firms using unique CUSIP identifier codes, supplemented with extensive checks “by hand” to remove duplicates, and to “connect” firms that changed codes during the period under observation, or whose codes were incorrectly entered in one of the data sources.
15. We include among these four firms that joined the Global Compact but were ultimately delisted for failure to submit COP reports. While the Compact did begin delisting larger numbers of firms for such failures in the latter years of our study, few of these delistings were in the United States, and fewer still among publicly listed firms eligible for inclusion in our analysis.
16. These divisions are: Agriculture, Forestry and Fishing; Mining; Construction; Manufacturing; Transportation, Communications, Electric, Gas and Sanitary; Wholesale Trade; Retail Trade; Finance, Insurance and Real Estate; Services; and Public Administration. However, since so many firms fall into the Manufacturing division, we divide them into two groups: “Manufacturing” (SIC 4-digit codes 2000 through 3499) and “High-Tech Manufacturing” (SIC 4-digit codes 3500 through 3999).
17. In order to make this variable more fine-grained, we use sectors at the SIC 2-digit level.
18. This variable is calculated using sectors at the SIC division level (with Manufacturing split into two divisions as noted earlier), as it was not possible to match the UN Global Compact’s own sector categories with anything more fine-grained.
19. Note that while institutional isomorphism contributes to firms undertaking the low-cost step of joining the Global Compact, we do not expect it to play a role in shaping actual changes in CSR performance. This is in line with the conclusions of research in the World Society school (Lim & Tsutsui, 2012; Meyer, Boli, Thomas, & Ramirez, 1997).
20. The inflation equation includes the independent variables UN Global Compact membership, Firm Size, Firm Profits, and year fixed effects. It is not computationally feasible to include the sectoral fixed effects in the inflation equation as well as in the outcome equation.

References

- Andrews, Richard N. L., Deborah Amaral, Nicole Darnall, Deborah Rigling Gallagher, Daniel Edwards Jr., Andrew Huston et al. 2003. *Environmental Management Systems: Do They Improve Performance?* Chapel Hill: University of North Carolina at Chapel Hill and Environmental Law Institute.
- Bartley, Tim, and Curtis Child. 2011. “Movements, Markets and Fields: The Effects of Anti-Sweatshop Campaigns on U.S. Firms, 1993–2000.” *Social Forces* 90 (2): 425–51.
- Berliner, Daniel. 2013. *Follow your Neighbor? Regional Emulation and the Design of Transparency Policies*. KFG Working Papers Series, Freie Universität Berlin.
- . 2014. “The Political Origins of Transparency.” *The Journal of Politics* 76 (2): 479–91.
- Berliner, Daniel, and Aseem Prakash. 2013. “Signaling Environmental Stewardship in the Shadow of Weak Governance: The Global Diffusion of ISO 14001.” *Law & Society Review* 47 (2): 345–73.
- . 2015. “How Domestic Regulatory Institutions Influence the Adoption of Global Private Regimes: Firm-level Evidence from Eastern Europe and Central Asia.” *International Studies Quarterly*, Forthcoming.
- Berliner, Daniel, and Aseem Prakash. 2012. “From Norms to Programs: The United Nations Global Compact and Global Governance.” *Regulation & Governance* 6 (2): 149–66.
- Bernhagen, Patrick, and Neil Mitchell. 2010. “The Private Provision of Public Goods: Corporate Commitments and the United Nations Global Compact.” *International Studies Quarterly* 54: 1175–87.
- Bertrand, Marianne, and Sendhil Mullainathan. 2003. “Enjoying the Quiet Life? Corporate Governance and Managerial Preferences.” *Journal of Political Economy* 111 (5): 1043–75.
- Borck, Jonathan C., and Cary Coglianese. 2009. “Voluntary Environmental Programs: Assessing their Effectiveness.” *Annual Review of Environment and Resources* 34: 305–24.
- Börzel, Tanja, and Thomas Risse. 2010. “Governance Without a State: Can it Work?” *Regulation and Governance* 4: 113–34.

