Multinational enforcement of labor law: Experimental evidence from Bangladesh's apparel sector

Laura Boudreau*

May 15, 2020

Abstract

Western stakeholders are increasingly demanding that multinationals sourcing from developing countries be accountable for labor rights and working conditions upstream in their supply chains. In response, many multinationals privately enforce labor standards in these countries, but the effects of their interventions on local firms and workers are unknown. I partnered with a set of multinational retail and apparel firms to enforce local labor laws on their suppliers in Bangladesh. I implemented a randomized controlled trial with 84 Bangladeshi garment factories, randomly enforcing a mandate for worker-manager safety committees in 41 supplier establishments. The intervention significantly improves compliance with the labor law. It also has a positive effect on indicators of factory safety, including measures of physical safety and awareness. These improvements do not appear to come at significant costs to suppliers in terms of efficiency. Factories with better managerial practices drive these improvements. In contrast, factories with poor managerial practices do not improve compliance or safety, and in these factories, workers' job satisfaction declines. JEL Codes: F61, J53, J81, L14, O12, O14

Keywords: Labor Regulation, Enforcement, Global supply chains, Corporate Social Responsibility, Private regulation, Working Conditions, Worker voice, Safety, Garment Industry, Bangladesh

*Columbia University & CEPR (l.boudreau@columbia.edu). I thank Noam Yuchtman, Reed Walker, Ernesto Dal Bó, Paul Gertler, and Ted Miguel for their guidance and encouragement. This research has benefited from comments and suggestions from Matt Amengual, Marianne Bertrand, Nick Bloom, Greg Distelhorst, Fred Finan, John de Figueiredo, David Levine, Rocco Macchiavello, Jeremy Magruder, Felix Oberholzer-Gee, Raul Sanchez de la Sierra, Daniela Scur, Abu Shonchoy, Mike Toffel, and numerous seminar and conference participants. This project is funded by the Alliance for Bangladesh Worker Safety, the Private Enterprise Development in Low-income Countries (PEDL) Initiative, the Weiss Family Program Fund for research in Development Economics, the International Growth Centre (IGC), and the Abdul Latif Jameel Poverty Action Lab (J-PAL) Governance Initiative. The Sylff Research Abroad Program also contributed travel funding.

1 Introduction

Over the past four decades, developing countries have become increasingly incorporated into global value chains (GVCs), which has raised economic growth and reduced poverty (World Bank, 2020). Despite these gains, poor working conditions and labor rights commonly persist (International Labor Organization, 2016). In light of weak local enforcement capacity and corruption (La Porta et al., 1999, Fisman and Wang, 2015, Amirapu and Gechter, forthcoming), Western stakeholder are increasingly demanding accountability from multinational corporations (MNCs) that source from developing countries. In response, many MNCs implement "Corporate Social Responsibility" (CSR) programs that claim to privately enforce local labor laws on their suppliers. It is an open question, however, whether these interventions achieve their stated compliance objectives.

This paper provides the first experimental evidence on the effects of private enforcement of labor law by MNCs in a developing country. MNCs' voluntary adoption of labor standards and their increasing participation in multistakeholder enforcement initiatives suggests that MNCs may have incentives to prevent violations (Kitzmueller and Shimshack, 2012, O'Rourke, 2014), at least those that could pose reputational risks (Macchiavello and Morjaria, 2015, Bai, Gazze and Wang, 2019). On the other hand, increased protections for workers may erode suppliers' competitiveness (Botero et al., 2004, Besley and Burgess, 2004), which suggests that without effective monitoring, MNCs' promises may not be credible (Besley and Ghatak, 2007). Finally, even if MNCs are motivated to improve labor standards, it's unclear whether their suppliers are capable of implementing better practices.

To test these possibilities, I partnered with a group of multinational retail and apparel firms, the Alliance for Bangladesh Worker Safety (hereafter, the Alliance), that expressed a desire to improve the safety performance of its shared Bangladeshi supplier base.[1] The

[1] As per the Alliance Members' Agreement, the Alliance ceased operations on December 31, 2018. Concurrently, many Alliance members supported the formation of Nirapon, an organization tasked with a similar set of safety oversight functions as the Alliance. 22 out of 29 Alliance Members joined Nirapon.

Alliance's membership included 29 multinational retail and apparel firms representing the majority of North American imports from Bangladesh (e.g., Wal-Mart, Gap, Target).[2] The Alliance committed to enforce a local labor law that mandates safety committees (SCs), which are joint worker-manager bodies tasked with helping to create and to maintain a safe workplace. The Alliance's stated objective was to hold suppliers accountable for SCs, and in particular, to validate that workers were empowered to report safety risks without fear of retaliation.

In collaboration with the Alliance, I implemented a randomized controlled trial (RCT) in which I randomly assigned supplier factories to the Alliance's enforcement of the SC law. The Alliance's intervention entailed six months of intensive monitoring of factories' adherence to the SC law through mandated reporting of SCs' activities and monitoring via email, phone calls, and onsite visits.[3] I estimate the intervention's effects on suppliers' compliance with the law and on indicators of safety. I also assess its effects on factories' labor productivity, wages, and employment and on their workers' well-being.

The RCT was implemented with 84 garment and garment-related factories over 2017-2018 as part of the SC Program roll-out. I randomly assigned 41 factories to immediate participation (treatment) and 43 factories to deferred participation approximately 11 months later (control).[4] The research team made three full-day visits to factories. The team collected a pre-intervention baseline and two post-intervention rounds about 5 and 9 months later, respectively. For treatment factories, the second visit occurred toward the end of the 6-month enforcement intervention. I also implemented a retrospective

[2]Alliance Members: Ariela and Associates International LLC; Bon Worth; Canadian Tire Corporation, Limited; Carter's Inc.; The Children's Place Retail Stores Inc.; Costco Wholesale Corporation; Fruit of the Loom, Inc.; Gap Inc.; Giant Tiger; Hudson's Bay Company; IFG Corp.; Intradeco Apparel; J.C. Penney Company Inc.; Jordache Enterprises, Inc.; The Just Group; Kate Spade & Company; Kohl's Department Stores; L. L. Bean Inc.; M. Hidary & Company Inc.; Macy's; Nordstrom; One Jeanswear Group; Public Clothing Company; Sears Holdings Corporation; Target Corporation; The Warehouse; VF Corporation; Wal-Mart Stores, Inc.; and YM Inc.

[3]Post-Alliance, the organization Nirapon is implementing a safety committee program that is based on the Alliance's Safety Committee Program.

[4]Factories were not aware of their experimental status. Due to logistical constraints, the Alliance rolls out all of its programs in stages, so this design aligns with the Alliance's standard operating procedures.

questionnaire to collect factories' business-related data. Finally, the Alliance provided its administrative records. The consolidated datasets are unique in their comprehensiveness and depth. I analyze them according to a pre-analysis plan (PAP), which is registered on the American Economic Association's Social Science Registry.

I find that the MNCs' enforcement program significantly increases factories' compliance with Bangladesh's SC law, which I measure using an index of compliance outcomes. The intervention improves factories' compliance by 0.20 standard deviations (sds) on average. Most factories begin with SCs that are formed correctly but largely inactive. The intervention significantly increases SCs' level of activity; for example, they begin to meet more frequently and are nearly four times more likely to conduct risk assessment. The increase in compliance translates into a statistically significant improvement in an index of factory safety indicators (0.14 sds), which is mostly driven by an improvement in treatment factories' performance on an independent spotcheck of safety conditions. Finally, medical clinic records available for 62 factories document that the proportion of the workforce seeking medical care declines by between 15-16%. Evidently, MNCs' enforcement interventions can meaningfully improve working conditions.

There is no evidence of adverse effects on supplier competitiveness, including labor productivity, wages, and employment. Estimated treatment effects on labor productivity are actually positive but are small. Estimated treatment effects on wages and employment are negative but are close to zero and not statistically significant. In light of these null results, I report ex post Minimum Detectable Effect Sizes (MDEs) for these outcomes. I am underpowered to detect all but large effects on labor productivity, but I can rule out moderate or larger effects on wages and employment. Viewed together, the results on suppliers' business competitiveness suggest that the intervention significantly improved safety without negative effects on suppliers' competitiveness.

Finally, I explore whether providing firms with incentives to comply with the SC law is sufficient to increase compliance or whether firms' capacity to respond to incentives

also matters. The existing research on improving adherence to regulation in developing countries focuses on strengthening state-supplied enforcement in order to increase firms' incentives for compliance (Duflo et al., 2013, 2014, Dal Bó and Finan, 2016). There has been little to no consideration, however, of whether the capacity of the private sector also constrains regulatory efficacy. I test whether the intervention's effects depend on firms' capacity using a measure of their management practices.

I find that factories' baseline managerial practices are an important factor in determining the enforcement intervention's effects.[5] The treatment has large, positive effects on compliance and on safety indicators in factories with better baseline managerial practices. In contrast, factories with worse practices do not significantly improve. The treatment effect heterogeneity is also present in the medical clinic results. These results suggest that there may be complementarity between labor regulation and managerial practices. Increasing compliance may depend not only on providing firms with appropriate incentives, but also on their capacity to respond to these incentives.

Further, the enforcement intervention has a negative effect on indicators of workers' job satisfaction in poorly-managed factories, measured using an index of survey and revealed preference measures. I provide suggestive evidence of a mechanism in which the intervention raises workers' expectations about what SCs will deliver, and in poorly-managed factories, these expectations are not met. This mechanism would be consistent with experimental evidence on low-wage workers' response to an upgrade in employer-provided housing from Adhvaryu, Nyshadham and Xu (2018). Objective measures of housing quality improve, but do not meet workers' expectations, and turnover increases.

This research makes four primary contributions. First, this paper contributes to the literature on labor regulation and economic development, and in particular, their interaction with GVCs. Several studies have found that heavier *de jure* labor regulation is associated with worse economic performance and adverse consequences for workers (Fishback

[5]For certain variables, due to power limitations, I am unable to reject that the estimated treatment effects are different for the two groups.

and Kantor, 1996, Botero et al., 2004, Besley and Burgess, 2004, Aghion et al., 2008). We also know, however, that weak state capacity and political capture by elites results in socially suboptimal enforcement in many developing countries (Fisman and Wang, 2015, Dal Bó and Finan, 2016, Amirapu and Gechter, forthcoming). Scholars have raised GVCs as a possible mechanism to bring about improved regulation and enforcement. Harrison and Scorse (2010) show that anti-sweatshop campaigns led the Indonesian government to raise minimum wages, which caused large real wage increases with some costs for firms but no significant effects on employment. Tanaka (forthcoming) provides evidence of trade-induced social upgrading among firms in Myanmar. This is the first study, however, to test the potential for multinational enforcement in a context where state enforcement is lacking. My findings demonstrate that private enforcement can improve compliance and contribute to achieving the law's objectives. I further contribute by identifying enforcement's causal effects on labor productivity and workers' well-being. Finally, my results suggest an under-explored constraint on regulatory efficacy: Managerial capacity of the private sector.

Second, it contributes to a burgeoning literature on the economics of CSR. Economists have long espoused Friedman (1970)'s view that markets should produce private goods and governments should provide public goods and correct failures. Recent theoretical and empirical work, however, highlights two reasons why this dichotomy may blur. First, governments, particularly in developing countries, frequently fail to fulfill their aforementioned roles; further, governments' jurisdiction is limited to their territories, and they are often constrained in their ability to police production abroad (Besley and Ghatak, 2007, Bénabou and Tirole, 2010, Dal Bó and Finan, 2016). Second, consumers, shareholders, and workers have social and ethical motivations and often value production that occurs socially and environmentally responsible ways (Besley and Ghatak, 2007, Dragusanu, Giovannucci and Nunn, 2014, Hainmueller, Hiscox and Sequeira, 2015, Burbano, 2016, Hart and Zingales, 2017). The existing economic literature on CSR primarily pro-

vides the bases for its existence and desirability. This paper joins Dragusanu and Nunn (2018) and Macchiavello and Miquel-Florensa (2019), both of which study the welfare effects of CSR programs on coffee farmers, in beginning to build a body of empirical evidence on the efficacy of CSR. I provide the first experimental evidence that firms' CSR initiatives can generate public goods/curtail public bads.[6]

Third, this paper contributes to the literature on collective worker voice and intra-firm institutions, including occupational safety and health (OSH) committees. The empirical literature on this topic has generally suffered from selection bias, and until recently, available causal evidence has largely been limited to marginal firms (Addison, Schnabel and Wagner, 2001, DiNardo and Lee, 2004, Lee and Mas, 2012, Yao and Zhong, 2013, Jäger, Schoefer and Heining, 2020). A much smaller literature on OSH committees examines correlations among the presence and features of OSH committees and injury rates or stakeholder satisfaction with them (see Yassi et al. (2013) for a review). My contribution is to randomize enforcement of worker-manager OSH committees to provide causal evidence of their effects on factories' and workers' outcomes.

Finally, this paper contributes to a nascent literature on the role of *private* sector capacity in determining the efficacy of *public* regulation. Existing research focuses on determinants of public sector capacity – the selection of public servants (Dal Bó, Finan and Rossi, 2013), their financial incentives (Ashraf, Bandiera and Jack, 2014), and their human capital and public sector motivation (Dal Bó, Finan and Rossi, 2013, Callen et al., 2018, Ashraf et al., forthcoming). In contrast, my analysis of heterogeneous treatment effects by management practices shows that the effects of regulation also depend on private sector capacity. Regulatory compliance is a function of both firms' incentives for compliance

[6] A literature spanning political science and management science asks related questions. This literature is largely skeptical, but as more causal evidence emerges, it is updating its view. In particular, Amengual and Distelhorst (2019), who also provide an excellent summary of the literature, use a regression discontinuity design to study Gap Inc's supplier code of conduct for labor. They find that a failing audit grade only improves compliance if coupled with the threat of a reduction in Gap's orders.

and their managerial capacity to respond to these incentives.[7]

The remainder of this paper is organized as follows: Section 2 describes the context, the Alliance, and the SC Program. Section 3 discusses the conceptual framework. Section 4 presents the research design. Section 5 presents the results. Section 6 concludes.

2 Background

Bangladesh's garments sector

Bangladesh plays a critical role in the global apparel supply chain. It is the second largest exporter of clothing in the world behind China (World Trade Organization, 2017). MNCs rely on Bangladesh for its combination of low prices and large production capacity (McKinsey & Company, 2011).[8]

Apparel is also a critical sector for Bangladesh's economy. Bangladesh is one of the most rapidly industrializing countries in the world (Central Intelligence Agency, 2016), and the garments sector has been and continues to be the major driver of its industrial transformation. In 2016, apparel exports constituted 81% of Bangladesh's total exports and 13% of its Gross Domestic Product.[9] The sector directly employs between 4-5 million of Bangladesh's 66.6 million workers.

Bangladesh has been infamous for its weak legal protections for workers, its lack of enforcement, and its low minimum wages for many years.[10] In a 2011 McKinsey survey of western buyers, buyers list lack of social compliance and economic and political instability as two of the top five risks to sourcing from the country (McKinsey & Company,

[7] Recent work by Almunia et al. (2020) finds that firms' "capacity," broadly defined, also plays a role in determining firms' adherence to tax codes, with some firms disadvantageously over-reporting their taxes even when enforcement is low.

[8] A Chief Procurement Officer (CPO) of a major Western retail firm states it simply in a 2011 McKinsey survey, "There is no alternative to Bangladesh."

[9] Author's calculations using data from the World Trade Organization and the World Bank.

[10] Garment sector jobs are not without benefits to Bangladeshi society. Heath and Mobarak (2015), for example, show that the growth in these jobs contributed to decreasing fertility, increasing age at marriage, and increasing educational attainment among Bangladeshi girls in recent decades.

2011). Decades of rapid industrial growth and weak state institutions culminated in a series of high fatality industrial accidents in 2012-13, including the collapse of the Rana Plaza building (see Figure II), that killed at least 1,273 workers and injured at least 3,812 workers at exporting factories (Solidarity Center, 2016). In the aftermath of these events, world leaders rebuked the Government of Bangladesh (GoB) for "not taking steps to afford internationally recognized worker rights to workers in that country," and western governments penalized the country by removing trade benefits (Greenhouse, 2013*a*).

Government and buyer response to the Rana Plaza collapse

Following the collapse, the GoB and MNCs faced intense pressure from the international community, consumers, and activists to ensure workers' safety and basic rights. The GoB promised to introduce labor reform and to work with the International Labor Organization (ILO) and other stakeholders to prevent another tragedy. European buyers quickly signed an agreement between buyers and labor unions to improve OSH in Bangladesh's garments sector. The coalition was called the Accord on Fire and Building Safety in Bangladesh (hereafter, the Accord). Several U.S. retailers refused to sign the Accord due to labor unions' participation and to the clause that buyers are subject to legally-binding arbitration (Greenhouse, 2013*b*, Bhattacharjee, 2013). A group of U.S. retailers formed the Alliance for Bangladesh Worker Safety (hereafter, the Alliance) shortly thereafter.

In July 2013, the GoB amended the labor law to strengthen OSH and freedom of association provisions. One of the amendment's key provisions was the requirement for SCs.[11] The GoB also agreed to a a multi-stakeholder action plan to strengthen its capacity and to improve the sector's OSH outcomes (Ministry of Labour and Employment, 2013). To fulfill the action plan, the GoB closely coordinated with the ILO, the Accord, and the Alliance. The Accord and the Alliance were responsible for overseeing safety for the 60-70% of the sector that they covered. The GoB, with the ILO's support, was responsible for

[11]The mandate for SCs applies to establishments with 50 or more workers.

the remaining 30-40% of the sector (International Labor Organization, 2017).

In September 2015, the GoB published implementation rules for the SC mandate that articulate requirements for SCs' formation, operations, and responsibilities. Table I summarizes the key aspects. Establishments had six months to form and to operationalize their SCs. Despite the *de jure* requirement to implem SCs, *de facto*, enforcement was low. According to an International Labor Organization (2017) report, from 2015-2017, the GoB's focus was primarily on physical safety remediation.[12] Unsurprisingly, as captured by the headline of a news article from late 2017, *"Half of all apparel units flout needs for safety committees,"* compliance with the regulation was low (Munni, 2017).

The Alliance & the SC Program

The Alliance was a coalition of 29 multinational retail and apparel firms (e.g., Wal-Mart, Gap, Target, Costco), which are displayed in Figure III. The Alliance's members represented the majority of North American garment imports from Bangladesh. They committed to a five-year agreement to improve the safety performance of their Bangladeshi supplier bases, which included between 600-700 factories and 1.21 million workers. Online Appendix C provides an overview of the Alliance's Member Agreement, its programs, and the nature of this research collaboration.[13] As per its Members' Agreement, the Alliance ceased operations on December 31, 2018. Concurrently, many Alliance members supported the formation of Nirapon, an organization tasked with a similar set of safety oversight functions. 22 out of 29 Alliance Members joined Nirapon, which as of early 2020, is fighting a court battle in Bangladesh to continue its operations.

The Alliance required all supplier factories to participate in its building safety audit and remediation and worker training and empowerment programs. Failure to comply with one or more programs resulted in suspension from all Alliance Members' supplier

[12] As of mid-2017, the ILO had supported the GoB to form SCs at 210 of the 1,549 garment factories under the government's purview (i.e., not including Alliance or Accord-covered factories).

[13] More information is also available on the Alliance's website: www.bangladeshworkersafety.org.

bases; the Alliance suspended 179 factories over its five-and-a-half year term.[14]

The Alliance's intervention to enforce the SC law was its SC Program, which included four phases:[15]

1. If necessary, the Alliance worked with the factory to reform the SC through compliant processes.
2. The Alliance provided SC members with training on their roles and responsibilities, on occupational safety and health, and on leadership and communication skills.
3. The SC prepared an action plan for required activities.
4. Once the Alliance approved the action plan, the Alliance intensively monitored the SC on its completion.

The SC Program's central feature was the SC's preparation and fulfillment of its action plan. The plan used an Alliance-provided template and included a detailed schedule of required activities. Several members of management, the SC President, and the SC Vice President had to sign off on it. Before approving the plan, the Alliance reviewed it and worked with the factory to make revisions. The factory then implemented the plan and provided evidence of its activities to the Alliance. The Alliance required factories to submit evidence of these activities by e-mail within 2-3 days. Repeated failure to submit evidence resulted in escalation of the factory's status toward suspension. The Alliance reviewed submissions and investigated by phone calls, e-mails, and onsite audits that are unannounced or announced within a certain time period. At the end of the six-month period, the Alliance reviewed the factory's progress. If the Alliance found it to be insufficient, the factory could be required to repeat parts of the SC Program or its status could be escalated toward suspension. If found to be sufficient, the factory returned to the pool of factories being monitored through the Alliance's general monitoring program.[16]

The Alliance implemented the SC Program with factories that did not supply to Ac-

[14]The Alliance was also a member of a Private sector-GoB Factory Closure Panel for cases of imminent danger due to structural integrity, which fully or partially closed 35 factories that supplied to the Alliance.

[15]Nirapon continues to require supplier factories to participate in a similar SC Program.

[16]More information about the SC Program, including many of the Program's materials, is available on the Alliance website.

cord buyers; this is because the Accord also implemented a SC Program with its suppliers, including those that were covered by the Alliance.

3 Conceptual framework

The most basic question about the impact of MNCs' CSR enforcement programs is whether they actually produce compliance with labor standards in their supply chains. The answer is theoretically ambiguous. On one hand, a Becker (1968)-style model of illicit behavior predicts that if MNCs' increase monitoring and penalties, suppliers will increase compliance. On the other, models of CSR that are consistent with these MNCs' behavior, such as (Besley and Ghatak, 2007), predict that profit maximizing firms have an incentive to under-provide monitoring and penalties in the absence of binding contracts over their provision or sufficiently high reputational or legal costs.[17] As such, the question in this setting is whether MNCs' CSR enforcement increases monitoring and penalties enough to generate material improvements in factories' compliance.

If multinational enforcement increases compliance, a natural next question is how this affects workers' safety and well-being. Figure IV displays the Alliance's proposed causal chain from its training materials. The chain suggests that SCs affect these outcomes through increased knowledge and awareness, through improved relations with management, and through improved factory safety. To test these possibilities, I first test objective measures of workers' safety, including physical indicators of factory safety, workers' safety knowledge, and workers' visits to their factories' medical clinics. I then turn to subjective measures of workers' well-being. These measures may be positively affected

[17] Besley and Ghatak (2007)'s model of CSR is particularly relevant for this context. They model CSR as the joint production of a public good (or curtailment of a public bad) with a private good. In their model, firms in a competitive product market provide CSR as a profit-maximizing response to consumers who value the public good. CSR curtails the public bad at the same level as a voluntary contribution equilibrium, which is excessive compared to first-best, but a Pareto improvement in case of government failure. They also show that CSR has the greatest advantage (e.g., vis-à-vis NGOs) when curtailment of the bad is naturally bundled with the production of a private good. Their leading example is child labor in the manufacture of sneakers, which is analogous to the case of worker safety in the manufacture of apparel.

by safety improvements, but this does not directly follow if, for example, more stringent safety protocols impose costs on workers or if employers reduce other benefits in response to the pressure to implement SCs.

Next, I turn to the economic impacts on supplier factories. First, I examine effects on labor productivity. As displayed in Figure IV, there are multiple mechanisms through which implementation of SCs may improve labor productivity. For example, workers may make fewer visits to the medical clinic or miss work less often. It's unclear, though, whether SCs' productivity-enhancing effects outweigh their productivity-decreasing ones. As with broader OSH regulations, there is a concern that SCs reduce production speed and/or capacity, for example, due to stringent safety protocols and floor-plan requirements. Further, barring increases in working hours, SCs' meetings and activities with workers reduce the amount of time allocated toward production. In sum, it is ambiguous whether, on net, SCs positively or negatively affect labor productivity.

Second, I examine the effects on wages and employment. These depend on SCs' net costs and benefits to factories and on the extent to which labor markets are competitive. A primary concern is that well-intentioned enforcement of SCs may actually make workers worse off overall by adversely affecting wages and employment. Suppose that on net, SCs increase factories' costs and labor markets are competitive. In this case, employers lower wages, and if SCs' costs exceed workers' valuation of them, employment will eventually fall.[18,19] If employers have some power in the labor market, however, barring SCs' imposing extremely high costs to employers, wages and employment will not fall, and under less likely conditions, may actually increase (Manning, 2003).

Finally, I propose that the aforementioned effects may depend on firms' managerial

[18]This study was implemented from January 2017-December 2018. In relation to the study period, Bangladesh's minimum wage was increased to 5300 Bangladeshi Taka per month (US$63) in 2013. It was not increased again until December 2018, which was the final month of data collection for 11 study factories. There is reason to believe that the minimum wage was not binding during the study period.

[19]Wages may also directly fall if compensation includes production-based incentives. While it varies across factories, compensation often includes a base wage and some degree of production-based incentives. If the intervention lowers productivity, wages could be directly negatively impacted.

capacity. As described in Section 1, the existing literature on strengthening adherence to regulation in developing countries focuses on strengthening state-supplied enforcement to increase firms' incentives for compliance. I suggest that the potential for complementarity between state enforcement capacity and managerial capacity of the private sector could provide insights into persistently low levels of compliance in lower-income countries. I note that this analysis is not an aspect of the randomization; I test for heterogeneous treatment effects by suppliers' baseline managerial practices.

4 Research design

4.1 Randomized assignment to the SC Program

The 84-factory sample is drawn from the population of SC Program-eligible supplier factories. In order to be eligible, factories must have a separate committee that is formed in compliance with Bangladeshi labor law.[20] In most factories, this is the Participation Committee (PC), and it is responsible for appointing worker representatives to the SC.[21] PC worker representatives must be elected through free, fair, and competitive elections. The Alliance verified a factory's election process as part of determining its eligibility for the SC Program. Often, the brand(s) sourcing from the factory had to oversee a new election. Once the Alliance verified that the PC election was compliant, a factory became eligible.

The RCT was built into the roll out of the SC Program.[22] From January-December 2017, when the Alliance had a batch of eligible factories, it sent the list to me. Within batch, I randomly assigned 50% of factories to the treatment condition and 50% to the control condition. The result is a stratified randomized experiment with six strata, where each stra-

[20] If a factory has a trade union, then it selects the worker representatives to the SC. Few garments factories in Bangladesh have trade unions. In the 84-factory sample, only two have trade unions.

[21] PCs are legally-required for all factories with 50 or more workers outside of Export Processing Zones (EPZs). EPZ factories are subject to different labor laws. The Alliance implemented an analogous process with EPZ factories. The worker representation structure in EPZs is a Workers' Welfare Association (WWA).

[22] The Alliance rolls out all of its programs in phases, so from the experimental factories' perspective, it would not be apparent that the factory was part of a treatment or control group.

tum is a batch, and a total of 41 treatment factories and 43 control factories.[23] In 11 cases in which multiple factories shared ownership and location (building or compound), I randomly selected one factory to participate in the RCT.[24] All other factories at the same location were non-experimental but shared the assignment status of the randomly-selected factory. Online Appendix Table DI reports summary statistics for sample factories.

4.2 Data collection and measurement

This analysis uses three main sources of data. First, it uses several types of data collected during three separate, day-long visits to factories implemented over nearly one year. Second, it uses monthly production, human resource, and other business performance-related data collected using a retrospective questionnaire administered following the final data collection visit. Third, it uses numerous types of administrative data from the Alliance. The Alliance invited factories to cooperate with data collection. Factories were told that I was conducting a general impact evaluation and were not told that I was specifically interested in the SC Program.[25]

The onsite visits included three types of data collection: Surveys of stakeholders, document collection and verification, and spotchecks of safety conditions. Surveys included 20 randomly selected workers, the SC President, two randomly-selected SC worker representatives, the factory's most senior manager, and up to 20 randomly-selected lower-level managers. The document verification process entailed checking legally-required and Alliance-required documentation. It also included photographing records for digitization by the research team. At the second and third visits, a trained assessor visited the production floor to check physical safety conditions using a checklist. The team leader was an assessor, who was responsible for managing interactions with management, veri-

[23] All control factories were required to participate in the SC Program after completing the study period.
[24] A compound is a plot of land housing multiple factories at the same address.
[25] More information about communications with factories and the data collection protocol is available upon request.

fying documentation, and implementing the safety spotchecks. A junior assessor oversaw the the survey process, photographed records, and supported survey implementation. Three enumerators conducted surveys.

Figure I displays the experiment's timeline. The first visit established factories' baselines. The second visit, approximately five months later, aimed to measure outcomes immediately after treatment factories completed the most intensive portion of the SC Program. The third visit, approximately 10 months after baseline, aimed to measure outcomes a few months after treatment factories completed the SC Program. To minimize experimenter demand effects specific to the SC Program, the research team designed data collection protocols to minimize the risks of non-truthful reporting and manipulation. For example, in addition to SC-related documentation, the research team verified several other types of safety-related documentation.

4.2.1 Pre-specified primary outcomes

I analyze two groups of primary outcome variables. The first measures the intervention's effects on compliance and safety outcomes, while the second measures its economic effects. Beignning with the first group, the outcomes are:

1. Compliance with Bangladesh SC Regulation (index);
2. Safety indicators (index);

The first outcome is a standardized index variable that summarizes factories' compliance with the SC regulation. I use an index variable for this outcome because compliance with the SC Regulation is many-dimensional. The regulation includes three categories of requirements, including requirements for how SCs are formed, how they operate, and their responsibilities. Within each category, there is numerous stipulations (see Table I). To determine the variables included in the index, I enumerated the regulation's stipulations. Whenever relevant, I measure a factory's compliance with a stipulation using multiple sources information. For example, to determine how worker representatives to the SC

were selected, I combine reports from both the SC President (a member of management) and from SC worker representatives. Table AI lists the index's sub-variables.

The second outcome measures safety indicators that capture SCs' effectiveness at fulfilling the law's intent. The regulation prescribes responsibilities for SCs related to management of physical factory safety, to training workers, and to safety culture. Correspondingly, the index includes both physical and cultural indicators of safety. It is comprised of the following sub-indexes and unique variables (see Table AII for all sub-variables):

- Physical safety:
  - Performance on an independent spotcheck of factory safety conditions.
  - Progress with required building safety remediation based on Alliance building safety audits (Alliance "Corrective Action Plan (CAP)" completion).[26]
- Factory safety culture:
  - Workers' awareness of SC.
  - Workers' safety knowledge.
  - Senior managers' awareness of the SC.

In light of the wide-ranging safety oversight that the SC regulation assigns to the SC, I worked with an OSH expert to determine the spotcheck checklist. The expert helped me to identify critical items from a checklist for typical OSH audits of the factory floor that a trained social compliance assessor could check during a 30-minute floor visit.[27]

The ideal measure of SCs' effects on OSH would be injuries and illnesses. There are multiple reasons, though, why I do not use these as my primary outcome. First, the in-

[26]Every Alliance-audited factory had a Corrective Action Plan (CAP) based on violations found in the Alliance's building safety audit. The CAP detailed the remediation actions that the factory would take to address the safety violations. The Alliance monitored factories' remediation progress and suspended factories that failed to make sufficient progress.

[27]We excluded items that the SC could not plausibly influence within the study's duration. We also identified several items that the social compliance assessor would only check during the 9-10 month visit. The rationale for this approach was twofold: First, the OSH expert identified eight items that required more than 3-4 months, but plausibly less than 8-9 months, for the SC to address. Second, I wanted to reserve some factory spaces (e.g., bathrooms) to only be visited during the third visit. I aimed to gain insight into the extent to which management was responding to the research team's visits. Due to an administrative error, the eight items were not included in the third visit checklist for 14 out of 80 factories. As such, I depart from my PAP by not including these items.

tervention aimed to empower workers to raise safety issues and concerns. Consequently, on net, it could increase reported injuries and illnesses, even if it reduced their true rates. While this problem affects all sources of information about injuries and illnesses, I identify factories' medical clinic visitor records as the source that is most likely closest to the truth. I provide supporting evidence from analysis of these records.[28]

Turning to the economic primary outcomes, these include:

3. Workers' job satisfaction and mental well-being (index);
4. Labor productivity;[29]
5. Employment;
6. Wages.

The third outcome is an index variable that summarizes the enforcement interventions' effects on self-reported and revealed preference measures of workers' job satisfaction and mental well-being. I construct it using survey questions and administrative data on worker turnover and absenteeism. Figure AIII lists the index's sub-variables.

The final three primary outcomes measure the intervention's effects on factories' business competitiveness. Labor productivity is measured as the log of the physical quantity of output (e.g, pieces of clothing) per person-hour. Person-hours are calculated as number of workers times the average weekly working hours times 4 weeks per month plus the number of management-level employees times average weekly working hours for management staff times 4 weeks per month.[30] In six factories that produce multiple products, output is measured at the product-level. For these factories, I include the main product in the analysis and determine the share of labor allocated to this product using employee

[28] I am only able to analyze records for 62 factories. This is because numerous factories, in particular those located in EPZs and in multi-factory compounds, use centralized facilities and do not maintain factory-specific records. Among factories that do maintain records, their information content and legibility varies.

[29] I pre-specified that I would analyze total factor productivity or labor productivity. I indicated that I would analyze labor productivity if I determined that I could not measure non-labor inputs to production with sufficiently high quality. Ultimately, I decided that I could measure labor productivity for more factories and with less measurement error.

[30] Table DV presents treatment effects on the log of the physical quantity of output and average weekly working hours.

lists.[31] Employment is the total number of people employed at the factory in a month. Finally, wages are the log of gross wages paid to all employees in a month. These three outcomes are measured using administrative data provided by the factories.

To construct the index variables, I follow Casey, Glennerster and Miguel (2012) and Haushofer and Shapiro (2016) in using the methodology proposed by Anderson (2008) based on O'Brien (1984). The method entails an average of a family of variables that have each been oriented to be unidirectional, standardized, and weighted by the sum of its row in the inverse variance-covariance matrix calculated using the control group. The weighting maximizes the amount of information captured by the index, as it places less weight on highly correlated outcomes and more weight on less correlated outcomes. This approach is particularly well-suited for this study because, due to the staggered roll-out, I was not able to collect a complete baseline before committing to the construction of my indexes. Summary index variables also have the benefit of reducing the number of hypotheses being tested, which reduces the risk of overrejection of the null hypothesis. Finally, it increases my ability to detect marginally statistically significant effects on multiple outcomes that, aggregated, achieve statistical significance (Anderson, 2008).

I also pre-specified secondary outcome variables to explore possible mechanisms underlying the effects on primary outcome variables. Online Appendix Table DII displays index components for secondary index variables. Online Appendix Table DIII reports baseline balance tests, and Tables DIV and DV present results for workers and factories, respectively. When relevant, I reference these results to support interpretation.

In the interest of transparency, I report all deviations from my PAP and their rationales in Online Appendix Table DVI. Overall, I adhere very closely to my PAP.

[31] I determine a factory's primary product using quantities of physical output.

4.2.2 Econometric analysis

Regression models: I estimate the intervention's average treatment effects using two simple regression models. For the main analysis, I use the following regression model:

$$Y_j = \alpha + \beta\, T_j + \theta\, Y_{j,t=0} + \gamma_j + \epsilon_j \tag{1}$$

where Y_j is the outcome of interest for factory j. T_j is the treatment indicator, $Y_{j,t=0}$ is a control for the baseline value of the outcome variable. γ_j is a stratum indicator, and ϵ_j is the residual. β is the coefficient of interest. All statistical tests are two-sided.

To test for heterogeneous treatment effects, I use the following regression model:

$$Y_j = \alpha + \beta_1\, T_j + \beta_2\, R_j + \beta_3\, T_j * R_j + \theta\, Y_{j,t=0} + \gamma_j + \epsilon_j \tag{2}$$

where R_j is an indicator for above median baseline value of a pre-specified interaction variable. The notation for equation 2 is otherwise analogous to that for equation 1. In this specification, β_1 is the estimated treatment effect on factories with a below median baseline value of the interaction variable, $\beta_1 + \beta_3$ is the estimated treatment effect on factories with an above median baseline value of the interaction variable, and β_3 is the difference between these two effects. I report β_1 and $\beta_1 + \beta_3$ as well as the p-value for β_3.

For business competitiveness outcomes, I also show panel regression results in the Online Appendix. I use a panel regression model with five months of pre-intervention and five months of post-intervention data. I include factory and calendar fixed effects in certain specifications. I report the estimated coefficient on the interaction between an indicator for being in the treatment group and an indicator for being post-treatment.

Statistical inference: I use randomization inference, which is increasingly the recommended way to analyze data from RCTs, in particular for small samples (Athey and Imbens, 2016, Young, 2015, Heß, 2017). In addition to using summary index variables for multi-measure outcome categories, I show multiplicity-adjusted p-values. Across my

primary outcome variables, I control the False Discovery Rate (FDR), the expected proportion of rejections that are false positives. I report FDR-sharpened p-values for my preferred specification for all primary outcomes (Anderson, 2008). For index variables, I also show p-values adjusted to control the FDR across each variable's sub-indexes.

4.2.3 Integrity of the Experiment

Baseline Balance: Table II shows baseline balance between control and treatment groups. The sample size is indicated in each row. Certain variables are not available for all factories. In particular, factories that attrited from the sample did not provide their administrative data on business outcomes. Among non-attrited factories, five declined to provide production data, and eight declined to provide wage data. In sum, the randomization successfully generated two groups that are balanced along observable characteristics. There is one variable with a statistically significant difference at the 10% level, which is the proportion of randomly-selected surveyed workers who are female. This difference is not statistically significant among non-attrited factories.

Although the difference is not statistically significant, treatment factories' performance on the job satisfaction and well-being index is 0.11 sds lower than controls'. This gap is largely due to a treatment factory whose performance negatively deviates markedly from other factories' - its index value is more than 4 sds below the mean. There are many different methods for identifying and handling outliers (see Aguinis, Gottfredson and Joo (2013) for a review). I take a common approach, which is to present results including the outlier and to include an appendix that shows baseline balance and the main the results after dropping the outlier (Online Appendix E). The results are robust to dropping this factory and to controlling for the baseline value of the dependent variable.

Finally, turning to labor productivity, although not statistically significant, there is a qualitatively large difference between treatment and control factories. This difference is due to small differences in factory types between the groups. The treatment group has

somewhat more non-sewing factories (e.g., washing and accessories) that tend to be more capital intensive. For this reason, I also show that there are no differences in labor productivity between treatment and control factories that produce the same type of product.

Compliance: Three treatment factories did not receive treatment by the second data collection visit. One of these did not participate due to a critical member of management being on an extended leave at the time that the factory was due to begin. The other two factories are located in the Chittagong Region, where the Alliance implemented the SC Program in batches to ensure cost effectiveness, and it did not have a sufficient number of factories to implement it with these factories. Once we identified this issue, we resolved it for other factories that could have been impacted. A fourth factory began the SC Program less than two weeks before its second data collection visit. All other factories complied with the treatment. I address non-compliance by presenting Intent to treat (ITT) estimates. I also present a the Local Average Treatment Effect (LATE) estimates for primary outcome variables in Online Appendix DVII.

Attrition: Four factories, two treatment and two control, attrited from the sample. Three of the four were suspended by the Alliance for failure to make progress with physical building safety remediation. One control factory refused to participate in the second data collection visit. I address attrition by reporting Lee (2009) bounds on the treatment effects for primary variables with statistically significant treatment effects (Online Appendix Table DVIII). There is minimal difference between the upper and lower bounds of the treatment effects, and with the exception of the lower bound for the safety indicators index, all estimates are statistically significant at the 5% or 10% level.

5 Results

I present the intervention's effects in four sub-sections. First, I present the effects of multinational enforcement on factories' compliance and on safety indicators. Next, I assess its

effects on workers' job satisfaction and well-being and on factories' business competitiveness. Third, I discuss whether the treatment effects are heterogeneous across factories. Finally, I test for persistence after MNCs cease intensive enforcement.

5.1 Multinational enforcement , factories' compliance, and safety

5.1.1 Factories' compliance at baseline

Before a factory began the SC Program, the Alliance aimed to verify that its SC was formed correctly. Specifically, the Alliance conducted verification visits and worked with its members to verify that the bodies responsible for nominating worker representatives to the SC are democratically elected. When a factory began the SC Program, the Alliance checked that the SC was formed correctly again and reformed it if necessary.

For the purpose of this study, factories needed to be eligible for the SC Program in order to participate. Consequently, all except one factory in the sample had a SC at baseline, at least on paper. Factories were legally required to establish SCs by March 15, 2016; 20% of factories met this requirement. The median factory formed its SC in November 2016, although dates range from October 2015 to December 2017. Relative to its participation in baseline data collection, the median factory established its SC about 5 months prior. 73% of SCs were of the correct size and composition; among those that were not, issues included too few worker worker representatives, too few female worker representatives, and/or too few total members.[32] Despite these issues, there was high consistency between factory documentation and SC presidents' reports of SC size and composition ($\rho = 0.93$). Compliance was worse for requirements for democratic selection of worker representatives: 20% of SC presidents and 41% of worker representatives reported non-

[32]In one control factory, the SC was found to be comprised only of managers. In this case, compliance index outcomes related to correct formation of the committee are coded as non-compliant. At the second visit, the same factory provided the names of workers whom it indicated were members of the SC. Through the SC worker representative survey, it emerged that these workers were not members of the SC. Management had instructed them to participate because the composition of SC remained all managers. Again, the compliance index outcomes related to correct formation of the committee are coded as non-compliant.

compliant selection procedures (mainly, selection by management) or did not know how worker representatives had been selected.

In most factories, SCs were just becoming active. In 10% of factories, the SC had not yet met; in a further 16%, the SC had met once. 84% of factories SCs' had met at least once in the previous three months. Among SCs that had met, 89% maintained legally-required meeting minutes. In 77% of factories, there was no legally-required policy describing the SC's role and responsibilities. There was less consistency in the information about SCs' operations across different sources: Presidents' reports matched members' and factory documents in about 57% and 58% of cases, respectively. There were some reports of interference: In 10% of factories, at least one worker representative reported that management had offered bribes or attempted to block SC activities. 5% of presidents and 7% of worker representatives reported that they were not considered on duty for SC activities.

Many SCs were not implementing their legally-required safety responsibilities. For example, a key responsibility in the labor law is risk assessment. SCs are supposed to regularly inspect factories, to identify risks, and to develop an action plan for their resolution, including making recommendations to senior management. At baseline, only 15% of SCs had ever conducted a risk assessment. Relatedly, SCs are required to submit reports/recommendations on safety issues to senior management at least once per three months, which 73% of senior managers report receiving. SCs' reported fulfillment of other legally-required responsibilities varied. According to SC presidents, SCs were most likely to participate in fire prevention and preparedness activities (84%) and least likely to participate in accident investigation (55%).[33]

Does the SC Program actually increase monitoring?

In Section 3, I questioned whether the Alliance's SC Program results in increased monitoring. I can partially answer this question using the Alliance's administrative records.