- Brehm, John, and James T. Hamilton. 1996. "Noncompliance in Environmental Reporting: Are Violators Ignorant, or Evasive of the Law?" *American Journal of Political Science* 40: 444–47.
- Cashore, Benjamin, Graeme Auld, and Deanna Newsom. 2004. *Governing Through Markets*, New Haven, CT: Yale University Press.
- Chatterji, Aaron K., David I. Levine, and Michael W. Toffel. 2009. "How Well Do Social Ratings Actually Measure Corporate Social Responsibility?" *Journal of Economics & Management Strategy* 18: 125–69.
- Coglianesi, Cary, and Jennifer Nash, eds. 2001. *Regulating from the Inside: Can Environmental Management Systems Achieve Policy Goals?* Washington, DC: Resources for the Future.
- Dahlström, Kristina, Chris Howes, Paul Leinster, and Jim Skea. 2003. "Environmental Management Systems and Company Performance: Assessing the Case for Extending Risk-Based Regulation." *European Environment* 13: 187–203.
- Darnall, Nicole, and Younsung Kim. 2012. "Which Types of Environmental Management Systems are Related to Greater Environmental Improvements?" *Public Administration Review* 72 (3): 351–65.
- Darnall, Nicole, G. Jason Jolley, and Robert Handfield. 2008. "Environmental Management Systems and Green Supply Chain Management: Complements for Sustainability?" *Business Strategy and the Environment* 17: 30–45.
- Dasgupta, Susmita, Hemamala Hettige, and David Wheeler. 2000. "What Improves Environmental Compliance? Evidence from Mexican Industry." *Journal of Environmental Economics and Management* 39: 39–66.
- Dashwood, Hevina. 2012. *The Rise of Global Corporate Social Responsibility: Mining and the Spread of Global Norms*. New York: Cambridge University Press.
- Deckop, John R., Kimberly K. Merriman, and Shruti Gupta. 2006. "The Effects of CEO Pay Structure on Corporate Social Performance." *Journal of Management* 32 (3): 329–42.
- Delmas, Magali A., and Vanessa Cuerel-Burbano. 2011. "The Drivers of Greenwashing." *California Management Review* 54: 64–87.
- Delmas, Magali A., and Arturo Keller. 2005. "Strategic Free Riding in Voluntary Programs: The Case of the US EPA Wastewater Program." *Policy Sciences* 38: 91–106.
- DeLeon, Peter. 1988. "The Contextual Burdens of Policy Design." *Policy Studies Journal* 17 (2): 297–309.
- DeLeon, Peter, and Jorge E. Rivera, eds. 2009. *Voluntary Environmental Programs: A Policy Perspective*. Lanham, MD: Rowman & Littlefield for The Policy Studies Organization.
- DeVaro, Jed, and Oliver Gürtler. 2012. *Strategic Shirking in Promotion Tournaments* [Online]. Available at SSRN: <http://ssrn.com/abstract=2174394>. Accessed January 10, 2014.
- DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48 (2): 147–60.
- Djankov, Simeon, Caralee McLiesh, and Andrei Shleifer. 2007. "Private Credit in 129 Countries." *Journal of Financial Economics* 84 (2): 299–329.
- Doh, Jonathan, Shawn Howton, Shelly Howton, and Donald Siegel. 2010. "Does the Market Respond to an Endorsement of Social Responsibility? The Role of Institutions, Information, and Legitimacy." *Journal of Management* 36 (6): 1461–85.
- Downs, George, David Rocke, and Peter Barsoom. 1996. "Is the Good News About Compliance Good News About Cooperation?" *International Organization* 50 (3): 379–406.
- EthicalCorp.com. 2008. *Debate: UN Global Compact—Is the Compact Raising Corporate Responsibility Standards?* [Online]. <http://www.ethicalcorp.com/content/debate-un-global-compact-%E2%80%93compact-raising-corporate-responsibility-standards>. Accessed January 10, 2014.
- Fiorino, Daniel. 2006. *The New Environmental Regulation*. Cambridge, MA: MIT Press.
- Fischer, Frank. 2003. *Reframing Public Policy: Discursive Politics and Deliberative Practices: Discursive Politics and Deliberative Practices*. Oxford: Oxford University Press.
- Gamper-Rabindran, Shanti, and Stephen Finger. 2013. "Does Self-Regulation Reduce Pollution? Responsible Care Program in the Chemical Industry." *Journal of Regulatory Economics* 43 (1): 1–30.
- Girma, Sourafel, Avid Greenaway, and Richard Kneller. 2004. "Does Exporting Increase Productivity? A Microeconomic Analysis of Matched Firms." *Review of International Economics* 12 (5): 855–66.