[33] Although 44 SC presidents reported that the SC was responsible to investigate in case of an accident, only 7 indicated that the SC had actually participated in an accident investigation.

The most informative but costly form of monitoring is onsite audits. According to administrative records, during these audits, an Alliance staff member reviewed SC-related documentation, interviewed SC members and workers, and inspected the factory. The Alliance conducted audits during and after factories' participation in the SC Program. Among the treatment factories, 10% were audited during the SC Program and 15% were audited in the six months after completing the Program. Evidently, the SC Program entailed a material threat of being monitored through onsite audits.

5.1.2 The effects of multinational enforcement on compliance

Figure V and Table III present the results for the index of compliance with the SC regulation. Figure V compares treatment and control factories' performance on the compliance index pre-treatment and during the treatment phase. Both groups start off performing similarly. Control factories' compliance improves slightly but is mostly unchanged five months later. In contrast, treatment factories' compliance markedly improves under the MNCs' enforcement. Panel A of Table III shows that the ITT effect on compliance is 0.20-0.21 sds (FDR p=0.004). The MNCs' enforcement intervention increases factories' compliance with the SC regulation above and beyond the effects of state-supplied enforcement and of the MNCs' other compliance programs.

Panel B of Table III displays the results for the formation, operations, and responsibilities sub-indexes (Table AI lists index components).[34] While treatment factories outperform control factories on all sub-indexes, the largest treatment effect is on the SC responsibilities sub-index. Treatment factories outperform control factories on this sub-index by 0.44 sds (RI p=0.000, FDR p=0.002). The large, positive effect on this sub-index is consistent with the Alliance requiring treatment factories to complete several legally-required activities. For example, at the second visit, only 15% of control SCs had conducted a risk assessment, while 56% of treatment SCs had conducted at least one. According to reports

[34]Treatment and control factories are balanced on all sub-indexes at baseline (Online Appendix Table DIX).

by SC presidents, worker representatives, and senior managers, treatment SCs also made more regular reports and recommendations to senior management and followed up on these reports more often.

There are not statistically significant effects on the SC formation or operations sub-indexes. The lack of an effect on SC formation is perhaps unsurprising in light of the Alliance's engagement with factories prior to their becoming eligible for the SC Program. For the operations sub-index, both groups improve their performance between the baseline and second visits. The improvement is consistent with SCs' becoming active around the baseline visit. In the absence of the MNCs' enforcement intervention, SCs still would have become more incorporated into factories' OSH policies and procedures.

While the intervention does not significantly affect factories' compliance with the requirement that SCs meet at least once per quarter, it does increase their meeting frequency by 58%, from an average of 1.27 to 1.95 meetings per three months.[35] Further, an exploratory text analysis of SCs' meeting minutes reveals that, despite meeting with greater frequency, treatment SCs' meeting minutes are 23% longer (RI p=0.094) (Online Appendix Table DX). Given that control SCs' meeting minutes are, on average, less than half of a page of text, longer meeting minutes suggest more substantive discussions.

To provide more qualitative insights into changes in SCs' discussions, I examine the meeting minutes' textual content. I prepare the text using standard methods (see Gentzkow, Kelly and Taddy (2019)). First, I use Monte Carlo methods from topic modeling to identify the number of distinct topics in the minutes.[36] I expect that treatment SCs' discussions will be more specific, and consequently more diverse, as well as more action-oriented than controls'. Online Appendix Figure DI shows the densities for the number of topics identified using three common metrics. While pre-treatment measures are highly variable, across metrics during the treatment phase, all three methods consistently identify a

[35]Most SCs met with the minimum required frequency, 88% of control SCs and 93% of treatment SCs at the second visit.

[36]These methods often do not perform well with small text samples, hence I apply them by treatment group and phase.

larger number of topics in treatment SCs' minutes compared to controls'. Second, I represent discussion content in a form that is easier to interpret: Two-word phrases, known as bi-grams. Figure BI presents the top 15 bi-grams for treatment and control groups, respectively, prior to and during the enforcement intervention. There are a few interesting observations. First, treatment SCs' more substantive discussions are reflected in the count distribution's outward shift, which reflects the broader sets of bi-grams with higher frequency counts. Related to this point, general safety terms, such as "fire safety," disappear from treatment SCs' top bi-grams. Second, "risk assessment" is one of the most frequent bi-grams for post-treatment treatment group minutes but are absent for others. Third, decision terms such as "discussion decision" and "unanimous consent" appear for post-treatment treatment group minutes but are absent for others.

Together, the compliance results and the analysis of SCs' Meeting Minutes suggest that the enforcement intervention is increasing SCs' effort and engagement. This interpretation is further supported by an analysis of the treatment effects on workers' perception of SCs' compliance and effectiveness, which I measure using a pre-specified index. In column (2) of Table DIV, the first column shows that the intervention improves workers' perception of SC compliance and effectiveness by about 0.20 sds (RI p-val=0.094).[37]

It is worth noting that the MNCs' intervention does not achieve full compliance. For example, out of 36 treatment SCs that participated in the Program before the second data collection visit, 13 had not conducted a risk assessment. In a couple of cases, the research team determined that the factory falsified the record, and in a few others, managers had conducted risk assessments, but not the SC. In all 13 cases, the Alliance's program records show that the SC had conducted a risk assessment before the visit date. Evidently, the MNCs' monitoring of compliance is imperfect, either due to information frictions or to MNCs' own choice of how much monitoring to provide.

[37]Table DIII presents baseline balance tests for worker secondary outcome variables. While there is only one statistically significant difference between the treatment and control groups out of several variables, the treatment groups' means are mostly more negative. To minimize possible bias, I focus on results controlling for the baseline value of the dependent variable.

5.1.3 The effects of multinational enforcement on safety

The next critical question is whether increased compliance translates into improvements in factory safety. Figure VI and Table IV present the results for the index of safety indicators. As can be seen in Figure VI, treatment and control factories perform similarly at baseline. Under multinational enforcement, treatment factories again outperform the controls, in this case, by about 0.14 sds. Table IV shows that this difference is statistically significant at the 5% level for the RI p-values and is marginally statistically significant according to the FDR-adjusted p-values (FDR p=0.113). This result provides causal evidence that MNCs' interventions to increase compliance with safety labor laws can improve safety.

Figure DII illustrates support for the statistical extremeness of this result. The figure plots the joint distribution of treatment effects on compliance and on safety indicators under the null hypothesis. The actual parameter estimates, indicated in red, are one of the most extreme points in the joint distribution. The chance of jointly observing these effect sizes under the null hypothesis is extremely small.

Panel B of Table IV presents the treatment effects for each sub-index. Baseline balance tests for these sub-indexes are presented in Table DIX. There is one baseline imbalance, which is on worker awareness of SCs. Workers' awareness at treatment factories is lower, although this difference lessens and is not significant at the 5% level when the outlier factory is dropped (Online Appendix Table EI). Estimated treatment effects on this sub-index should be interpreted with appropriate caution.

The first row of Panel B shows that the treatment improves factories' performance on the safety spotcheck conducted by the research team. Treatment factories outperform controls on the safety spotcheck by 0.22 sds (RI p=0.015 and FDR p=0.083). Table V shows the treatment effects on each component of the spotcheck index.[38] Treatment factories outperform controls on nearly every sub-component. For example, workers in treatment

[38] Four spotcheck checklist variables drop from the analysis because all factories were found to comply.

factories are 9-18% more likely to be found using machines with appropriate guards for dangerous components and to be wearing required personal protective equipment (PPE) for their tasks.[39] Although none of the individual differences between treatment and control groups is statistically significant, aggregated, they indicate that the intervention has a small, positive effect on physical indicators of factory safety.

The safety improvement is consistent with the dramatic increase in treatment SCs' implementation of risk assessment. Table BI provides support for this interpretation. It shows the result of a two-stage least squares (2SLS) analysis in which I instrument for SC risk assessment using factories' assignment to treatment. Among factories whose SCs are induced to conduct a risk assessment by the intervention, risk assessment has a large, positive, statistically significant effect on their safety spotcheck performance.

Returning to Panel B of Table IV, the second row shows that the SC Program does not increase factories' progress on completing their CAPs for building safety violations. There are two likely reasons why. First, concurrently with the RCT's roll out, the Alliance intensified pressure on factories to complete their CAPs. Consequently, 25% of sample factories had completed 90% or more of required remediation actions by baseline. Second, the outstanding violations often required significant financial and time investment, and even if the MNCs' intervention increases SCs' ability to push for these investments, it may require more time for the effect to materialize.

Finally, the enforcement intervention does not have statistically significant effects on the safety culture sub-indexes (Panel B, rows 3-5). Workers' awareness of SCs increases compared to controls', but the difference is not statistically significant. For these outcomes, it is relevant that both treatment and control factories are required to participate in the Alliance's Fire Safety and Worker Helpline Training Program. This Program includes information about the factory's SC and likely helps to explain workers' high baseline level of awareness of SCs and the null result: At baseline, 81% of workers reported being aware

[39] PPE includes equipment such as eye guards, finger guards, gloves, goggles, and boots.

of SCs' role and responsibilities, and 89% knew that their factory had a SC. As shown in Table BII, even with this very high awareness, the intervention increases workers' awareness for both of these outcomes and for some other measures of awareness.

Visitors to medical clinics: I analyze visitor records from factories' medical clinics to test for effects on injuries and illnesses. This outcome is not one of my primary outcomes, and these results should be interpreted as suggestive. Records are available for 62 factories. The records were photographed during onsite visits for later digitization. Photographing and digitizing the records were time intensive, so social compliance assessors were instructed to photograph a sample of three to six days per month when records were large.[40] Due to a misunderstanding about how the records were to be used, assessors sometimes did not always photograph complete daily records. For this reason, and due to the aforementioned legibility issue, the number of days observed per factory varies. To address this issue, I present results with and without probability weights based on the number of days observed (pre-baseline).

I focus on the best-measured outcome available, which is the daily count of visitors to the medical clinic. Using monthly data on the number of employees, I calculate the proportion of the workforce that visits the clinic per day. Table II shows that treatment and control factories are balanced on this outcome. To smooth noise in the daily records, to address differences in the number of day-level observations per factory, and to present results at the same level as the business competitiveness outcomes, I average by month.[41]

Panel A of Table BIII presents the results. Column (1), which presents weighted results, shows that on average, 1.1% of the workforce visited the medical clinic per day in control factories. In treatment factories, however, the proportion of the workforce visiting the clinic is 15.4% lower (RI p=0.107). The unweighted estimate is very similar, a 15.5%

[40] Social compliance assessors were to photograph the fifth, fifteenth, and twenty-fifth days of month. If these days were weekends or holidays, they were instructed to photograph the preceding and following days.

[41] There is no statistical difference in the average number of observations for treatment or control factories, either at the day- or month-levels (results not reported).

decrease (RI p=0.135) (column (2)). The marginal statistical significance may be due to the smaller sample size. Further, in light of the possibility that the intervention may increase workers' willingness to report injuries and illnesses, the estimates may be upward biased. Together, though, the results provide suggestive evidence that the intervention reduces the workforce's need for medical attention by between 15-16%.

5.2 Multinational enforcement, workers' well-being, and factories' competitiveness

5.2.1 The effects of multinational enforcement on workers

Figure VII and Table VI present the results on workers' job satisfaction and well-being. Beginning with Figure VII, sub-figure (a) shows the full non-attrited sample and sub-figure (b) drops the negative outlier in the treatment group. Although the baseline difference including the outlier is not statistically significant, sub-figure (b) shows that the outlier does not drive the result. The Figure shows a statistically significant decrease in the job satisfaction and mental well-being index at treatment factories relative to controls. Turning to Table VI, the estimated treatment effect is -0.15 sds (RI p= 0.061; FDR p=0.113). The estimated treatment effect is unchanged when the outlier factory is dropped (Online Appendix Table EIV).

Panel B of Table VI displays the estimated treatment effects for the job satisfaction, mental well-being, turnover, and absenteeism sub-indexes/variables.[42] The results reveal that the negative effect on the primary index is driven by the job satisfaction sub-index (-0.39 sd effect, FDR p=0.075). The estimated treatment effects on mental well-being, turnover, and absenteeism are all negative, but they are smaller in magnitude and are not statistically significant.[43] Consistent with the null effect for turnover, the interven-

[42]Treatment and control factories are balanced on all sub-indexes at baseline (Online Appendix Table DIX).

[43]For inclusion in the index, the absenteeism and turnover sub-variables are constructed by collapsing five pre- and post-intervention monthly observations into one pre- and post-observation, respectively. They

tion does not affect workforce composition (Online Appendix Table DXI). Together, these results rule out the possibility that changes in workforce composition drive the negative effect.

To further unpack the negative effect on job satisfaction, Table BIV shows the estimated treatment effect on each variable in this sub-index. Panel A shows that the negative effect on job satisfaction is driven by an increase in the proportion of workers considering leaving their factory for safety-related reasons. While the effect is large, a 79% increase, it is on a control group mean of only 2.4% of workers.

Why does the enforcement intervention negatively affect some measures of workers' job satisfaction? In Section 5.3, I show that the effect is driven by factories with poor managerial practices where the intervention does not improve compliance or safety. I provide suggestive evidence that it may be related to the intervention raising workers' expectations about what SCs will deliver, and SCs' actual performance not meeting these expectations. I have also checked for evidence of other mechanisms, such as workers learning about unsafe conditions. The data do not provide evidence in favor of learning about unsafe conditions driving the negative effect (results not reported).

5.2.2 The effects of multinational enforcement on factories' business competitiveness

In this section, I analyze the intervention's effects on labor productivity, gross wages, and employment. There is one control factory that partially or fully suspends production over three months between the first and second data collection visits.[44] Because this type of temporary suspension is part of business, it does not mean that this factory should be removed from the analysis. But due to the timing of the partial shut down and my smaller sample size, my results may be sensitive to its inclusion. I present results for the full sample, for the full sample trimming the 1st and 99th percentiles of observations, and for the

are then multiplied by -1 in order to be unidirectional with other outcomes.

[44]The factory does not dramatically cut employment, and gross wages do not dramatically fall during the three month period. Employment and wage results are very similar when this factory is dropped. Other main results are also unchanged.

sample dropping the factory that partially shuts down.

Panel A of Table VII shows the estimated treatment effects on labor productivity using the main regression model (equation 1).[45] All three specifications include product-type fixed effects. In column (1), which includes the full sample, the estimated treatment effect is large and positive, a 11.5% increase. In column (2), which includes the trimmed sample, the estimated effect remains positive but is smaller, an 8.7% increase. Finally, in column (3), which includes all observations except those for the factory that partially shuts down, the estimated effect is a 3.6% increase. None of these estimated effects is statistically significant. The large decrease in the estimated effect when the distribution is trimmed or the partial shutdown factory is dropped supports the interpretation that the intervention did not affect labor productivity.[46] In light of the null results, Table BV reports the ex post minimum detectable effect size (MDE) that would be detectable ex post under standard assumptions for power calculations (80% power and 5% statistical significance level). I am underpowered to detect small to moderate effects.

Panel B of Table VII presents the estimated treatment effect on gross wages and on employment. Column (1) shows the estimated treatment on wages, which is a -1.5% decrease in wages (RI p=0.612). Turning to employment, the estimated treatment is a -1.1% decline in employment (RI p=0.635). Again, Table BV reports the MDEs for both variables. I am powered to detect moderate effects on employment and wages.

Taken together, the estimated treatment effects suggest that the intervention improves safety without adversely affecting labor productivity, employment, or wages. These results are consistent with the intervention imposing only small costs on suppliers with possible benefits to labor productivity. The results support the potential for MNCs' enforcement interventions to improve labor standards without coming at significant costs in terms of suppliers' efficiency.

[45] Table DV presents effects on physical output and on working hours; the estimated treatment effects on both are close to zero and are not statistically significant.

[46] Treatment effect estimates for all business competitiveness variables are very similar using a panel regression model (Online Appendix Table DXII).

5.3 Heterogeneous treatment effects by managerial capacity

In this section, I explore heterogeneity in the intervention's effects depending on factories' baseline managerial capacity. In my PAP, I specified three other dimensions of heterogeneity to explore: Factory size, compliance with the SC regulation, and location in an EPZ. I find the most compelling pattern of heterogeneous treatment effects for managerial practices, so I present the results for the other dimensions of heterogeneity in the Online Appendix (Tables DXIV, DXV, and DXVI).[47]

My measure of managerial practices is a variable that summarizes senior and lower-level managers' reported frequency of holding production-related meetings with workers.[48] This question is a variant of questions asked in the World Management Survey (WMS) and in studies on managerial practices by Bloom et al. (2013) and Macchiavello et al. (2015).[49] It measures one specific managerial practice, as it was not feasible to conduct a complete management diagnostic. As such, one may question whether this measure reflects broader managerial capacity. Figure BII provides evidence that it does. The Figure presents a binned scatterplot that includes all apparel manufacturers from all countries included in the WMS. It shows that apparel firms' score on the WMS's meeting question is highly correlated with their average overall WMS Management Index excluding this question.[50] Evidently, this question captures meaningful information about firms' overall managerial practices.

[47] For space reasons, I omit HTE analysis for these dimensions for business competitiveness variables. For location in an EPZ, as Online Appendix Table DXIV shows, there are large differences between the seven treatment and the seven control factories in EPZs. For this reason, I depart from the PAP and do not analyze this dimension. Results for both analyses are available upon request.

[48] The measure places 25% weight on the factory's most senior manager's report and 75% weight on the lower-level managers' reports. On average, 15 lower-level managers were surveyed.

[49] In Bloom et al. (2013)'s study of managerial practices in Indian textile factories, "Daily meetings to discuss efficiency with the production team" is one of the management practices that they consider. 5% of factories implement this practice at baseline. This practice has one the highest adoption rates among factories randomly assigned to receive management consulting services (70% adopt). Several years later, Bloom et al. (2018) find that these meetings are also one of the stickiest practices: 80% of treatment factories implement the meetings.

[50] The WMS question asks whether performance is reviewed with appropriate frequency and communicated to staff ((World Management Survey, n.d.)). The WMS Management Index is the average score on all other questions.

I partition the sample into above/below median groups using the baseline value of managerial practices. I refer to the below median group as poorly-managed and to the above median group as better-managed. Online Appendix Table DXIII shows baseline balance within each interaction-term subgroup for primary outcome variables for non-attrited factories. There are no statistically significant differences between subgroups.

Panel A of Table VIII shows the main results. Each column considers a different primary outcome variable. In this and other tables that present results on heterogeneous treatment effects, the first row of the panel displays the estimated treatment effect for the below median group, and the second row displays the treatment effect for the above median group. The final row displays the p-value of the difference in the treatment effects on the subgroups. The regression specification is equation 2.

Beginning with column (1), the estimated treatment effect on SC compliance for poorly-managed factories is 0.11 sds (not statistically significant). In contrast, the estimated effect on better-managed factories is 0.31 sds (RI $p \approx 0.000$). The difference between these estimates approaches marginal statistical significance (RI p=0.133). For poorly-manged factories, the small improvement in compliance translates into a very small improvement in safety indicators (0.05 sd effect, RI p=0.566). In contrast, consistent with the large effect on compliance, better-managed factories improve safety indicators by 0.24 sds (RI p=0.049). Again, though, the difference between the estimates is not statistically significant. The pattern of results suggests that the enforcement intervention leads to large improvements in factories' compliance and in SCs' effectiveness, but only among factories with better baseline managerial practices.

Turning to column (3), again, there is a stark difference in the estimated effects on job satisfaction and mental well-being for poorly- and better-managed factories. The estimated treatment effect on job satisfaction and mental well-being for poorly-managed factories is -0.27 sds (RI p=0.027). The estimated effect for better-managed factories, though, is close to zero (-0.030 sd effect, RI p=0.754)). The RI p-value for the difference in the treat-

ment effects for these groups is p=0.145. Evidently, the decline in job satisfaction found in Section 5.2.1 is driven by poorly-managed factories, for which the intervention does not result in meaningful improvements in compliance or in safety.

I can try to increase statistical power to detect differences between the treatment effects for poorly- and better-managed factories by pooling the treatment and post-treatment rounds of data. It is not clear that this approach will help because the treatment effects for each subgroup may exhibit different dynamics. Panel C of Table VIII displays the results of pooling the treatment and post-treatment rounds. Column (1) shows that the estimated treatment effects on compliance for above and below median managerial practice factories remain stable. I can not, however, reject that treatment effects for above and below median groups are equal (RI p=0.173). Turning to column (2), the estimated treatment effects on safety indicators, although attenuated, exhibit the same pattern. I remain unable, however, to reject that the effects for both groups are equal (RI p=0.332). Finally, in column (3), the difference in effects is fairly stable, and I reject the null of equality with RI p=0.093. In sum, while not conclusive, the results support the interpretation that MNCs' enforcement has differential effects on poorly- and better-managed factories.

Next, I examine whether the pattern of heterogeneous treatment effects is present for visitors to factories' medical clinics. Panel B of Table BIII presents the results. In both columns, the pattern of estimates is consistent with the decline in visitors being driving by better-managed factories. While I am underpowered to detect differences in the treatment effects, these results are consistent with the other evidence presented on the important role of managerial practices.

Finally, I test for heterogeneous treatment effects on business competitiveness outcomes. Table BVI presents the results. There is not strong evidence of heterogeneous effects. The results allay concerns that the null overall treatment effects mask negative treatment effects on these outcomes for better-managed factories.

Robustness checks for heterogeneity results

There is correlation in factories' characteristics: Better-managed factories tend to be somewhat larger and less compliant. These correlations raise the possibility that only one of these characteristics is actually important in determining the intervention's effects. To examine this possibility, I regress each outcome on the treatment indicator, an indicator for each dimension of heterogeneity, and interactions between each dimension and the treatment indicator. This specification demands a lot of the data, but it provides qualitative insight into the relative importance of each dimension. Table BVII presents the results. In column (1), in which compliance is outcome, the only interaction term that is large in magnitude and statistically significant at the 10% level is above median managerial practices (RI p=0.087). In column (2), in which safety indicators is outcome, the above median managerial practices interaction term is again largest in magnitude (RI p=0.155). Finally, in column (3), in which job satisfaction and mental well-being is the outcome, managerial practices remains an important dimension after controlling for other dimensions of heterogeneity (RI p=0.068). Together, the results show that management practices are an important dimension of heterogeneity after controlling for factories' other characteristics and their interaction with the treatment.

Another concern is that MNCs may more intensively monitor less compliant factories and that this generates the heterogeneous effects. In this case, one would expect the Alliance to be more likely to audit factories that, at baseline, are less compliant with the SC law. The Alliance audited five treatment factories during the study period, but all of the audits occurred after the 4-5 month data collection visit. As such, differential auditing could not drive the heterogeneous effect patterns in Panel A of Table VIII.

Finally, I implement the analysis using an alternative measure of management practices. This measure captures a different dimension of managerial capacity: Human Resource (HR) management. I measure HR practices using an index of worker-reported HR practices and relations with managers that I pre-specified as a secondary outcome vari-

able (Online Appendix Table DII lists index components)). I find a qualitatively similar pattern of heterogeneous treatment effects using this variable as with my main managerial practices measure. See Online Appendix Tables DXIV, DXV, and DXVI.

To summarize, the heterogeneous treatment effect results show that managerial capacity plays an important role in determining the MNCs' enforcement intervention effects. The intervention improves compliance and safety indicators only in factories with better managerial capacity. These improvements do not come at the cost of negative effects on workers' job satisfaction and well-being. For factories with poor management practices, however, the intervention does not improve compliance or safety-related outcomes and has a negative effect on indicators of workers' job satisfaction.

Why do indicators of job satisfaction decline at poorly-managed treatment factories?

Why are indicators job satisfaction at poorly-managed factories negatively affected when the intervention has little to no effect on compliance and safety indicators? One plausible mechanism is that the MNCs' intervention raises workers' expectations for SCs, but in poorly-run factories, workers' expectations are not met, and they are disappointed. As discussed in Section 1, this effect would be consistent with recent experimental findings by Adhvaryu, Nyshadham and Xu (2018).

I cannot directly test for this possibility, as I did not collect data on workers' expectations for SCs. I find support for an important role for workers' expectations and for learning, though, from qualitative evidence from eight interviews with compliance managers at treatment factories.[51] Multiple managers reported that it took several months after their factory's SC became active for workers to understand what issues they could report to the SC and expect to have resolved. In particular, managers reported that it was initially common for workers to raise issues to the SC that were outside of its authority (e.g., working hours or wage-related concerns). If these concerns go unresolved, it is easy

[51] I selected two treatment factories per quartile of management practice scores for interviews. Selection was not done randomly but based on factories' proximity to Dhaka and managers' availability.

to see why workers may be disappointed. It is also unsurprising that workers may not initially understand the scope of SCs' authority, as the SC Program is possibly the first time that workers have been informed that there is an institution inside the factory with democratically-selected worker representatives that is responsible to address workers' concerns.[52] I do find that workers' perception of SCs' effectiveness remains unchanged in poorly-managed treatment factories and only improves in better-managed factories (results not reported). If unmet expectations play a role, it suggests that the negative effect on indicators of job satisfaction at at poorly-managed factories may be temporary. I test this possibly using the third data collection round.

5.4 Do the effects persist after multinationals' cease intensive enforcement?

The Alliance's SC Program aims to bring factories into compliance with the law through intensive enforcement for a period of six months. The Alliance then continues to monitor factories under its general monitoring activities. Do factories maintain improvements in compliance and safety indicators after the MNCs cease intensive enforcement?

Table IX presents the estimated treatment effects on the first three primary outcomes measured 3-4 months after the end of the intensive enforcement period for treatment factories. The first row of Column (2) shows that the estimated treatment effect on compliance persists. The effect remains around 0.21 sds (FDR p=0.078). Treatment factories continue to outperform controls on the responsibilities sub-index and also surpass controls on the operations sub-index (results not reported).

While treatment factories continue to significantly outperform controls on the compliance index, they no longer do on the safety indicators index. Treatment factories perform 0.06 sds better than controls (RI p=0.371). The difference in treatment and control fac-

[52] I do not have data with which I can test whether workers raise non-safety issues to SCs. I have access to records of issues raised to SCs, but SCs only recorded the safety-related issues in these documents. Further, we did not ask workers about the specific issues that they raised to the SC.

tories' performance on the research team's spotcheck of safety conditions is no longer statistically significant.

While treatment factories' performance on the safety indicators index attenuates slightly, control factories' performance also improves. I argue that the convergence is likely due to control factories' expectations about future enforcement by the Alliance. In particular, the Alliance rolls out its programs in a staggered fashion, and factory managers generally know that they will be required to participate. For example, in a survey that I conducted with compliance and HR managers of factories under shared-ownership with experimental factories, 46% of managers whose factories had participated in the SC Program report taking actions in advance to prepare.[53]

Interestingly, the negative effect on the job satisfaction and mental well-being index disappears, and the estimated treatment effect is actually positive (0.12 sd effect, RI p=0.206). Panel B of Table VIII shows that the effect for poorly-managed factories is zero, while the effect for better-managed factories is 0.24 sds (RI p=0.065), although the estimates are not statistically different. This pattern of results is consistent with the possibility that the short-run negative effect for poorly-managed factories is due to disappointment. For better-managed factories, the positive estimate is consistent with workers' valuing SCS, but it requiring time for them to learn about SCs' role or for SCs' benefits to materialize.

Together, the results suggest that under less intensive monitoring, treatment factories maintain compliance improvements that are arguably more easily observable. Less observable improvements in safety indicators, however, attenuate slightly. Control factories, possibly expecting future enforcement by the MNCs, also begin to improve compliance and safety indicators. One reason why treatment factories may maintain compliance improvements while safety indicators improvements attenuate slightly is that that treatment factories are subject to continued audits that are likely to reveal more observable aspects

[53]Further, the Alliance agreed to delay controls' participation in the SC Program until the end of the data collection period. No control factories were treated during the experiment, but it's possible that Alliance personnel communicated to control factories that they would eventually be required to participate.

of compliance.

Finally, I do not find evidence of delayed adverse effects on labor productivity, employment, or wages. The effects for all variables are close to zero. See Table DXVII.

6 Conclusion

In this paper, I analyze the effects of a coalition of MNCs' CSR program to enforce a local labor law on their Bangladeshi suppliers. This study is a "first" in multiple streams of literature. It is the first study to provide experimental evidence on whether firms' CSR programs generate meaningful social benefits, which in this case is improved labor standards in GVCs. It is also the first study to provide experimental evidence on the effects of enforcing labor regulation on factories' competitiveness and workers' well-being. Further, it is the first study to experimentally intervene to increase collective worker voice inside the firm.

I find that the MNCs' enforcement intervention is successful at increasing factories' compliance with Bangladesh's labor law. Specifically, their intervention to enforce a labor law that mandates worker-manager SCs improves compliance and indicators of factory safety. There is evidence that the treatment effect on compliance persists beyond the period of intensive enforcement by multinational buyers. These findings demonstrate that MNCs' enforcement interventions can meaningfully improve working conditions and contribute to achieving the law's objectives.

I do not find evidence of negative effects on labor productivity, employment, and wages. Together, these results help to allay concerns that enforcement of labor regulation necessarily entails trade-offs between competitiveness and improved working conditions. Further, they can help economists to update their views on enforcement of labor regulation and economic outcomes in developing countries.

Finally, I find that providing firms with incentives to comply with the SC law is not

sufficient to increase compliance but that firms' capacity to respond to incentives also matters. In particular, the estimated treatment effects on compliance and safety are large for better-managed factories but are small for poorly-managed factories.

My findings raise several important directions for future research. First, this research suggests that *private* sector capacity plays an important role in determining the efficacy of *public* regulation. Future research can more fully investigate how firms' capacity supports compliance with labor and other regulations. Second, in the short-run, I do not find evidence of workers differentially sorting in response to improvements in firms' compliance. It is possible, though, that if improvements in compliance are sustained, it may affect workers' mobility and sorting into factories. Boudreau, Heath and McCormick (2020) provide evidence that garment workers who begin their careers with poor information about factories' working condition exhibit a revealed preference for improving their working conditions compared to their wages. More research is needed, however, on how workers in developing countries trade-off between wages and non-pecuniary amenities. Third, a critical question is what the general equilibrium effects of multinational enforcement of labor law are on compliance and competitiveness of the targeted sector. Finally, there is generally a dearth of empirical evidence in economics on the welfare effects of firms' CSR activities. CSR programs, including private enforcement programs and other types of programs, are becoming increasingly common and increasingly large-scale. These interventions merit more attention.

References

Addison, John T., Claus Schnabel, and Joachim Wagner. 2001. "Works councils in Germany: their effects on establishment performance." *Oxford Economic Papers*, 53(4): 659–694.

Adhvaryu, Achyuta, Anant Nyshadham, and Huayu Xu. 2018. "Hostel Takeover: Living Conditions, Reference Dependence, and the Well-being of Migrant Workers." *Working Paper*.

Aghion, Philippe, Robin Burgess, Stephen J Redding, and Fabrizio Zilibotti. 2008. "The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India." *The American Economic Review*, 98(4): 1397–1412.

Aguinis, Herman, Ryan K Gottfredson, and Harry Joo. 2013. "Best-Practice Recommendations for Defining, Identifying, and Handling Outliers." *Organizational Research Methods*, 16(2): 270–301.

Alliance for Bangladesh Worker Safety. 2015. "Protecting and Empowering Bangladesh's Garment Workers: Second Annual Report." Dhaka, Bangladesh.

Almunia, Miguel, Jonas Hjort, Justine Knebelmann, and Lin Tian. 2020. "Strategic or Confused Firms? Evidence from "Missing" Transactions in Uganda." *Working Paper*.

Amengual, Matt, and Greg Distelhorst. 2019. "Can Sourcing Help Enforce Global Labor Standards? Evidence from the Gap Inc Supply Chain." *Working Paper*.

Amirapu, Amrit, and Michael Gechter. forthcoming. "Indian Labor Regulations and the Cost of Corruption: Evidence from the Firm Size Distribution." *Review of Economics and Statistics*.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481 –1495.

Arun, Rajkumar, V Suresh, Veni Madhavan, and Narasimha Murthy. 2010. "On finding the natural number of topics with latent dirichlet allocation: Some observations." In *Advances in knowledge discovery and data mining*. , ed. Mohammed J Zaki, Jeffrey Xu Yu, Balaraman Ravindran and Vikram Pudi, 391–402. Berlin Heidelberg,:Springer.

Ashraf, Nava, Oriana Bandiera, and B. Kelsey Jack. 2014. "No margin, no mission? A field experiment on incentives for public service delivery." *Journal of Public Economics*, 120: 1–17.

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S Lee. forthcoming. "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the delivery of Public Services." *American Economic Review*.

Athey, Susan, and Guido Imbens. 2016. "The Econometrics of Randomized Experiments."

Bai, Jie, Ludovica Gazze, and Yukun Wang. 2019. "Collective Reputation in Trade: Evidence from the Chinese Dairy Industry." *NBER Working Paper No. 26283*.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 169–217.

Bénabou, Roland, and Jean Tirole. 2010. "Individual and Corporate Social Responsibility." *Economica*, 77: 1–19.

Besley, Timothy, and Maitreesh Ghatak. 2007. "Retailing public goods: The economics of corporate social responsibility." *Journal of Public Economics*, 91: 1645–1663.

Besley, Timothy, and Robin Burgess. 2004. "Can Labor Regulation Hinder Economic Performance? Evidence from India." *The Quarterly Journal of Economics*, 119(1): 91–134.

Bhattacharjee, Nivedita. 2013. "Gap willing to sign Bangladesh safety accord, if tweaked." *Reuters*.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie, and John Roberts. 2018. "Do Management Interventions Last? Evidence from India." *Working Paper*.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *The Quarterly Journal of Economics*, 128(1): 1–51.

Botero, Juan C, Simeon Djankov, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. 2004. "The Regulation of Labor." *The Quarterly Journal of Economics*, 119(4): 1339–1382.

Boudreau, Laura, Rachel Heath, and Tyler McCormick. 2020. "Migrants, Information, and Working Conditions in Bangladeshi Garment Factories." *Working Paper*.

Burbano, Vanessa C. 2016. "Social Responsibility Messages and Worker Wage Requirements: Field Experimental Evidence from Online Labor Marketplaces." *Organization Science*, 27(4): 1010–1028.

Callen, Michael, Saad Gulzar, Ali Hasanain, Yasir Khan, and Arman Rezaee. 2018. "Strengthening state capabilities: The role of financial incentives in the call to public service." *NBER Working Paper No. 21180*.

Cao, Juan, Tian Xia, Jintao Li, Yongdong Zhang, and Sheng Tang. 2009. "A density-based method for adaptive LDA model selection." *Neurocomputing*, 72(7-9): 1775–1781.

Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *Quarterly Journal of Economics*, 127: 1755–1812.

Central Intelligence Agency. 2016. "The World Factbook." Central Intelligence Agency.

Dal Bó, Ernesto, and Frederico Finan. 2016. "At the intersection: A review of institutions in economic development." *Working Paper*.

Dal Bó, Ernesto, Frederico Finan, and Martín A Rossi. 2013. "Strengthening state capabilities: The role of financial incentives in the call to public service." *The Quarterly Journal of Economics*, 128(3): 1169–1218.

DiNardo, John, and David S Lee. 2004. "Economic Impacts of New Unionization on Private Sector Employers: 1984–2001." *The Quarterly Journal of Economics*, 119(4): 1383–1441.

Donaghey, Jimmy, and Juliane Reinecke. 2018. "When Industrial Democracy meets Corporate Social Responsibility – A Comparison of the Bangladesh Accord and Alliance as responses to the Rana Plaza disaster." *British Journal of Industrial Relations*, 56(1): 14–42.

Dragusanu, Raluca, and Nathan Nunn. 2018. "The Effects of Fair Trade Certification: Evidence from Coffee Producers in Costa Rica." *Working Paper*.

Dragusanu, Raluca, Daniele Giovannucci, and Nathan Nunn. 2014. "The Economics of Fair Trade." *Journal of Economic Perspectives*, 28: 217–236.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. "Truth-Telling by Third Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." *The Quarterly Journal of Economics*, 128(4): 1499–1545.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2014. "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India." *Working Paper*.

Fishback, Price V, and Shawn E Kantor. 1996. "Did Workers Pay for Workers Compensation Laws?" *The Quarterly Journal of Economics*, 110(3): 713–742.

Fisman, Raymond, and Yongxiang Wang. 2015. "The Mortality Costs of Political Connections." *Review of Economic Studies*, 82: 1346–1382.

Friedman, Milton. 1970. "The Social Responsibility of Business is to Increase its Profits." *The New York Times Magazine*, 1–19.

Gentzkow, Matthew, Bryan T Kelly, and Matt Taddy. 2019. "Text as Data." *Journal of Economic Literature*, 57(3): 535–574.

Government of Bangladesh. 2015. "Bangladesh Gazette: Ministry of Labor and Employment Notification." *English translation*, S.R.O. No. 291.

Greenhouse, Steven. 2013*a*. "Obama to Suspend Trade Privileges with Bangladesh." *The New York Times*.

Greenhouse, Steven. 2013*b*. "U.S. Retailers See Big Risk in Safety Plan for Factories in Bangladesh." *The New York Times*.

Griffiths, Thomas L, and Mark Steyvers. 2004. "Finding scientific topics." *Proceedings of the National Academy of Sciences*, 101: 5228–5235.

Hainmueller, Jens, Michael J Hiscox, and Sandra Sequeira. 2015. "Consumer demand for fair trade: Evidence from a multistore field experiment." *The Review of Economics and Statistics*, 97: 242–256.

Harrison, Ann, and Jason Scorse. 2010. "Multinationals and anti-sweatshop activism." *The American Economic Review*, 100(1): 247–273.

Hart, Oliver, and Luigi Zingales. 2017. "Companies Should Maximize Shareholder Welfare Not Market Value." *Journal of Law, Finance, and Accounting*, 2: 247–274.

Haushofer, Johannes, and Jeremy Shapiro. 2016. "The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya." *Quarterly Journal of Economics*, 1973–2042.

Heath, Rachel, and Ahmed Mushfiq Mobarak. 2015. "Manufacturing growth and the lives of Bangladeshi women." *Journal of Development Economics*, 115: 1–15.

Heß, Simon. 2017. "Randomization inference with Stata: A guide and software." *Stata Journal*, 17(3): 630–651.

International Labor Organization. 2016. "Sectoral Studies on Decent Work in Global Supply Chains." International Labor Organization.

International Labor Organization. 2017. "Towards safer working conditions in the Bangladesh ready-made garment sector." International Labor Organization.

Jäger, SImon, Benjamin Schoefer, and Jörg Heining. 2020. "Labor in the Boardroom." *Working Paper*.

Kitzmueller, Markus, and Jay Shimshack. 2012. "Economic Perspectives on Corporate Social Responsibility." *Journal of Economic Literature*, 50: 51–84.

La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Schleifer, and Robert Vishny. 1999. "The Quality of Government." *Journal of Law, Economics, & Organizations*.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, 76(3): 1071–1102.

Lee, David S., and Alexandre Mas. 2012. "Long-Run Impacts of Unions on Firms: New Evidence from Financial Markets, 1961–1999 *." *The Quarterly Journal of Economics*, 127(1): 333–378.

Macchiavello, Rocco, and Ameet Morjaria. 2015. "The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports." *American Economic Review*, 105: 2911–2945.

Macchiavello, Rocco, and Josepa Miquel-Florensa. 2019. "Buyer-Driven Upgrading in GVCs: The Sustainable Quality Program in Colombia." *Working Paper*.

Macchiavello, Rocco, Andreas Menzel, Atonu Rabbani, and Christopher Woodruff. 2015. "Challenges of Change: An Experiment Training Women to Manage in the Bangladeshi Garment Sector." *Working Paper*.

Manning, Alan. 2003. *Monopsony in Motion : Imperfect Competition in Labor Markets*. Princeton University Press.

McKinsey & Company. 2011. "Bangladesh's ready-made garments landscape: The challenge of growth." McKinsey & Company.

Ministry of Labour and Employment. 2013. "National Tripartite Plan of Action on Fire Safety and Building Integrity in the Ready-Made Garment Sector in Bangladesh." Ministry of Labour and Employment, Dhaka, Bangladesh.

Munni. 2017. "Half of apparel units flout needs for safety committees." *The Financial Express*.

O'Brien, Peter C. 1984. "Procedures for Comparing Samples with Multiple Endpoints." *Biometrics*, 40(4): 1079–1087.

O'Rourke, Dara. 2014. "The Science of Sustainable Supply Chain." *Science*, 344(6188): 1124–1127.

Solidarity Center. 2016. "Bangladesh Garment and Textile Factory Safety Incidents Since Tazreen Fashion Factory Fire, November 24, 2012." Solidarity Center.

Tanaka, Mari. forthcoming. "Exporting Sweatshops? Evidence from Myanmar." *Review of Economics and Statistics*.

The Economist. 2013. "Accord, alliance, or disunity?" *The Economist*.

World Bank. 2020. "World Development Report 2020: Trading for Development in the Age of Global Value Chains." World Bank, Washington DC.

World Management Survey. n.d.. "World Management Survey Manufacturing Instrument." https://worldmanagementsurvey.org/survey-data/methodology/.

World Trade Organization. 2017. "World Trade Statistical Review 2017." World Trade Organization, Geneva, Switzerland.

Yao, Yang, and Ninghua Zhong. 2013. "Unions and Workers' Welfare in Chinese Firms." *Journal of Labor Economics*, 31(3): 633–667.

Yassi, Annalee, Karen Lockhart, Mona Sykes, Brad Buck, Bjorn Stime, and Jerry M Spiegel. 2013. "Effectiveness of Joint Health and Safety Committees: A Realist Review." *American Journal of Industrial Medicine*, 56: 424–438.

Young, Alwyn. 2015. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *Working Paper*.