- Greenhill, Brian. 2010. "The Company You Keep: International Socialization and the Diffusion of Human Rights Norms." *International Studies Quarterly* 54 (1): 127–45.
- Guler, Isin, Mauro F. Guillén, and John Muir Macpherson. 2002. "Global Competition, Institutions, and the Diffusion of Organizational Practices: The International Spread of ISO 9000 Quality Certificates." *Administrative Science Quarterly* 47 (2): 207–32.
- Gunningham, Neil, Robert Kagan, and Dorothy Thornton. 2004. "Social License and Environmental Protection: Why Businesses Go Beyond Compliance." *Law and Social Inquiry* 29 (2): 307–41.
- Hamann, Ralph, Paresha Sinha, Farai Kapfudzaruwa, and Christoph Schild. 2009. "Business and Human Rights in South Africa: An Analysis of Antecedents of Human Rights due Diligence." *Journal of Business Ethics* 87 (2): 453–73.
- Kell, Georg, and David Levin. 2003. "The Global Compact Network: An Historic Experiment in Learning and Action." *Business and Society Review* 108 (2): 151–81.
- Khanna, Madhu, and Lisa A. Damon. 1999. "EPAs Voluntary 33/50 Program: Impact on Toxic Releases and Economic Performance of Firms." *Journal of Environmental Economics and Management* 37: 1–25.
- King, Andrew. 2007. "Cooperation between Corporations and Environmental Groups: A Transaction Cost Perspective." *Academy of Management Review* 32: 889–900.
- King, Andrew, and Michael Lenox. 2000. "Industry Self-Regulation Without Sanctions: The Chemical Industry's Responsible Care Program." *Academy of Management Journal* 43: 698–716.
- Levine, David I., and Michael W. Toffel. 2010. "Quality Management and Job Quality: How the ISO 9001 Standard for Quality Management Systems Affects Employees and Employers." *Management Science* 56 (6): 978–96.
- Lim, Alwyn, and Kiyoteru Tsutsui. 2012. "Globalization and Commitment in Corporate Social Responsibility: Cross-National Analyses of Institutional and Political-Economy Effects." *American Sociological Review* 77: 69–98.
- Lim, Sijeong, and Aseem Prakash. 2014. "Voluntary Regulations and Innovation: The Case of ISO 14001." *Public Administration Review* 74 (2): 233–44.
- Linder, Stephen, and B. Guy Peters. 1984. "From Social Theory to Policy Design." *Journal of Public Policy* 4: 237–59.
- Locke, Richard. 2013. *The Promise and Limits of Private Power: Promoting Labor Standards in a Global Economy*. New York: Cambridge University Press.
- Lundgren, Tommy. 2003. "A Real Options Approach to Abatement Investments and Green Goodwill." *Environmental and Resource Economics* 25 (1): 17–31.
- MacGarvie, Megan. 2006. "Do Firms Learn From International Trade?" *Review of Economics and Statistics* 88 (1): 46–60.
- March, James G., and Johan P. Olsen. 2006. "The Logic of Appropriateness," In *The Oxford Handbook of Public Policy*, eds. M. Moran, M. Rein, and R. E. Goodin. Oxford: Oxford University Press, 689–708.
- Mattingly, James E., and Shawn L. Berman. 2006. "Measurement of Corporate Social Action Discovering Taxonomy in the Kinder Lydenburg Domini Ratings Data." *Business & Society* 45 (1): 20–46.
- May, Peter J. 1991. "Reconsidering Policy Design: Policies and Publics." *Journal of Public Policy* 11: 187–206.
- Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of sociology* 103 (1): 144–81.
- Mitchell, Ronald. 1994. "Regime Design Matters: Intentional Oil Pollution and Treaty Compliance." *International Organization* 48: 425–58.
- Morgenstern, Richard, and William Pizer, eds. 2007. *Reality Check: The Nature and Performance of Voluntary Environmental Programs in the United States, Europe, and Japan*. Washington, DC: RRF Press.
- Mwangi, Wagaki, Lothar Rieth, and Hans Peter Schmitz. 2013. "Encouraging Greater Compliance: Local Networks and the United Nations Global Compact." In *The Persistent Power of Human Rights: From Commitment to Compliance*, eds. Thomas Risse, Stephen C. Ropp, and Kathryn Sikkink. New York: Cambridge University Press, 203–21.
- Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.
- Ostrom, Elinor. 1990. *Governing the Commons*. New York: Cambridge University Press.