7 Figures and Tables

Figure I: RCT timeline

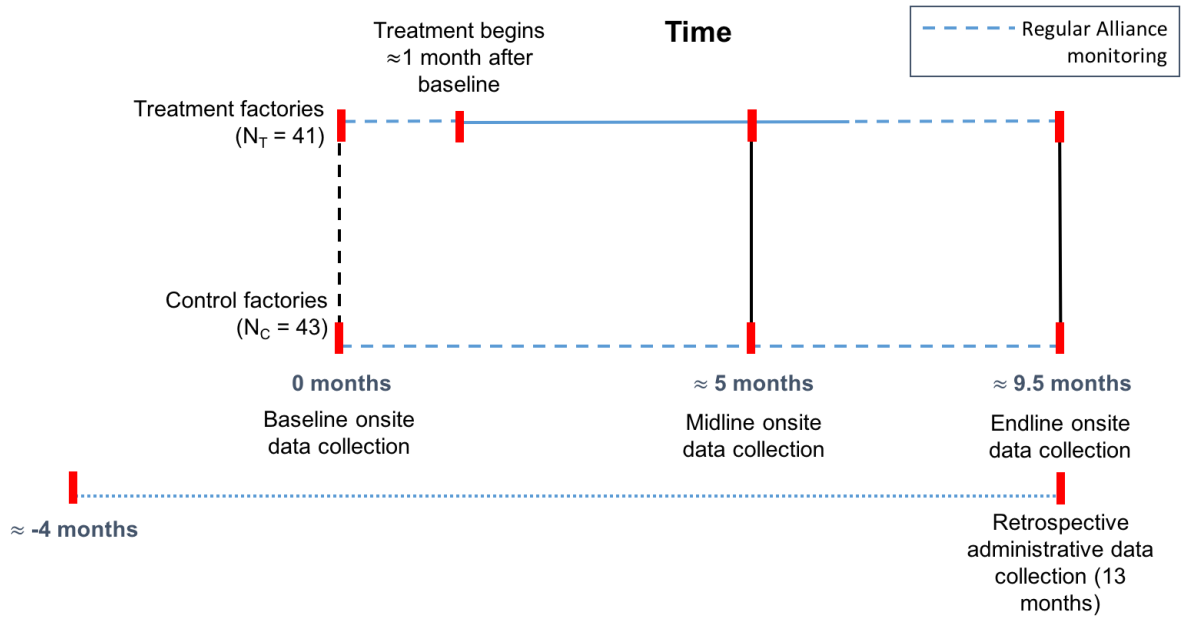
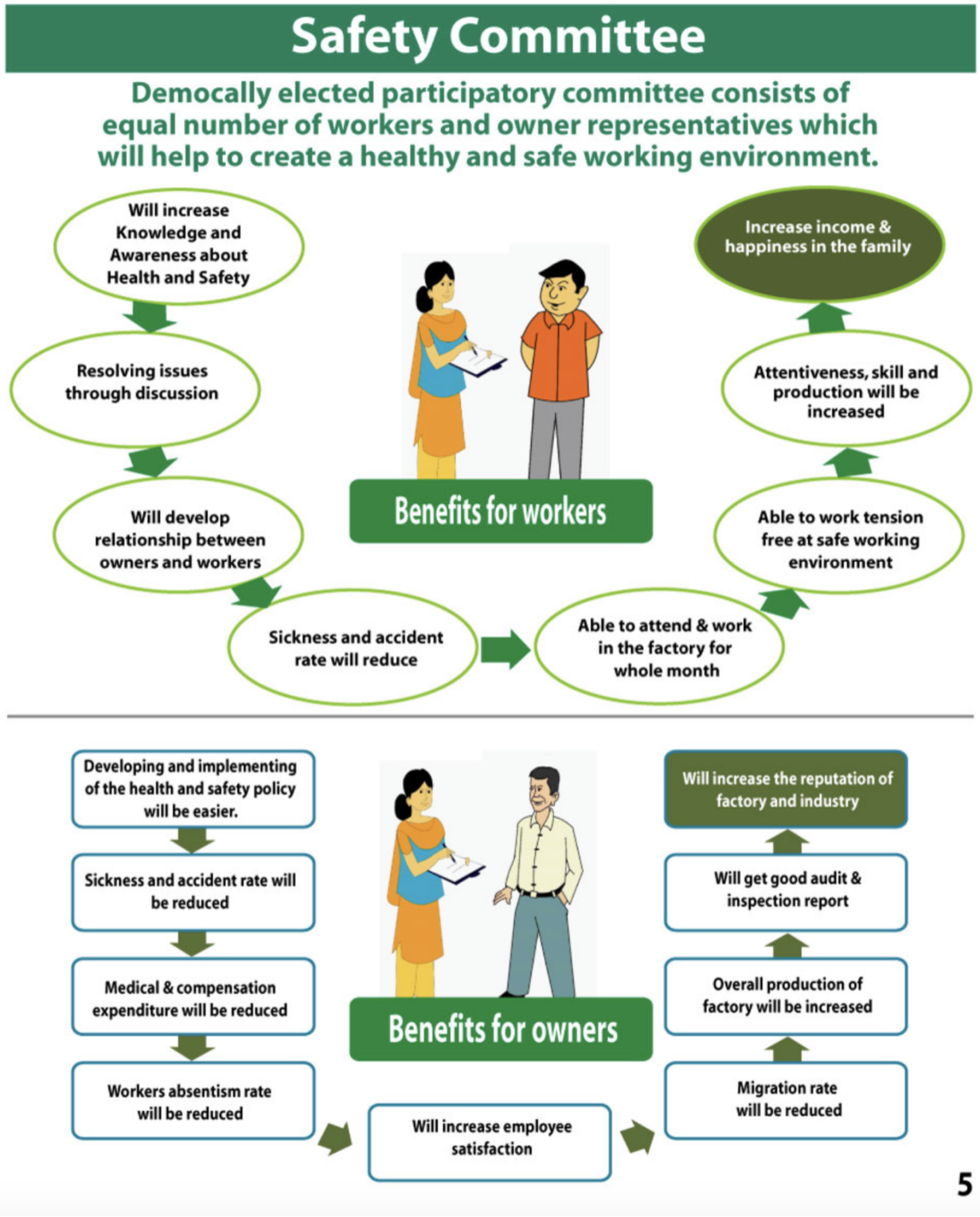


Figure II: Rana Plaza building collapse

Source: Wikipedia.

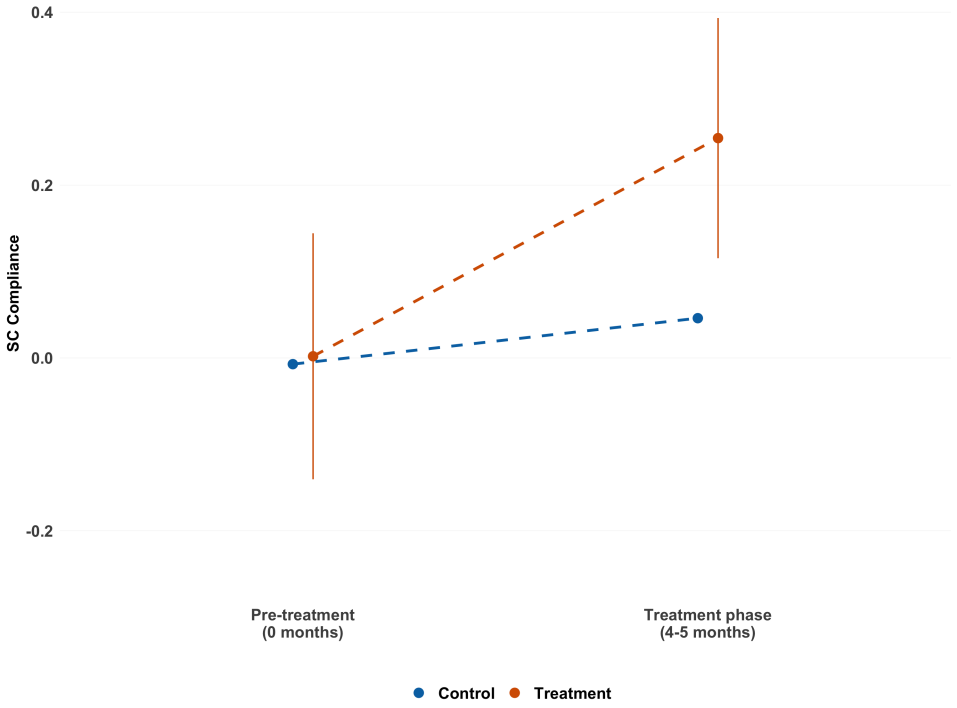
Figure III: Alliance member companies

Figure IV: Alliance SC Theory of Change



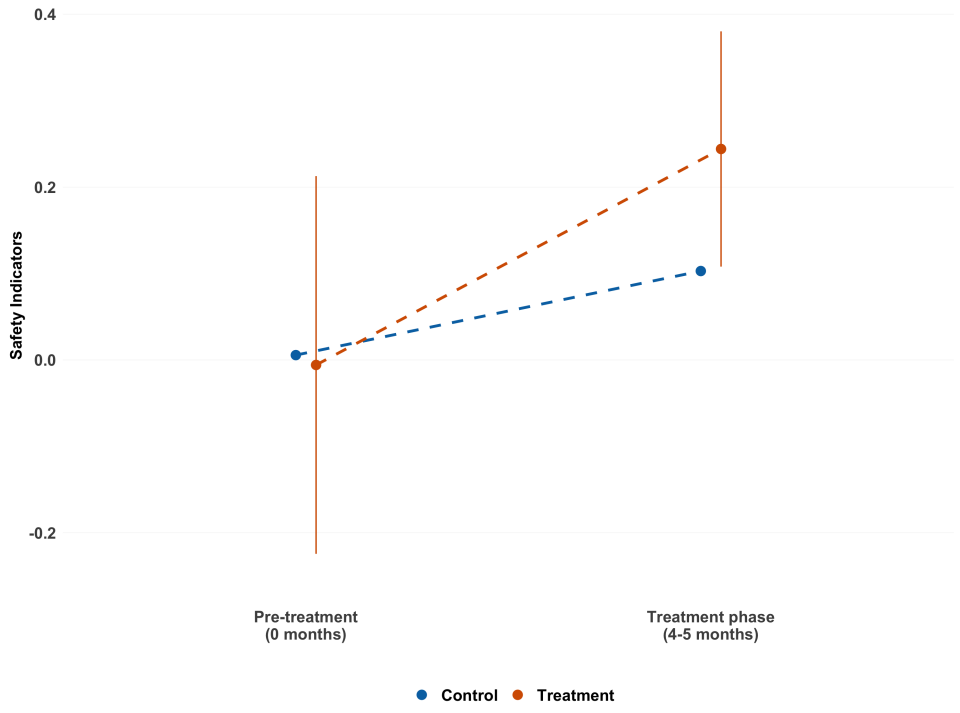
Source: Alliance training materials for SC members (English translation).

Figure V: SC Compliance Index



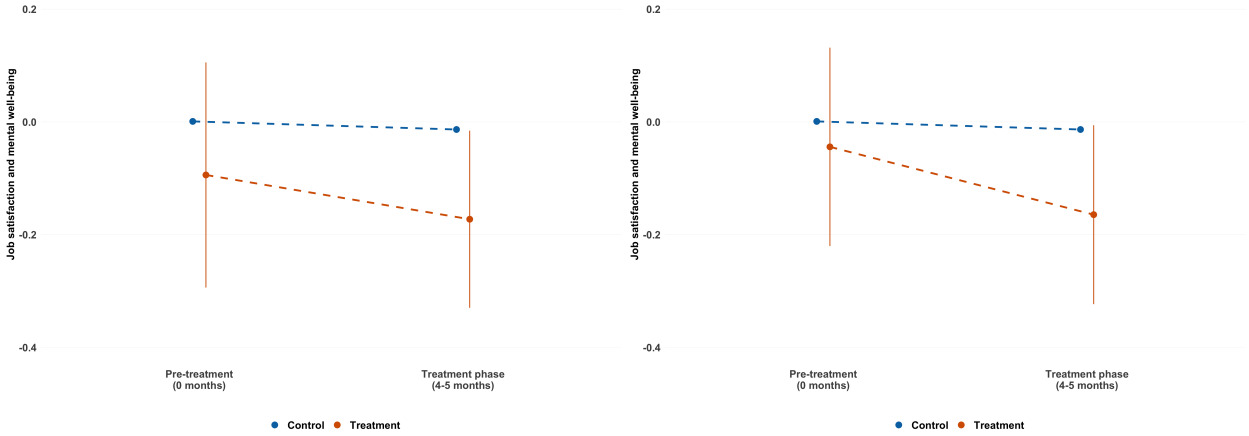
Notes: FDR p-val=0.020 for difference in post-treatment means. Whiskers show the the 95% confidence interval calculated from regressions of the outcome variable on a treatment indicator and stratification variables separately for pre-treatment and post-treatment rounds using robust standard errors. Summary index variable is constructed using methodology from Anderson (2008).

Figure VI: Safety Indicators Index



Notes: FDR p-val=0.089 for difference in post-treatment means. Whiskers show the the 95% confidence interval calculated from regressions of the outcome variable on a treatment indicator and stratification variables separately for pre-treatment and post-treatment rounds using robust standard errors. Summary index variable is constructed using methodology from Anderson (2008).

Figure VII: Worker job satisfaction and mental well-being Index



(a) Full sample

(b) Dropping 1 outlier

Notes: FDR p-val=0.089 for difference in post-treatment means. Whiskers show the the 95% confidence interval calculated from regressions of the outcome variable on a treatment indicator and stratification variables separately for pre-treatment and post-treatment rounds using robust standard errors. Summary index variable is constructed using methodology from Anderson (2008).

Table I: Key Safety Committee Requirements

Category	Requirements
Formation	6-12 committee members depending on factory size
	Equal worker-manager representation
	Appointment of worker representatives by collective bargaining agent or Participation Committee*
	SC president appointed by management, SC vice president appointed by worker representatives
	In establishments with >33% female workforce, at least 33% of worker representatives must be female
Operations	Establishments must maintain a written policy on the SC
	SCs must meet at least once per quarter
	SCs must maintain written meeting minutes
	Employers must provide SC members adequate time during working hours to fulfill their duties
	Employers must provide SC members with occupational health and safety training
Responsibilities	SCs must implement factory risk assessment at least once per quarter
	SCs must make safety-improvement recommendations to the employer
	SCs must arrange training and awareness-raising for workers
	SCs will participate in the oversight of the following safety management systems: Management of equipment and work procedure; Management of dangerous fumes, explosives, or flammable items; Fire safety management; Management of dangerous operations, occupational disease, poisonous disease; Emergency Planning
	SCs will investigate accidents and occupational disease and can submit recommendation to employer for treatment and compensation
	SCs will organize regular fire, earthquake, and other disaster management drills

*In factories with a collective bargaining agent (CBA), the CBA selects worker representatives to the safety committee. In factories where there is not a CBA, a Participation Committee (PC) selects worker representatives to the safety committee. A PC is legally required for all factories with 50 or more workers located outside of Export Processing Zones (EPZs). A PC has equal worker-manager representation that aims to promote trust and cooperation between employers and workers. It also aims to ensure application of labor laws.

Source: Translation based on Government of Bangladesh (2015).

Table II: Baseline balance tests

	Control mean	T-C diff	RI p-value	Number of factories
	(1)	(2)	(3)	(4)
Panel A: Primary outcome variables				
Compliance index	0.000	-0.072	0.453	84
Safety indicators index	0.002	-0.053	0.621	84
Job satisfaction & well-being index	-0.005	-0 .113	0.257	84
Number of employees	1166	-150	0.609	84
Gross wages (log)	15.820	-0.190	0.471	72
Labor productivity (log)	1.575	0.398	0.272	75
Labor productivity (log)†	1.575	-0.049	0.747	75
Labor productivity (log)†, trimmed sample	1.538	-0.043	0.782	75
Panel B: Factory characteristics				
Trade union at factory (1=Yes)	0.047	-0.045	0.514	84
EPZ(1=Yes)	0.163	0.036	0.775	84
Sewing (only)	0.465	-0.129	0.271	84
Number product types	1.163	0.053	0.688	84
Monthly absenteeism (%)	4.859	-0.667	0.439	80
Monthly turnover (%)	3.920	0.012	0.989	84
Prop. employees visit medical clinic (daily)	0.012	0.005	0.447	62
Participation in Alliance training (6 mo pre-baseline)	0.070	-0.021	1.000	84
Number Alliance remediation visit to factory (6 mo pre-baseline)	0.186	-0.014	1.000	84
Panel C: Worker survey respondent characteristics				
Age	27.179	0.185	0.821	84
Proportion female	0.568	-0.111	0.080*	84
Education (yrs)	6.222	-0.428	0.233	84
Tenure (yrs)	3.852	-0.180	0.707	84
Prior industry experience (yrs)	1.535	0.041	0.857	84

Notes: This table reports OLS estimates of baseline differences between control and treatment groups. For each outcome or covariate, I report the baseline control group mean in column (1). In column (2), I report the estimated coefficient for the treatment indicator from a regression of the outcome or covariate on the treatment indicator and stratification variables. In column (3), I report the randomization inference (RI) p-value for the coefficient reported in column (2) based on 5000 draws. The regression sample remains the same in all rows unless otherwise indicated. † The regression also includes product-type fixed effects. The trimmed sample drops factory-month observations in the 1st and 99th percentiles of labor productivity. *p<0.1; **p<0.05; ***p<0.01.

Table III: Treatment effects: Compliance with SC regulation

	Control mean	ITT Effect
	(1)	(2)
Panel A: Primary outcome		
SC Compliance Index	0.046	0.203 [0.001]*** {0.004}***
Panel B: Sub-indexes		
Formation sub-index	0.118	0.080 [0.298] {0.424}
Operations sub-index	0.184	0.055 [0.522] {0.534}
Responsibilities sub-index	-0.149	0.442 [0.000]*** {0.002}***
Observations		80
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on the index of compliance with the SC regulation and its sub-indexes. Outcome variables are listed on the left. In all cases, higher values of the index correspond to "positive" outcomes. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. In Panel A, *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. In Panel B, *p*-values adjusted to control the False Discovery Rate (FDR) across this primary outcome's sub-indexes are reported in curly brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table IV: Treatment effects: Safety Indicators Index

	Control mean	ITT Effect
	(1)	(2)
Panel A: Primary outcome		
Safety indicators	0.103	0.143 [0.046]** {0.113}
Panel B: Sub-indexes and sub-variables		
Factory safety spotcheck index†	-0.000	0.217 [0.015]** {0.083}*
CAP completion sub-variable	0.345	0.023 [0.794] {0.936}
Worker SC awareness sub-index	0.049	0.197 [0.189] {0.606}
Worker safety knowledge sub-index	0.551	-0.094 [0.503] {0.936}
Senior manager awareness sub-variable	0.086	0.075 [0.805] {0.936}
Observations		80
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on the safety indicators index and its sub-indexes. Outcome variables are listed on the left. In all cases, higher values of the index correspond to "positive" outcomes. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. In Panel A, *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. In Panel B, *p*-values adjusted to control the False Discovery Rate (FDR) across this primary outcome's sub-indexes are reported in curly brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. †This regression does not include a control for the baseline value of the dependent variable, as the factory safety spotcheck was not conducted at baseline. *p<0.1; **p<0.05; ***p<0.01.

Table V: Treatment effects: Physical indicators of factory safety

	Control mean (1)	ITT Effect (2)
Factory safety spotcheck index	0.000	0.217 [0.015]**
Sewing: Machines have guards *and* workers wear PPE† for their task	0.500	0.076 [0.619]
Cutting: Machines have knife guards *and* workers wear PPE for their task	0.792	0.071 [0.557]
Dyeing and jobs handling chemicals: Safety masks, goggles, gloves, aprons, and boots worn by workers handling chemicals	0.545	0.102 [0.668]
All PPE appropriate size, functional, and well-maintained	0.951	0.050 [0.492]
Aisles clearly marked and markings visible	0.780	0.025 [1.000]
Aisles clear of sewing scrapes and debris	0.951	0.048 [0.503]
Aisles clear of obstruction	0.854	0.014 [1.000]
Machines in good working order & dangerous parts properly covered	0.927	0.070 [0.247]
Work stations maintained in tidy condition (no loose materials close to electrical appliances)	0.976	0.022 [1.000]
One or more easily accessible first aid kit in section	0.976	0.022 [1.000]
Physical separation between storage & production areas	0.976	0.022 [1.000]
Drinking water easily accessible for all workers	1.000	-0.025 [1.000]
Drinking water provided appears clean (visual check)	1.000	-0.025 [1.000]
Stratification variables		Y

Notes: This table reports OLS estimates of treatment effects on the spotcheck sub-index and for each variable in the spotcheck index. Four variables on the spotcheck checklist drop from the analysis because all factories were found to comply with these variables (see Table AII). Sub-variables are listed on the left. Results are shown for the sub-variables *prior* to standardizing them for inclusion in the index. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator and stratification variables. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. †PPE stands for personal protective equipment. PPE vary by task and include equipment such as eye guards, finger guards, chain mesh gloves, goggles, boots, etc. *p<0.1; **p<0.05; ***p<0.01.

Table VI: Treatment effects: Workers' job satisfaction and mental well-being

	Control mean	ITT Effect
	(1)	(2)
Panel A: Primary outcome		
Worker job satisfaction & mental well-being (well-being index)	-0.013	-0.149 [0.061]* {0.113}
Panel B: Sub-indexes and sub-variables		
Job satisfaction sub-index	-0.130	-0.386 [0.017]** {0.075}*
Mental well-being sub-index	0.011	-0.059 [0.709] {0.792}
Turnover sub-variable	0.115	-0.010 [0.884] {0.792}
Absenteeism sub-variable	0.088	-0.084 [0.162] {0.321}
Observations		80
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on the workers' job satisfaction and mental well-being index and its sub-indexes. Outcome variables are listed on the left. In all cases, higher values of the index correspond to "positive" outcomes. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. In Panel A, *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. In Panel B, *p*-values adjusted to control the False Discovery Rate (FDR) across this primary outcome's sub-indexes are reported in curly brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table VII: Treatment effects: Business competitiveness outcomes

	(1)	(2)	(3)
Panel A	Log(Labor productivity)		
Treatment effect	0.115	0.087	0.036
	[0.148]	[0.189]	[0.392]
			{0.418}
Factories	75	75	74
Observations	375	368	370
Stratification variables	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y
Product type FE	Y	Y	Y
Trimmed sample	N	Y	N
Dropping outlier	N	N	Y
Panel B	Log(Gross wages)	Log(Employment)	
Treatment effect	-0.015	-0.011	
	[0.612]	[0.635]	
	{0.466}	{0.466}	
Factories	72	80	
Observations	360	400	
Stratification variables	Y	Y	
Control, baseline dep. var.	Y	Y	

Notes: This table reports OLS estimates of treatment effects on labor productivity, employment, and gross wages. Each column in the table reports the estimated coefficient from a separate regression. Panel A reports results for labor productivity. In column (1), the regression includes five post-treatment observations per factory. In column (2), the sample is trimmed at the 1st and 99th percentile of all factory-month labor productivity observations. In column (3), a factory in the control that partially shut down during the study is dropped. In Panel B, each regression includes five post-treatment observations per factory, where each observation is one month. The regression sample changes across columns due to differential data availability. The dependent variable in each column is regressed on the treatment indicator, stratification variables, and a control for the baseline value of the dependent variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table VIII: Heterogeneous treatment effects by managerial practices

	SC Compliance	Safety Indicators	Job Satisfaction & Mental Well-being
	(1)	(2)	(3)
Panel A: Treatment Phase			
Below median	0.105	0.050	-0.271
	[0.230]	[0.566]	[0.027]**
Above median	0.311	0.235	-0.030
	[0.000]***	[0.049]**	[0.754]
p-val, diff	[0.133]	[0.239]	[0.145]
Observations	80	80	80
Stratification variables	Y	Y	Y
Control, base. dep. var.	Y	Y	Y
Panel B: Post-treatment phase			
Below median	0.127	0.031	0.009
	[0.333]	[0.789]	[0.941]
Above median	0.292	0.090	0.240
	[0.016]**	[0.307]	[0.065]*
p-val, diff	[0.364]	[0.702]	[0.216]
Observations	80	80	80
Stratification variables	Y	Y	Y
Control, base. dep. var.	Y	Y	Y
Panel C: Pooled			
Below median	0.116	0.041	-0.131
	[0.195]	[0.613]	[0.216]
Above median	0.302	0.162	0.105
	[0.001]***	[0.084]*	[0.239]
p-val, diff	[0.173]	[0.332]	[0.093]*
Observations	160	160	160
Factories	80	80	80
Stratification variables	Y	Y	Y
Control, base. dep. var.	Y	Y	Y

Note: This table reports OLS estimates of heterogeneous treatment effects on primary outcome index variables. Each outcome variable is indicated at the top of the table. Each panel reports the results for data collection during the treatment phase, after the treatment phase, and pooling both rounds of data collection. In each panel, the "Below median" row reports the estimated treatment effect for the subgroup with below median baseline managerial practices. In each panel, the "Above median" row reports the estimated treatment effect for the subgroup with above median baseline managerial practices. The final row in each panel reports the *p*-value of the difference between the estimated treatment effects for below and above median subgroups. All regressions include stratification variables and a control for the baseline value of the dependent variable. All subgroups have 40 observations. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table IX: Persistence of treatment effects after end of intensive enforcement: Primary Outcome Index Variables

	Control mean	ITT Effect
	(1)	(2)
SC Compliance	0.149	0.213
		[0.012]**
		{0.078}*
Safety Indicators	0.156	0.063
		[0.371]
		{1.000}
Job satisfaction & mental well-being	-0.099	0.115
		[0.206]
		{1.000}
Observations		80
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of the persistence of treatment effects on primary outcome index variables measured 3-4 months after the end of the intensive enforcement period. Outcome variables are listed on the left. In all cases, higher values of the index correspond to "positive" outcomes. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator and stratification variables. Column (3) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Main Appendices

A: Index Variable Components

Table AI: SC Compliance Index

	Sub-Index	Variable	Variable Source 1	Variable Source 2
1	Formation	Equal worker-manager representation (or more workers than managers)	Factory Documentation	
2	Formation	Number of members is greater than or equal to mandated number of members	Factory Documentation	
3	Formation	Compliant worker representative selection process: CBA, PC, or WWA as required	SC President Survey	SC Worker Rep. Survey
4	Formation	Management does not select worker representatives on SC	SC President Survey	SC Worker Rep. Survey
5	Formation	In factories with >= (1/3) female workforce, at least (1/3) worker representatives are female	Factory Documentation	
6	Formation	Factory maintains list of SC Members	Factory Documentation	
7	Formation	Correlation between SC President's reports and factory documentation	Factory Documentation	SC President Survey
8	Operations	Factory maintains description of SC Members' roles and responsibilities	Factory Documentation	
9	Operations	Factory Safety Policy includes a section on the Safety Committee's role and responsibilities	Factory Documentation	
10	Operations	Safety Committee meets at least 1 time per 3 months	Factory Documentation	
11	Operations	Frequency of meetings per 3 months	Factory Documentation	
12	Operations	Meeting minutes are available for all Safety Committee meetings in past three months	Factory Documentation	
13	Operations	Meeting attendance lists are available for all Safety Committee meetings in past three months	Factory Documentation	
14	Operations	Safety Committee members have received training in their role on SC	SC President Survey	SC Worker Rep. Survey
15	Operations	Safety Committee members considered on duty during the time they spend on Committee meetings and Committee-related activities	SC President Survey	SC Worker Rep. Survey
16	Operations	Safety Committee uses compliant decision rule (unanimous or majority vote)	SC President Survey	SC Worker Rep. Survey
17	Operations	Count of number of questions out of prespecified questions where agreement between SC President's reports and factory documentation	Factory Documentation	SC President Survey
18	Operations	Count of number of questions out of prespecified questions where agreement between SC President's reports and SC worker member reports	SC President Survey	SC Worker Rep. Survey

19	Operations	Management interference in SC operations: Members of management provided any payments to worker representatives on the SC in return for not raising or pursuing safety issues; Members of management have interfered with or attempted to block SC efforts to improve factory safety	SC Worker Rep. Survey	
20	Responsibilities	Safety Committee has completed a risk assessment of the factory	Factory Documentation	
21	Responsibilities	Safety Committee has developed an action plan for safety improvements	Factory Documentation	
22	Responsibilities	Safety Committee makes regular safety reports/recommendations to management	SC President Survey	SC Worker Rep. Survey
23	Responsibilities	Frequency of follow-up: Regular reports and recommendations to management	SC President Survey	SC Worker Rep. Survey
24	Responsibilities	Senior management frequency of reports from SC (should be minimum 1x month or quarter)	Senior Manager Survey	
25	Responsibilities	Safety Committee organizes training and fire drills	SC President Survey	SC Worker Rep. Survey
26	Responsibilities	Number of fire drills, previous 3 months	Factory Documentation	
27	Responsibilities	Proportion of workers who report participation in safety-related training	Worker Survey	
28	Responsibilities	Proportion of workers who report participation in fire drill	Worker Survey	
29	Responsibilities	Safety Committee regularly inspects the factory's machinery and equipment and make suggestions to senior management in case of faulty operation or insufficient safety	SC President Survey	SC Worker Rep. Survey (midline, endline only)
30	Responsibilities	Safety Committee participation in the oversight and implementation of the factory's management of flammable and/or dangerous materials and goods	SC President Survey	SC Worker Rep. Survey (midline, endline only)
31	Responsibilities	Safety Committee participation in the oversight and implementation of the factory's fire prevention and preparedness activities	SC President Survey	SC Worker Rep. Survey (midline, endline only)
32	Responsibilities	Safety Committee participation in the oversight and implementation of the factory's health protection system	SC President Survey	SC Worker Rep. Survey (midline, endline only)
33	Responsibilities	Safety Committee investigates accidents and make recommendations to prevent future accidents	SC President Survey	SC Worker Rep. Survey
34	Responsibilities	In case of on-the-job worker injury or occupational disease, Safety Committee mediates between the worker and the factory.	SC President Survey	SC Worker Rep. Survey

Table AII: Safety Indicators Index

	Sub-Index	Variable	Variable Source
1	Floor Spotcheck	Aisles in section are clearly marked, and markings are easily visible	Floor Spotcheck (midline, endline only)
2	Floor Spotcheck	Aisles in section are clear of obstruction	Floor Spotcheck (midline, endline only)
3	Floor Spotcheck	Aisles in section are clear of sewing scraps or other materials	Floor Spotcheck (midline, endline only)
4	Floor Spotcheck	There is a physical separation between areas where materials are stored and areas where personnel are working (in this section)	Floor Spotcheck (midline, endline only)
5	Floor Spotcheck	Windows, fans, air conditioners or heaters are operational for air circulation, ventilation and provide an acceptable work floor temperature (in this section)	Floor Spotcheck (midline, endline only)
6	Floor Spotcheck	Machines are in good working order and points of operation and other potential dangerous parts are operated with proper machine guards and safety features (i.e., all reeling and dangerous parts of machines are covered) (machines in section)	Floor Spotcheck (midline, endline only)
7	Floor Spotcheck	Individual machines have an individual power shut-off switch within reach of the operator (machines in section)	Floor Spotcheck (midline, endline only)
8	Floor Spotcheck	Fire extinguisher and other fire-fighting materials are in clear view and easily accessible (in section)	Floor Spotcheck (midline, endline only)
9	Floor Spotcheck	Emergency exits are clearly marked with illuminated exit signs (in section)	Floor Spotcheck (midline, endline only)
10	Floor Spotcheck	Evacuation plan is easily visible in all production areas in section	Floor Spotcheck (midline, endline only)
11	Floor Spotcheck	At least one easily accessible first aid kit in section in section	Floor Spotcheck (midline, endline only)
12	Floor Spotcheck	Drinking water is easily accessible for all workers in section (within 100 meters for all workers in section)	Floor Spotcheck (midline, endline only)
13	Floor Spotcheck	Visual check of drinking water provided for workers appears clean (in section)	Floor Spotcheck (midline, endline only)
14	Floor Spotcheck	Sewing: Sewing machines are equipped with appropriate machine guards and workers wear appropriate PPE for their task (e.g., eye guards for button sewing, finger guards for pocket welt sewing) (in section)	Floor Spotcheck (midline, endline only)
15	Floor Spotcheck	Cutting: Cutting machines are equipped with knife guards and workes wear appropriate PPE for their task (e.g., chain mesh gloves for cutting tasks) (in section)	Floor Spotcheck (midline, endline only)

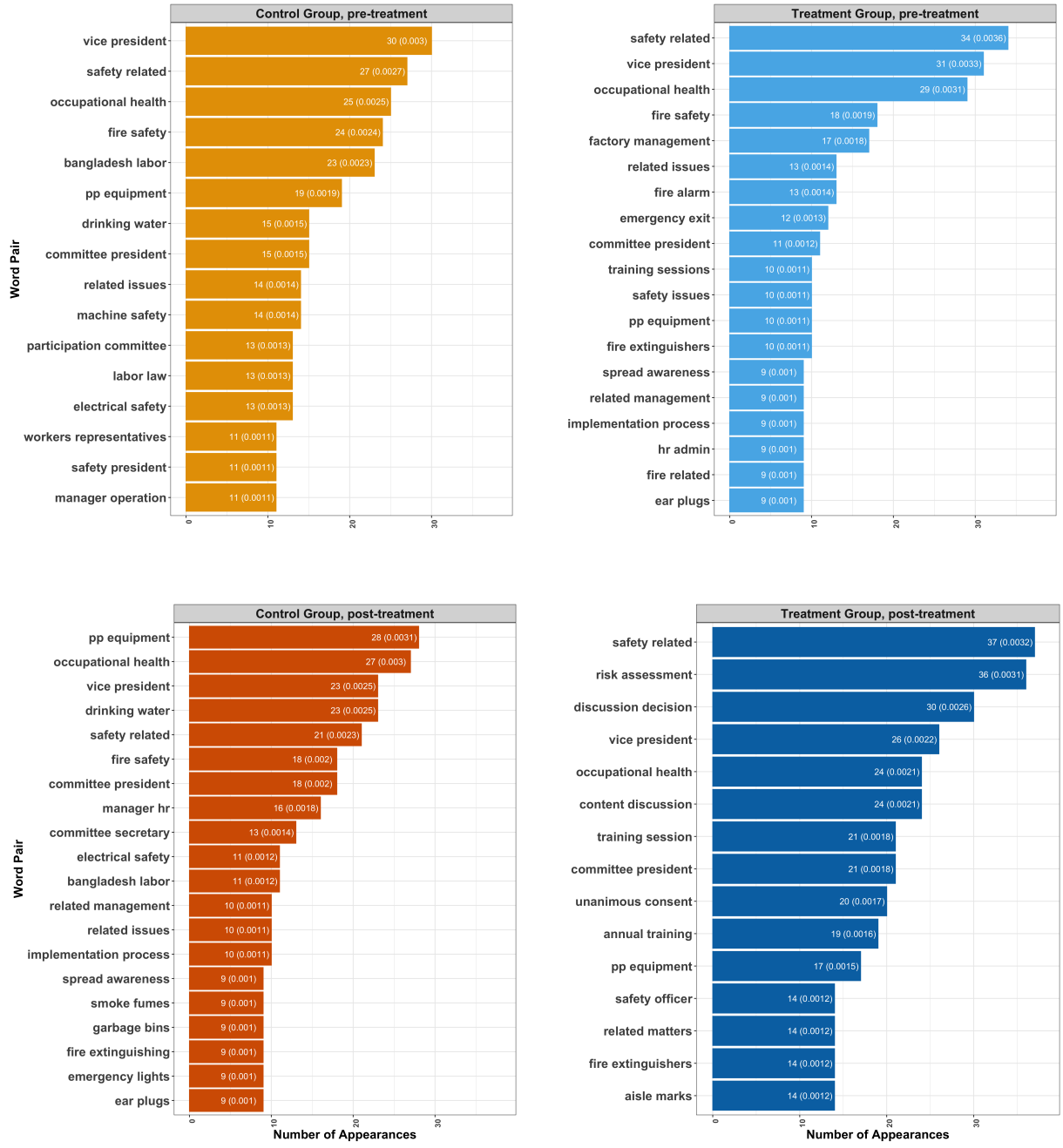
16	Floor Spotcheck	Dyeing and jobs handling chemicals: Safety masks, goggles, gloves, aprons, and boots are worn by workers handling chemicals (in section)	Floor Spotcheck (midline, endline only)
17	Floor Spotcheck	All PPE provided are of appropriate size, are functional, and appear well-maintained (in section)	Floor Spotcheck (midline, endline only)
18	Floor Spotcheck	All work stations are maintained in tidy condition, with no loose materials close to electrical appliances (in section)	Floor Spotcheck (midline, endline only)
19	CAP	Percent compliant	Alliance CAP data
20	Worker Awareness	Proportion of workers aware that factory has a SC	Worker Survey
21	Worker Awareness	Proportion of workers aware of SC's function and responsibilities	Worker Survey
22	Worker Awareness	Proportion of workers aware of how to contact SC member with issue	Worker Survey
23	Worker Awareness	Proportion of workers aware of SC topic-specific responsibilities	Worker Survey (midline, endline only)
24	Worker Safety Knowledge	Proportion of workers correctly answer fire question	Worker Survey
25	Worker Safety Knowledge	Proportion of workers correctly answer earthquake question	Worker Survey
26	Senior Manager Awareness	Senior management can provide at least one example of one issue identified by SC that has been resolved	Senior Manager Survey

Table AIII: Worker Job Satisfaction and Mental Well-being Index

	Sub-Index	Variable	Variable Source
1	Job Satisfaction	How satisfied are you with your job at your factory?	Worker Survey
2	Job Satisfaction	Have you suggested to or helped family or friends to get a job at your factory?	Worker Survey
3	Job Satisfaction	In the past three months or since you began working at this factory if less than three months ago, have you thought about leaving your job because of safety reasons?	Worker Survey
4	Mental Well-being	In general, how stressed are you about things in your life?	Worker Survey
5	Mental Well-being	How much control you feel that you have over the way your life turns out?	Worker Survey
6	Mental Well-being	How much control you feel that you have over your safety at the factory?	Worker Survey
7	Mental Well-being	How stressed are you about the risk of experiencing an accident or injury at your factory?	Worker Survey
8	Mental Well-being	How often do you feel unsafe when you are working at the factory?	Worker Survey
9	Absenteeism		Factory Questionnaire
10	Turnover		Factory Questionnaire

B: Figures and Tables

Figure BI: Most common two-word combinations in SC Meetings Minutes



Notes: To prepare the meeting minutes for text analysis, I strip the text of factory and participant names, the phrases "[health and] safety committee(s)" and "meeting(s)," English language stop words, numbers, and punctuation. I also replace the commonly used acronym of "ppe," which stands for personal protective equipment, and the complete phrase, with "pp equipment." Finally, I "stem" words, or replace them with their root, using the Porter stemmer. These approaches are common practice in text analysis (Gentzkow, Kelly and Taddy, 2019).

Figure BII: Correlation between WMS Management Index (excluding meeting question) and WMS Meeting-related Question, apparel firms in all countries

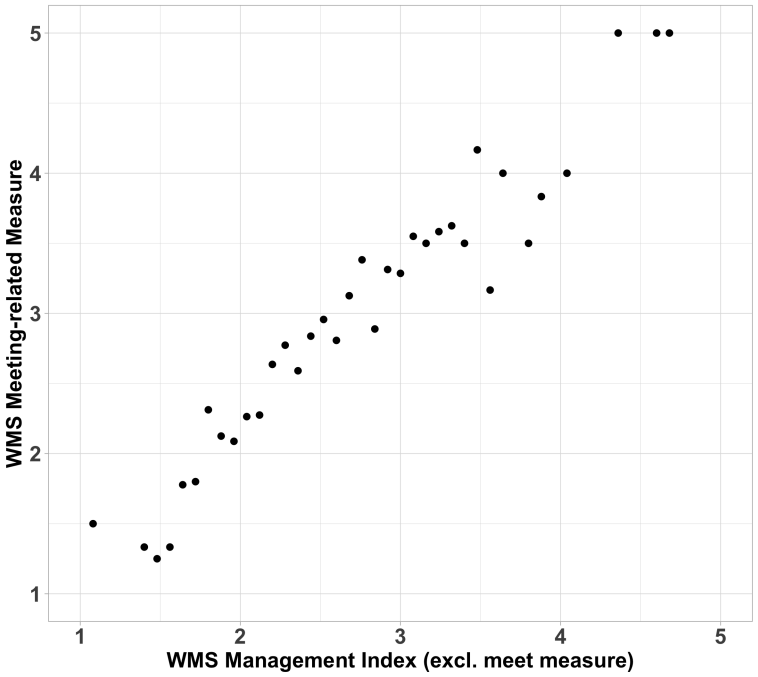


Table BI: Instrumental variable analysis: SC risk assessment and factory safety conditions

	(1)
	Performance on safety conditions spotcheck
SC Risk Assessment	0.520
	(0.226)**
Observations	80
Stratification variables	Y

Notes: 2SLS analysis in which SC risk assessment is instrumented by random assignment to the treatment group. *p<0.1; **p<0.05; ***p<0.01. This table reports a 2SLS estimate of the effects of SC risk assessment on spotcheck performance. SC risk assessment is instrumented by random assignment to the treatment group. Robust standard errors are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table BII: Treatment effects: Worker awareness outcome variables

	Control mean	ITT Effect
	(1)	(2)
Safety Indicators Index Sub-Variables		
Aware of SCs & their responsibilities	0.843	0.053
		[0.036]**
Knows factory has SC	0.945	0.040
		[0.017]**
Knows how to report safety concern to SC	0.920	0.011
		[0.651]
Reported num SC resp†	3.060	-0.118
		[0.345]
Other worker awareness variables		
Reports SC as channel for raising an issue	0.655	0.063
		[0.101]
Knows SC members†	0.689	0.073
		[0.028]**
Observations		80
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on all worker SC awareness variables from the baseline and 4-5 month surveys. The first four rows report outcomes included in the Safety Indicators Index (prior to standardization for inclusion in the index). Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. †Controls for the baseline value of these variables are not available. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. *p<0.1; **p<0.05; ***p<0.01.

Table BIII: Treatment effects: Visitors to Medical Clinic

	Mean proportion of workforce visits medical clinic (daily)	
	(1)	(2)
Panel A: Main treatment effects		
Treatment effect	-0.0017	-0.0019
	[0.107]	[0.135]
Factories	62	62
Observations	254	254
Control mean	0.011	0.012
Stratification variables	Y	Y
Control, baseline dep. var.	Y	Y
Weighted regression	N	Y
Panel B: Heterogeneous treatment effects by managerial practices		
Below Median	-0.0012	-0.0016
	[0.422]	[0.464]
Above Median	-0.0024	-0.0022
	[0.200]	[0.222]
p-val, diff	[0.626]	[0.834]
Factories	62	62
Observations	254	254
Control mean, below median	0.010	0.014
Control mean, above median	0.011	0.011
Stratification variables	Y	Y
Control, baseline dep. var.	Y	Y
Weighted regression	N	Y

Notes: This table reports OLS estimates of treatment effects on the proportion of employees that visit the medical clinic on a daily basis. Each column in the table reports the estimated coefficient from a separate regression. In column (1), probability weights based on the number of pre-treatment observations are applied. The dependent variable in each column is regressed on the treatment indicator, stratification variables, and a control for the baseline value of the dependent variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table BIV: Treatment effects: Workers' job satisfaction & mental well-being sub-variables

	Control mean	ITT Effect
	(1)	(2)
Panel A: Job Satisfaction		
Self-reported job satisfaction (qualitative scale, coded 1-5)	4.813	-0.045 [0.384] {0.345}
Respondent suggested/helped family or friends to get a job at their factory (previous 4 months)	0.573	-0.049 [0.266] {0.345}
Respondent has thought about leaving their job at factory for safety-related reasons (previous 3 months)	0.024	0.019* [0.064] {0.238}
Panel B: Mental Well-being		
Self-reported level of stress in life (qualitative scale, coded (-1)-(-5))	-1.760	-0.059 [0.474]
Self-reported perceived extent of control over their life (qualitative scale, coded 1-5)	4.083	-0.037 [0.521]
Self-reported perceived extent of control safety at factory (qualitative scale, coded 1-5)	4.368	-0.037 [0.520]
Self-reported stress about experiencing accident or injury at factory (qualitative scale, coded (-1)-(-5))	-1.489	0.041 [0.526]
Self-reported frequency of feeling unsafe at factory (qualitative scale, coded (-1)-(-5))	-1.236	-0.013 [0.691]
Panel C: Turnover and Absenteeism		
Turnover	3.356	-0.094 [0.779]
Absenteeism	4.457	0.040 [0.898]
Observations		80
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on each variable included in the worker job satisfaction and mental well-being index. Each panel reports the sub-variable results for a different sub-index. Sub-indexes and sub-variables are listed on the left. Results are shown for the variables *prior* to orienting them to be unidirectional and standardizing them for inclusion in the index. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table BV: Ex post minimum detectable effect sizes (MDEs): Business competitiveness outcomes

	Control mean (sd)	MDE
	(1)	(2)
Log(Labor productivity)†, trimmed sample	1.477	0.169
	(1.345)	
Log(Labor productivity), dropping labor productivity outlier	1.455	0.102
	(1.344)	
Log(Gross wages)	15.865	0.081
	(1.080)	
Log(Employment)	6.665	0.060
	(1.038)	

Notes: This table reports ex post power calculations and minimum detectable effect sizes for labor productivity, employment, and wage outcome variables with 80% power at the 5% significance level. Outcome variables are listed on the left. Column (1) reports the control group mean and standard deviation in column. Column (2) reports the ex post MDE. †Reported MDE is for sample trimmed at the 1st and 99th percentiles of all factory-month observations for labor productivity.