- Perez-Batres, Luis A., Van V. Miller, and Michael J. Pisani. 2011. "Institutionalizing Sustainability: An Empirical Study of Corporate Registration and Commitment to the United Nations Global Compact Guidelines." *Journal of Cleaner Production* 19 (8): 843–51.
- Perkins, Richard, and Eric Neumayer. 2010. "Geographic Variations in the Early Diffusion of Corporate Voluntary Standards: Comparing ISO 14001 and the Global Compact." *Environment and Planning A* 42 (2): 347–65.
- Potoski, Matthew, and Aseem Prakash. 2005. "Green Clubs and Voluntary Governance: ISO 14001 and Firms' Regulatory Compliance." *American Journal of Political Science* 49: 235–48.
- Prakash, Aseem, and Matthew Potoski. 2006. *The Voluntary Environmentalists*. Cambridge, UK: Cambridge University Press.
- . 2007. "Investing Up: FDI and the Cross-Country Diffusion of ISO 14001 Management Systems." *International Studies Quarterly* 51 (3): 723–44.
- Rasche, Andreas. 2009. "A Necessary Supplement: What the United Nations Global Compact is and is Not." *Business & Society* 48 (4): 511–37.
- Rasche, Andreas, Sandra Waddock, and Malcolm McIntosh. 2013. "The United Nations Global Compact Retrospect and Prospect." *Business & Society* 52 (1): 6–30.
- Ruggie, John. 1982. "International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order." *International Organization* 36: 379–415.
- . 2002. "The Theory and Practice of Learning Networks." *Journal of Corporate Citizenship* 2002 (5): 27–36.
- Ruggie, John, and Georg Kell. 1999. "Global Markets and Social Legitimacy: The Case of the Global Compact." *Transnational Corporations* 8: 101–20.
- Russo, Michael V. 2002. "Institutional Change and Organization Strategy: ISO 14001 and Emissions in the Electronics Industry." *Academy of Management Proceedings* 1: A1–6.
- Schneider, Anne, and Helen Ingram. 1993. "Social Construction of Target Populations: Implications for Politics and Policy." *American Political Science Review* 87 (2): 334–47.
- Schneider, Anne Larason, and Helen M. Ingram. 1997. *Policy Design for Democracy*, Lawrence, KS: University Press of Kansas.
- Scholz, John T. 1991. "Cooperative Regulatory Enforcement and the Politics of Administrative Effectiveness." *The American Political Science Review* 85 (1): 115–36.
- Sethi, S. Prakash, and Donald H. Schepers. 2014. "United Nations Global Compact: The Promise–Performance Gap." *Journal of Business Ethics* 122 (2): 193–208.
- Simmons, Beth. 1998. "Compliance with International Agreements." *Annual Reviews of Political Science* 1: 75–93.
- Simmons, Beth, Frank Dobbin, and Geoffrey Garrett. 2006. "Introduction: The International Diffusion of Liberalism." *International Organization* 60 (4): 781–810.
- Slaughter, Matthew J. 2001. "Trade Liberalization and Per Capita Income Convergence: A Difference-in-Differences Analysis." *Journal of International Economics* 55 (1): 203–28.
- Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 88: 355–74.
- Stone, Deborah A. 1997. *Policy Paradox: The Art of Political Decision Making*. New York: WW Norton.
- Tashman, Peter, and Jorge Rivera. 2010. "Are Members of Business for Social Responsibility More Socially Responsible?" *Policy Studies Journal* 38: 487–514.
- UN Global Compact. 2009. *United Nations Global Compact Annual Review 2008*. [Online]. http://www.unglobalcompact.org/docs/news_events/9.1_news_archives/2009_04_08/GC_2008AR_FINAL.pdf. Accessed July 1, 2009.
- Vidovic, Martina, Neha Khanna, and Michael Delgado. 2013. *Third Party Certification and the Effectiveness of Voluntary Pollution Abatement Programs: Evidence from Responsible Care*. Agricultural and Applied Economics Association 2013 Annual Meeting, August 4–6, 2013, Washington, DC.
- Vogel, David. 2005. *The Market for Virtue*. Washington, DC: The Brookings Press.
- Waddock, Sandra A., and Samuel B. Graves. 1997. "The Corporate Social Performance–Financial Performance Link." *Strategic Management Journal* 18: 303–17.

- Welch, Eric W., Allan Mazur, and Stuart Bretschneider. 2000. "Voluntary Behavior by Electric Utilities: Levels of Adoption and Contribution of the Climate Challenge Program to the Reduction of Carbon Dioxide." *Journal of Policy Analysis and Management* 19: 407–25.
- Winter, Søren C., and Peter J. May. 2001. "Motivation for Compliance with Environmental Regulations." *Journal of Policy Analysis and Management* 20: 675–98.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Definitions of KLD Indicators Included in the Analysis.

Table S1. Full Results of Models Presented in Table 2 in the Main Text.