Table BVI: Heterogeneous treatment effects by managerial practices, business competitiveness outcomes

	Log(Labor productivity)		Log(Gross wages)	Log(Employment)
	(1)	(2)	(3)	(4)
Panel A: Treatment Phase				
Below median	0.161	0.049	-0.015	-0.010
	[0.127]	[0.403]	[0.673]	[0.754]
Above Median	0.017	0.030	-0.018	-0.007
	[0.830]	[0.662]	[0.731]	[0.858]
p-val, diff	[0.281]	[0.828]	[0.968]	[0.941]
Factories	75	74	72	80
Observations	368	370	360	400
Stratification variables	Y	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y	Y
Product type FE	Y	Y	N	N
Trimmed sample	Y	N	N	N
Dropping outlier	N	Y	N	N
Panel B: Post-treatment phase				
Below median	0.005	0.029	-0.004	-0.009
	[0.957]	[0.712]	[0.929]	[0.850]
Above Median	-0.027	-0.038	-0.015	0.023
	[0.636]	[0.504]	[0.808]	[0.671]
p-val, diff	[0.758]	[0.500]	[0.877]	[0.624]
Factories	75	74	72	80
Observations	218	222	216	240
Stratification variables	Y	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y	Y
Product type FE	Y	Y	N	N
Trimmed sample	Y	N	N	N
Dropping outlier	N	Y	N	N
Panel C: Pooled				
Below median	0.101	0.042	-0.011	-0.010
	[0.166]	[0.499]	[0.759]	[0.801]
Above Median	0.001	0.004	-0.017	0.004
	[0.987]	[0.939]	[0.727]	[0.922]
p-val, diff	[0.279]	[0.661]	[0.928]	[0.784]
Observations	586	592	576	640
Factories	75	74	72	80
Stratification variables	Y	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y	Y
Product type FE	Y	Y	N	N
Trimmed sample	Y	N	N	N
Dropping outlier	N	Y	N	N

Note: This table reports OLS estimates of heterogeneous treatment effects on business competitiveness variables. Each outcome variable is indicated at the top of the table. For labor productivity, in column (1), the sample is trimmed at the 1st and 99th percentile of all factory-month observations. In column (2), a factory in the control that partially shut down during the study is dropped. Each panel reports the results for data collection during the treatment phase, after the treatment phase, and pooling both rounds of data collection. In each panel, the "Below median" row reports the estimated treatment effect for the subgroup with below median baseline managerial practices. In each panel, the "Above median" row reports the estimated treatment effect for the subgroup with above median baseline managerial practices. The final row in each panel reports the *p*-value of the difference between the estimated treatment effects for below and above median subgroups. All regressions include stratification variables and a control for the baseline value of the dependent variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. *p<0.1; **p<0.05; ***p<0.01.

Table BVII: Heterogeneous treatment effects: Testing the importance of each dimension of heterogeneity

	Dependent variable:		
	Compliance	Safety Indicators	Job satisfaction & mental well-being
	(1)	(2)	(3)
Treat	0.194	0.002	-0.572
	[0.225]	[0.990]	[0.013]**
Treat*Abv med Mgmt	0.235	0.207	0.326
	[0.087]*	[0.155]	[0.068]*
Treat*Abv med Size	-0.174	0.139	0.127
	[0.187]	[0.346]	[0.438]
Treat*Abv med Compliance	-0.040	-0.075	0.407
	[0.766]	[0.621]	[0.040]**
Observations	80	80	80
Stratification variables	Y	Y	Y

Notes: This table reports OLS estimates of heterogeneous treatment effects, controlling for all dimensions of heterogeneity. Each column in table the reports the estimated coefficients from a separate regression. The regression sample is the same in all columns in a panel. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. *p<0.1; **p<0.05; ***p<0.01.

Online Appendices

C: Background on the Alliance

The Alliance's Membership Agreement

The Alliance's governing document is its Members' Agreement, which retail and apparel firms seeking to join the Alliance were required to sign. The Members' Agreement dictates the legal and financial commitments entailed in joining the Alliance. It also outlines the Alliance's envisioned safety initiatives, including worker empowerment programs, fire and building safety training for workers, supervisors, and managers, development of common building safety standards, and building safety audits and remediation.

The Alliance has been criticized on the grounds that its Member Agreement is unenforceable (The Economist, 2013). According to an academic analysis of the Alliance's Member Agreement, the only legally-enforceable component of the agreement pertains to Members' financial commitments (Donaghey and Reinecke, 2018). The Alliance leadership, though, maintained that the full Membership Agreement is legally binding. I have not evaluated the legal enforceability of the Members' Agreement, although I have aimed to test adherence to certain clauses.

Sourcing Commitment: The Member Agreement states that members agree to "work with factories that ensure a safe working environment, with each Member committing not to source from any factory that the Member has deemed to be unsafe." There is a question of whether members adhere to this commitment for factories that the Alliance suspends. I aim to verify this using data on Alliance members' activations and deactivations of supplier factories, as they report them to the Fair Factories Clearinghouse (FFC). The FFC is a software platform that aims to standardize monitoring of social and other compliance aspects in global supply chains. These data were only observable to Alliance administrators. Using these data, I check whether any Alliance members reported that a suspended factory was active as of January 2017 and October 2018, respectively. I find two cases in which a buyer reports a suspended factory has an active status. It's possible that these cases arise because suppliers were completing orders placed prior to the suspension decision or because buyers were not adhering to the requirement not to source from the factory. Given 29 members and 129 and 168 respective suspensions at these dates, this appears to be a low rate of noncompliance. I can't rule out, though, that Alliance members may have misreported their suppliers to the FFC.

Information Sharing: The Agreement outlines information sharing requirements in terms of sourcing and factory safety among members and the Alliance's leadership, the FFC, and the public. According to the agreement, the Alliance committed to publicly disclose a list of all factories utilized by its members by the fifteenth day of each month. In September 2014, I started to collect the publicly-posted lists. The Alliance posted these lists, with a couple of exceptions, each month through December 2018.

Financial Commitment: The financial commitment entailed a tiered fee structure based on the value of a member's exports from Bangladesh in the prior year. The maximum annual contribution was US1$million for members sourcing greater than US$250 million in exports in the prior calendar year.

Termination of Membership: The Agreement's term was for five years, but it required a minimum two-year commitment to participate. If a member were to exit the Alliance before two years, it would be responsible to pay its full five-year financial commitment and its exit would be publicly announced. If a member were to exit the Alliance after two years, it would be responsible to pay its financial commitment through the exit date and its exit would be publicly announced. These clauses would not apply if termination is due to the member's no longer sourcing from Bangladesh.

The Alliance's Safety Programs

The Alliance was announced on July 10, 2013, and its operations launched in late 2013. According to the Alliance's Member Agreement, its building safety audit and monitoring, safety training, and worker empowerment programs were to be launched immediately, or as soon as they could be developed. In order to implement these programs, the Alliance established an office in Dhaka, Bangladesh. According to its annual reports, it initially relied on a combination of its own staff and third-party service providers to implement its programs. Over the course of its five-year term, it moved more of these activities in-house. It employed a physical safety remediation team comprised of engineers and a training and worker empowerment team comprised of trainers.

The Alliance's central program was its building safety audit and remediation program. According to the 2015 Annual Report, the Alliance's initial focus was to develop a comprehensive list of supplier factories, to establish a strategy for building safety inspections, to employ one set of standards for building safety, and to collaborate with other stakeholders in Bangladesh (Alliance for Bangladesh Worker Safety, 2015). The Alliance reported that it completed initial building safety inspections of all factories in its supplier base by the end of 2014. Based on its actual building safety audit data, most factories were audited during 2014, although there are a small number of factories that are in the supplier base during 2014 that were audited after 2014. Once factories completed their initial audits, they developed a Corrective Action Plan (CAP) that the Alliance had to approve. Factories could then begin building safety remediation in order to complete their CAP. According to the Alliance's administrative records, the Alliance's engineers monitored factories on completion of their CAPs through verification visits.

The Alliance also implemented worker empowerment and safety training programs with its suppliers. Its first worker empowerment program is a worker safety helpline named Amader Kotha (AK). The AK Helpline was managed by a group of three external service providers. Workers could call the helpline to make reports about safety and non-safety related issues. The AK Helpline operated from 2014 through mid-2018 under the Alliance. In mid-2018, it was spun off as a separate organization managed by the three external partners, and as of 2020, it continues to operate in Bangladesh.

The Alliance's fire and building safety training program was launched in early 2014. The program entailed a train-the-trainer (TtT) approach in which factory representatives were trained by the Alliance's in-house training staff. Factories participating in the program submitted an action plan for when and how they would train all of the employees in the factory, and the Alliance monitored factories on fulfilment of these action plans through spotchecks. In addition to this training for general employees, the Alliance also launched a security guard training program in 2014. According to the Alliance's 2016 Annual Report, the security guard training program aimed "...to equip security guards with the skills to help prevent fires, to facilitate the swift and safe evacuation of workers, and to protect life rather than property...in case of an emergency."

These programs comprised the Alliance's operational activities until the SC Program's rollout. The SC Program differed from the Alliance's previous interventions because it focused on enforc-

ing a legal requirement to implement an internal structure to address safety issues. In contrast, the physical safety remediation, anonymous worker helpline, and fire and building safety training programs all entailed externally-provided incentives and expertise.

The timing of the SC Program's rollout was influenced by the GoB's promulgation of the SC implementation rules. In July 2013, the GoB amended Bangladesh's Labor Act to mandate SCs; it was not until September 15, 2015, however, that it promulgated the legal rules for SCs' formation and implementation. During 2015, the Alliance collaborated with the International Labor Organization (ILO)'s Better Work Program to pilot SCs in 16 factories. Once the GoB issued the legal rules for SCs, development of the SC Program began in earnest. In the spring and summer of 2016, the Alliance conducted a second pilot with 39 factories. It then launched the SC Program at the beginning of 2017 in collaboration with this research.

Nature of the Research Collaboration

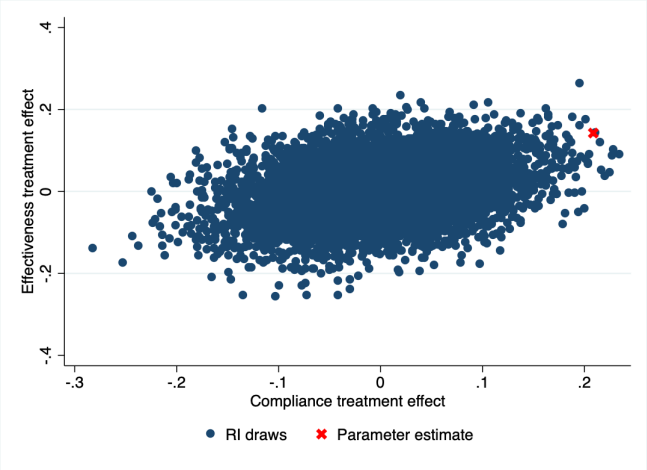
I initiated discussions with the Alliance in early 2015 that led to the research collaboration. At the time, the Alliance was beginning to develop its SC Program. The Alliance had received a lot of criticism on the grounds that it was not serious about improving safety or empowering workers and that the initiative was solely a marketing ploy. It's project managers, though, believed that their programs were having effects. They also indicated that the SC Program would be the most complex to implement and to measure its effects due to the multifaceted nature of SCs' responsibilities. I believe that these are the primary reasons why the Alliance was interested to engage in a rigorous research collaboration. Finally, it is worth noting that the SC Program was not the Alliance's largest or most publicized initiative. It's largest, most publicly-examined program was its building safety audit and remediation program. For this and other reasons, I think that it is unlikely that the research itself influenced the Alliance's allocation of effort toward the SC Program or that the Alliance opted into the research because it planned to allocate sufficient effort toward the SC Program to deliver marketable treatment effects.

D: Online Figures and Tables

Figure DI: Number of topics identified in SC Meeting Minutes

Notes: To prepare the meeting minutes for text analysis, I strip the text of factory and participant names, the phrases "[health and] safety committee(s)" and "meeting(s)," English language stop words, numbers, and punctuation. Finally, I "stem" words, or replace them with their root, using the Porter stemmer. These approaches are common practice in text analysis (Gentzkow, Kelly and Taddy, 2019). To identify the number of topics, I run three different probabilistic algorithms to estimate the number of topics 400 times (methods include Arun et al. (2010), Griffiths and Steyvers (2004), Cao et al. (2009)). After tuning, I allow each algorithm to identify between 2 and 50 topics per group of documents.

Figure DII: Joint distribution of compliance and safety indicators treatment effects under the null hypothesis with actual parameter estimates



Notes: The figure plots 5000 jointly generated estimates of treatment effects for SC compliance and safety indicators under the null hypothesis of no treatment effects. The actual parameter estimates are indicated in red.

Table DI: Sample Summary Statistics

	(1) Mean	(2) SD	(3) Minimum	(4) Maximum	(5) N
Panel A: Primary outcome variables					
Compliance index	-0.037	0.409	-2.478	0.585	84
Safety Indicators index	-0.020	0.494	-1.226	1.157	84
Job satisfaction & well-being index	-0.056	0.461	-2.329	0.772	84
Number of employees	1124	1321	52	7633	80
Gross wages (log)	15.72	1.09	13.35	18.26	72
Labor productivity (log)	1.761	1.501	0.079	6.483	75
Panel A: Factory characteristics					
Trade union at factory (1=Yes)	0.024	0.153	0	1	84
EPZ(1=Yes)	0.179	0.385	0	1	84
Sewing (only)	0.405	0.494	0	1	84
Number product types	1.321	0.779	0	4	84
Monthly absenteeism (%)	4.588	3.845	0.074	26.916	80
Monthly turnover (%)	3.911	4.789	0.000	29.950	84
Prop. employees visit medical clinic (daily)	0.013	0.019	0.001	0.109	62
Participation in Alliance training (6 mo pre-baseline)	0.060	0.238	0	1	84
Number Alliance remediation visit to factory (6 mo pre-baseline)	0.179	0.415	0	2	84
Panel B: Worker survey respondent characteristics					
Age	27.51	3.42	21.55	40.29	84
Proportion female	0.50	0.28	0	1	84
Education (yrs)	6.53	1.62	2.75	11.3	84
Tenure (yrs)	3.80	2.30	0.43	14.38	84
Prior industry experience (yrs)	1.54	0.97	0	5.68	84

Table DII: Secondary Outcome Index Variables

Index	Variable	Variable Source
Perceived SC Compliance & Effectiveness Index	Proportion of workers reporting compliant worker representative selection process: CBA, PC, or WWA as required.	Worker Survey
	Proportion of workers reporting management selects worker representatives on SC.	Worker Survey
	Proportion of workers reporting participation in activities organized by the Safety Committee.	Worker Survey
	Proportion of workers reporting SC provides reports of health and safety issues and/or recommendations on these issues to senior management.	Worker Survey
	Mean reported responsiveness of Safety Committee to workers' concerns.	Worker Survey
	Proportion of workers that think SC would be able to affect the factory's safety policies if learn their concerns.	Worker Survey
	Mean reported extent to which SC helps to improve safety at the factory.	Worker Survey
	Proportion of workers that report that factory management implemented one or more of the Safety Committee's recommendations.	Worker Survey
Perceived Human Resource Practices/Relations Index	If you or a worker like you told management an idea that could improve productivity at the factory, how likely do you think it is that management would implement the idea?	Worker Survey
	Can workers at your factory make anonymous reports/recommendations?	Worker Survey
	If you were to report unsafe conditions or make a safety suggestion, how likely do you think it is that management would address your concern?	Worker Survey
	To what extent does management care about workers' safety?	Worker Survey
	To what extent do workers and management discuss and make plans to improve safety together at your factory?	Worker Survey
	If you got hurt at work, would you be concerned that management may punish you because of reporting your injury to management?	Worker Survey
	If you got hurt at work, and you needed to take time off to get better, would you be able to take time off without a risk of losing your job?	Worker Survey
	If you got hurt at work, and you needed medical care because of the accident, do you think that management pay for part or all of the care?	Worker Survey
	Mean reported comfort sharing safety concern with senior managers.	Worker Survey
	Mean reported comfort sharing safety concern with direct supervisor.	Worker Survey
	I have seen supervisors verbally or physically abuse me or my coworkers.	Worker Survey

Worker Empowerment Index	How confident are you in your ability to identify unsafe conditions at your factory?	Worker Survey
	To what extent do you think that you or workers like you at your factory are capable of contributing ideas that can improve safety at the factory?	Worker Survey
	Have you reported a safety concern at your factory in the last year or since you began working at this factory if less than one year ago?	Worker Survey
	If you were to have a safety concern, would you report it?	Worker Survey
	If you were to get hurt at work, would you report the incident?	Worker Survey
	Mean reported comfort sharing safety concern with member of Safety Committee.	Worker Survey
	To what extent do you think that you or workers like you at your factory are capable of contributing ideas that can improve productivity at the factory?	Worker Survey
	Do you feel that if you wanted to change jobs, you could?	Worker Survey
	Do you have a goal for job promotion at your factory (e.g., operator level, process supervisor or production line manager)?	Worker Survey
Worker Organization Awareness Index	Familiar with the function and responsibilities: Trade union.	Worker Survey
	Familiar with the function and responsibilities: Participation Committee (non-EPZ only).	Worker Survey
	Familiar with the function and responsibilities: Workers' Welfare Association (EPZ only).	Worker Survey
	Awareness of factory's PC (non-EPZ only).	Worker Survey
	Awareness of factory's WWA (EPZ only).	Worker Survey
Supplier-buyer Relationships Index	Mean duration of relationship with Alliance-Member buyers (monthly).	Alliance administrative data
	Number of Alliance-member buyers (monthly).	Alliance administrative data
	Mean volume (in pieces) of orders per Alliance buyer.	Factory questionnaire

Table DIII: Baseline balance tests, secondary outcome variables

	(1) Control mean	(2) T-C diff	(3) RI p-value	(4) N
Panel A: Secondary outcomes for workers, full sample				
Perceived SC compliance & effectiveness index	0.000	-0.170	0.187	84
Perceived worker-manager relations index	0.000	-0.187	0.137	84
Worker empowerment index	0.000	-0.192	0.131	84
Worker organization awareness index	0.000	-0.165	0.315	84
Number non-pecuniary benefits	6.483	-0.404	0.035**	84
Monthly safety-related calls (per 1000 workers), Alliance Worker Helpline	0.057	0.025	0.875	80
Monthly non-safety-related calls (per 1000 workers), Alliance Worker Helpline	0.422	0.130	0.935	80
Panel B: Secondary outcomes for workers, dropping outlier on worker outcomes				
Perceived SC compliance & effectiveness index	0.000	-0.126	0.318	83
Perceived worker-manager relations index	0.000	-0.147	0.212	83
Worker empowerment index	0.000	-0.117	0.265	83
Worker organization awareness index	0.000	-0.121	0.444	83
Number non-pecuniary benefits	6.483	-0.388	0.053*	83
Monthly safety-related calls (per 1000 workers), Alliance Worker Helpline	0.057	0.025	0.873	79
Monthly non-safety-related calls (per 1000 workers), Alliance Worker Helpline	0.422	0.144	0.920	79
Panel C: Secondary outcomes for factories				
Average Weekly Working Hours	54.37	2.34	0.051*	79
Efficiency (sewing section)	0.530	0.070	0.215	33
Defects per hundred units	3.37	-1.01	0.127	72
Supplier-buyer relations index	0.018	-0.133	0.437	71

Notes: This table reports OLS estimates of baseline differences between control and treatment groups. For each outcome or covariate, I report the baseline control group mean in column (1). In column (2), I report the estimated coefficient for the treatment indicator from a regression of the outcome or covariate on the treatment indicator and stratification variables. In column (3), I report the randomization inference (RI) p-value for the coefficient reported in column (2) based on 5000 draws. The regression sample remains the same in all rows unless otherwise indicated. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table DIV: Treatment effects: Worker Secondary Outcome Variables

	Perceived SC compliance & effect index	Perceived worker-manager relations index	Worker empowerment index	Worker organization aware index	Number non-pecuniary benefits	Safety-related calls per 1k workers Alliance Helpline	Non-safety-related calls per 1k workers Alliance Helpline
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment effect	0.195	-0.039	0.067	0.016	-0.052	-0.002	-0.030
	[0.094]*	[0.630]	[0.487]	[0.842]	[0.761	[0.980]	[0.780]]
Control mean	-0.109	0.072	-0.178	0.073	6.802	0.107	0.303
Observations	80	80	80	80	80	400	400
Factories	80	80	80	80	80	80	80
Stratification variables	Y	Y	Y	Y	Y	Y	Y
Control, base. dep. var.	Y	Y	Y	Y	Y	Y	Y

Notes: This table reports OLS estimates of treatment effects on secondary outcome variables for workers. Each column in the table reports the estimated coefficient from a separate regression. The regression sample changes in Columns (6) and (7) due to a different source of data for these outcomes. The dependent variable in each column is regressed on the treatment indicator, stratification variables, and a control for the baseline value of the dependent variable. Each regression in Columns (1)-(5) includes one post-treatment observations per factory, while those in Columns (6)-(7) include five post-treatment observations per factory, where each observation is one month. For index variables, in all cases, higher values of the index correspond to "positive" outcomes. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table DV: Treatment effects: Output & Factory Secondary Outcome Variables

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Output)	Average Weekly Working Hours	Efficiency (Sewing section)	Defects per 100 units	Supplier-buyer relations index	
Treatment effect	0.026	-0.550	0.018	0.275	0.047	0.015
	[0.788]	[0.368]	[0.486]	[0.270]	[0.430]	[0.843]
Factories	75	79	33	72	71	80
Observations	374	395	165	360	355	400
Stratification variables	Y	Y	Y	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y	Y	Y	Y

Notes: This table reports OLS estimates of treatment effects on physical output, average weekly working hours, sewing section efficiency, defects per one hundred units, and the supplier-buyer relations index. For the supplier-buyer relations index, Column (4) includes all three variables from Table DII, and Column (5) drops the third variable, which is missing for nine factories. Each column in the table reports the estimated coefficient from a separate regression. The regression sample changes across columns due to differential data availability. Each regression includes five post-treatment observations per factory, where each observation is one month. The dependent variable in each column is regressed on the treatment indicator, stratification variables, and a control for the baseline value of the dependent variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table DVI: Pre-Analysis Plan Deviations

	Pre-Analysis Plan (PAP)	Modification
1)	The original PAP, which was posted on the AEA Registry in June 2017, included measures from survey question production line-level managers in the index of factory safety indicators (primary outcome variable). It also included survey measures from this group in certain secondary outcome index variables (perceived SC effectiveness and compliance index, worker-manager relations index, perception of worker capabilities index, and a perception of worker well-being index).	During baseline data collection, the research team determined that the line level managers were difficult to engage in surveys during the workday. As Safety Committees mostly aim to serve workers and to support senior management in occupational health and safety policies, after the baseline, I decided to remove variables from lower-level managers' survey data from these indices. I posted an updated PAP reflecting this change in December 2017.
2)	The SC Compliance Index included the following sub-variable: "President is management member and Vice President is worker member."	Due to an oversight, the information about the vice president's status was not collected for the first 57 baseline visits. I do not include this variable in the analysis.
3)	The Safety Indicators Index for the endline round included eight spotcheck variables that were only to be checked at endline.	Due to an administrative error, these additional items were not included in the checklist for the third visit for 14 out of 80 factories. As such, I do not include these variables in the analysis.
4)	The PAP indicated that I would also report outcomes measured at the individual-level at the individual level.	Consistent with best practice in econometric analysis for clusterd RCTs (Athey and Imbens, 2016), as the factory is the unit of analysis that is of interest, I omit individual-level regressions for space reasons.
5)	The PAP indicated that I would test for heterogeneous treatment effects by factories' location inside versus outside of an EPZ.	Ultimately, there were only 14 factories (7 treatment, 7 control) located in EPZs. There are large differences between these groups. As such, while I report baseline balance, I do not test for heterogeneous treatment effects by factory location.
6)	The PAP included hiring and machine downtime as factory-level secondary outcome variables.	I was not able to pilot the factory questionnaire until after I registered the PAP. I learned that many Alliance-covered factories did not systematically track these variables. Many factories in the sample had difficulty reporting them or indicated that they were unable to do. As such, I omit these secondary outcomes.
7)	The worker-manager relations index, a secondary outcome variable, include one variable to measure worker participation in strikes.	I decided that that participation in strikes was too sensitive to credibly measure in my setting, so I do not include the strike variable in the analysis.
8)	The PAP included an index of worker accidents and illness as a secondary outcome variable that included the medical clinic records, factory-reported illnesses and fires, and self-reported accidents and illness from workers.	The research team determined that factories' records of fires and accidents were often incomplete. For this reason, and due to the concern that the intervention may increase reporting of accidents by workers, I determined that the medical clinic records provide the most objective measure of accidents and illnesses.

Table DVII: Local Average Treatment Effects (LATEs): Primary Outcome Variables

	(1)	(2)	(3)
Panel A: Index Outcome Variables			
	SC Compliance	Safety Indicators	Job Satisfaction & Mental Well-being
LATE	0.220	0.154	-0.161
	(0.059)***	(0.069)**	(0.079)**
Control mean	0.046	0.103	-0.013
Observations	80	80	80
Stratification variables	Y	Y	Y
Control, base. dep. var.	Y	Y	Y
Panel B: Other Outcome Variables			
	Log(Labor Productivity)	Log(Gross Wages)	Log(Employment)
LATE	0.041	-0.017	-0.011
	(0.039)	(0.031)	(0.023)
Observations	370	360	400
Factories	74	72	80
Stratification variables	Y	Y	Y
Control, base. dep. var.	Y	Y	Y
Dropping outlier	Y	N	N

Notes: This table reports two stage least squares (2SLS) estimates of treatment effects on primary outcome variables. Each column in the table reports the estimated coefficient from a separate regression. In Panel A, all outcomes are index variables. In these cases, higher values of the index correspond to "positive" outcomes. Robust standard errors are reported in parentheses. Index variables constructed using Anderson (2008) variance-covariance weighted index. In Panel B, the regression sample changes across columns due to differential data availability. For labor productivity, results are shown after dropping the control factory that partially shuts down during the study. In this panel, each regression includes five post-treatment observations per factory, where each observation is one month. Robust standard errors clustered at the factory-level are reported in parentheses. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table DVIII: Lee (2009) bounds for primary outcome index variables

	Lower bound	Upper bound
SC compliance index	0.205	0.217
	(0.076)***	(0.069)***
Safety indicators index	0.140	0.141
	(0.091)	(0.072)**
Job satisfaction & mental well-being index	-0.159	-0.158
	(0.081)**	(0.084)*

Notes: This table reports Lee treatment effect bounds for sample selection. Outcome variables are listed on the left. Column (1) reports the lower bound. Column (2) reports the upper bound. Standard errors are reported in parentheses. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table DIX: Baseline balance tests, sub-index components

	(1) Control mean	(2) T-C diff	(3) RI *p*-value	(4) N
Panel A: SC Compliance				
Formation sub-index	0.000	-0.171	0.304	84
Operations sub-index	0.000	0.035	0.781	84
Responsibilities sub-index	0.000	-0.096	0.347	84
Panel B: Safety Indicators				
CAP completion sub-variable	0.017	0.092	0.655	84
Worker awareness sub-index	0.000	-0.547	0.027**	84
Worker knowledge sub-index	0.000	-0.082	0.661	84
Senior manager awareness sub-variable	0.000	0.348	0.136	84
Panel C: Worker Job Satisfaction and Mental Well-being				
Job satisfaction sub-index	0.000	-0.179	0.248	84
Mental well-being sub-index	0.000	-0.237	0.213	84
Turnover sub-variable	0.000	-0.002	0.989	84
Absenteeism sub-variable	-0.027	0.150	0.408	84

Notes: This table reports OLS estimates of baseline differences between control and treatment groups for the sub-indexes and sub-variables that comprise each primary outcome index. Each panel reports the sub-index/sub-variable balance tests for a different outcome variable. For each sub-index or sub-variable, column (1) reports the baseline control group mean. Column (2) reports the estimated coefficient for the treatment indicator from a regression of the sub-index or sub-variable on the treatment indicator and stratification variables. Column (3) reports the randomization inference (RI) *p*-value for the coefficient reported in column (2) based on 5000 draws. The regression sample remains the same in all rows. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table DX: Treatment effects: Meeting Minutes

	Log(Word Count)	
	(1)	(2)
Treatment effect	0.152	0.227
	[0.432]	[0.094]*
Observations	74	71
Stratification variables	Y	Y
Control, baseline dep. var.	N	Y

Notes: This table reports OLS estimates of treatment effects on the number of words in Meeting Minutes for Safety Committee meetings. Each column in the table reports the estimated coefficient from a separate regression. The regression sample changes in Column (2) because some factories' SCs did not meet before baseline. The dependent variable in each column is regressed on the treatment indicator, stratification variables, and in Column (2), a control for the baseline value of the dependent variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table DXI: Treatment effects: Workforce composition

	Dependent variable:				
	Age	Female	Tenure	Prior exp.	Yrs. Education
	(1)	(2)	(3)	(4)	(5)
Treatment effect	−0.200	−0.040	0.233	0.042	0.255
	[0.698]	[0.283]	[0.473]	[0.823]	[0.362]
Control mean	27.667	0.577	3.696	1.507	6.635
Observations	80	80	80	80	80
Stratification variables	Y	Y	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y	Y	Y

Notes: This table reports OLS estimates of treatment effects on workforce characteristics. Each column in the table reports the estimated coefficient from a separate regression. The regression sample is the same in all columns. The dependent variable in each column is regressed on the treatment indicator, stratification variables, and a control for the baseline value of the dependent variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. *p<0.1; **p<0.05; ***p<0.01.

Table DXII: Treatment effects: Business competitiveness outcomes (panel regression model)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A	Log(Labor productivity)					
Treatment x Post	0.097	0.108	0.076	0.087	0.036	0.043
	[0.233]	[0.171]	[0.273]	[0.200]	[0.389]	[0.292]
Factories	75	75	75	75	74	74
Observations	750	750	738	738	740	740
Factory FE	Y	Y	Y	Y	Y	Y
Calendar month FE	N	Y	N	Y	N	Y
Trimmed sample	N	N	Y	Y	N	N
Dropping outlier	N	N	N	N	Y	Y
Panel B	Log(Employment)		Log(Gross wages)			
Treatment x Post	-0.013	-0.012	-0.017	-0.018		
	[0.563]	[0.620]	[0.575]	[0.546]		
Factories	80	80	72	72		
Observations	800	800	719	719		
Factory FE	Y	Y	Y	Y		
Calendar month FE	N	Y	N	Y		

Notes: This table reports panel regression estimates of treatment effects on labor productivity, employment, and wages. Each column in the table reports the estimated coefficient from a separate regression. The regression sample changes across regressions due to differential data availability. In Panel A columns (1) and (2), each regression includes five 5 pre-baseline and 5 post-baseline observations per factory. In Panel A columns (3) and (4), the sample is trimmed at the 1st and 99th percentile of all factory-month labor productivity observations. In columns (5) and (6), a factory in the control that partially shut down during the study is dropped. In Panel B, in each regression, there are 10 observations per factory, 5 pre-baseline and 5 post-baseline. In all regressions, the dependent variable in each column is regressed on an interaction between the treatment indicator and a post-treatment indicator variable and factory fixed effects. Calendar month fixed effects are included in the second column for each variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table DXIII: Baseline balance tests within subgroups for management practices heterogeneity analysis, primary outcome variables

	(1) Control mean	(2) T-C diff	(3) RI p-value	(4) N
Below median subgroup:				
SC Compliance	0.025	0.113	0.233	40
Safety Indicators	0.099	-0.067	0.651	40
Job Satisfaction & Mental Well-being	0.070	-0.198	0.284	40
Log(Labor productivity)†, trimmed sample	2.158	-0.236	0.315	40
Log(Wages)	15.625	0.007	0.984	38
Log(Employment)	6.297	0.060	0.858	40
Above median subgroup:				
SC Compliance	-0.032	-0.114	0.255	40
Safety Indicators	-0.067	0.013	0.934	40
Job Satisfaction & Mental Well-being	-0.053	-0.021	0.838	40
Log(Labor productivity)†, trimmed sample	1.088	-0.087	0.731	35
Log(Wages)	16.004	-0.376	0.372	34
Log(Employment)	6.925	-0.514	0.166	40

Notes: This table reports OLS estimates of baseline differences between control and treatment groups within above- and below-median management subgroups for treatment effect heterogeneity analysis. For each outcome, within subgroup, column (1) reports the baseline control group mean. Column (2) reports the estimated coefficient for the treatment indicator from a regression of the outcome on the treatment indicator and stratification variables within that subgroup. Column (3) reports the randomization inference (RI) p-value for the coefficient reported in column (2) based on 5000 draws. Column (4) reports the number of observations in that subgroup. † The regression also includes product-type fixed effects. The trimmed sample drops factory-month observations in the 1st and 99th percentiles of labor productivity. *p<0.1; **p<0.05; ***p<0.01.

Table DXIV: Baseline balance tests within non-management subgroups for heterogeneity analysis, primary outcome variables

	(1) Control mean	(2) T-C diff	(3) RI *p*-value	(4) N
Panel A: Factory Size				
Below median subgroup:				
SC Compliance	-0.003	0.078	0.442	40
Safety Indicators	0.021	0.083	0.627	40
Job Satisfaction & Mental Well-being	0.004	0.026	0.868	40
Above median subgroup:				
SC Compliance	-0.010	-0.075	0.536	40
Safety Indicators	-0.007	-0.118	0.464	40
Job Satisfaction & Mental Well-being	-0.002	-0.133	0.452	40
Panel B: SC Compliance				
Below median subgroup:				
SC Compliance	-0.224	-0.061	0.369	40
Safety Indicators	-0.029	-0.059	0.688	40
Job Satisfaction & Mental Well-being	-0.063	0.030	0.808	40
Above median subgroup:				
SC Compliance	0.221	0.027	0.543	40
Safety Indicators	0.042	0.162	0.314	40
Job Satisfaction & Mental Well-being	0.068	-0.161	0.360	40
Panel C: Location in EPZ				
EPZ subgroup:				
SC Compliance	-0.190	0.385	0.188	14
Safety Indicators	-0.051	0.220	0.560	14
Job Satisfaction & Mental Well-being	-0.104	0.503	0.078*	14
Non-EPZ subgroup:				
SC Compliance	0.031	-0.013	0.866	66
Safety Indicators	0.017	-0.066	0.584	66
Job Satisfaction & Mental Well-being	0.022	-0.173	0.137	66
Panel D: HR Managerial Practices				
Below median subgroup:				
SC Compliance	-0.051	0.080	0.476	40
Safety Indicators	-0.181	0.045	0.798	40
Job Satisfaction & Mental Well-being	-0.011	-0.250	0.151	40
Above median subgroup:				
SC Compliance	0.031	-0.095	0.347	40
Safety Indicators	0.167	-0.055	0.680	40
Job Satisfaction & Mental Well-being	0.011	0.040	0.738	40

Notes: This table reports OLS estimates of baseline differences between control and treatment groups within non-management subgroups for treatment effect heterogeneity analysis. For compliance, size, and HR managerial practices, I partition the sample into above/below median subgroups using the baseline value of the variable. For location in Export Processing Zone (EPZ), I partition the sample using this variable. Each panel reports the within subgroup baseline differences for a different dimension of heterogeneity. For each outcome, within subgroup, column (1) reports the baseline control group mean. Column (2) reports the estimated coefficient for the treatment indicator from a regression of the outcome on the treatment indicator and stratification variables within that subgroup. Column (3) reports the randomization inference (RI) *p*-value for the coefficient reported in column (2) based on 5000 draws. Column (4) reports the number of observations in that subgroup. *p<0.1; **p<0.05; ***p<0.01.

Table DXV: Other heterogeneous treatment effects: Primary Outcome Index Variables

	SC Compliance	Safety Indicators	Job Satisfaction & Mental Well-being
	(1)	(2)	(3)
Panel A: Baseline Size			
Below median	0.261	0.072	-0.210
	[0.001]***	[0.397]	[0.108]
Above median	0.138	0.202	-0.083
	[0.142]	[0.101]	[0.377]
p-val, diff	[0.367]	[0.401]	[0.442]
Panel B: Baseline SC Compliance			
Below median	0.236	0.199	-0.299
	[0.026]**	[0.119]	[0.022]**
Above median	0.169	0.074	0.004
	[0.013]**	[0.294]	[0.964]
p-val, diff	[0.603]	[0.418]	[0.067]*
Panel C: Baseline HR Management Practices			
Below median	0.131	0.093	-0.177
	[0.110]	[0.374]	[0.153]
Above median	0.281	0.189	-0.107
	[0.001]***	[0.078]*	[0.293]
p-val, diff	[0.250]	[0.536]	[0.668]
Observations	80	80	80
Stratification variables	Y	Y	Y
Control, base. dep. var.	Y	Y	Y

Note: This table reports OLS estimates of heterogeneous treatment effects on primary outcome index variables. Each outcome variable is indicated at the top of the table. Each panel reports the results for a different dimension of heterogeneity. In each panel, the "Below median" row reports the estimated treatment effect for the subgroup with below median baseline values of the heterogeneity variable. In each panel, the "Above median" row reports the estimated treatment effect for the subgroup with above median baseline values of the heterogeneity variable. The final row in each panel reports the *p*-value of the difference between the estimated treatment effects for below and above median subgroups. All regressions include stratification variables and a control for the baseline value of the dependent variable. All subgroups have 40 observations. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table DXVI: Other heterogeneous treatment effects: Primary Outcome Index Variables, Pooling treatment and post-treatment rounds of data

	SC Compliance	Safety Indicators	Job Satisfaction & Mental Well-being
	(1)	(2)	(3)
Panel A: Baseline Size			
Below median	0.234	0.083	-0.114
	[0.004]***	[0.279]	[0.283]
Above median	0.163	0.109	0.084
	[0.060]*	[0.236]	[0.314]
p-val, diff	[0.575]	[0.831]	[0.131]
Panel B: Baseline SC Compliance			
Below median	0.236	0.098	-0.045
	[0.029]**	[0.304]	[0.678]
Above median	0.180	0.099	0.007
	[0.010]**	[0.202]	[0.929]
p-val, diff	[0.676]	[0.990]	[0.705]
Panel C: Baseline HR Management Practices			
Below median	0.129	0.024	-0.096
	[0.121]	[0.754]	[0.388]
Above median	0.288	0.181	0.064
	[0.002]***	[0.056]*	[0.388]
p-val, diff	[0.220]	[0.208]	[0.221]
Observations	160	160	160
Factories	80	80	80
Stratification variables	Y	Y	Y
Control, base. dep. var.	Y	Y	Y

Note: This table reports OLS estimates of heterogeneous treatment effects on primary outcome index variables, pooling treatment and post-treatment rounds of data. Each outcome variable is indicated at the top of the table. Each panel reports the results for a different dimension of heterogeneity. In each panel, the "Below median" row reports the estimated treatment effect for the subgroup with below median baseline values of the heterogeneity variable. In each panel, the "Above median" row reports the estimated treatment effect for the subgroup with above median baseline values of the heterogeneity variable. The final row in each panel reports the *p*-value of the difference between the estimated treatment effects for below and above median subgroups. All regressions include stratification variables and a control for the baseline value of the dependent variable. All subgroups have 40 factories. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table DXVII: Treatment effects after end of intensive enforcement: Business competitiveness outcomes

	(1)	(2)	(3)
Panel A	Log(Labor productivity)		
Treatment effect	-0.022	-0.017	-0.010
	[0.652]	[0.733]	[0.837]
			{1.000}
Factories	75	75	74
Observations	225	218	222
Stratification variables	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y
Product type FE	Y	Y	Y
Trimmed sample	N	Y	N
Dropping outlier	N	N	Y
Panel B	Log(Gross wages)	Log(Employment)	
Treatment effect	-0.009	0.003	
	[0.811]	[0.932]	
	{1.000}	{1.000}	
Factories	72	80	
Observations	216	240	
Stratification variables	Y	Y	
Control, baseline dep. var.	Y	Y	

Notes: This table reports OLS estimates of the persistence of treatment effects on labor productivity, employment, and gross wages measured 3-4 months after the end of the intensive enforcement period. Each column in the table reports the estimated coefficient from a separate regression. Panel A reports results for labor productivity. In column (1), the regression includes three post-treatment, post-enforcement intervention observations per factory. In column (2), the sample is trimmed at the 1st and 99th percentile of all factory-month labor productivity observations. In column (3), a factory in the control that partially shut down during the study is dropped. In Panel B, each regression includes three post-treatment, post-enforcement intervention observations per factory, where each observation is one month. The regression sample changes across columns due to differential data availability. The dependent variable in each column is regressed on the treatment indicator, stratification variables, and a control for the baseline value of the dependent variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. *p<0.1; **p<0.05; ***p<0.01.

E: Main Results Dropping Outlier on Worker Variables

Table EI: Baseline balance tests, dropping outlier on worker outcomes

	(1) Control mean	(2) T-C diff	(3) RI p-value	(4) N
Primary outcome variables				
SC Compliance index	0.000	-0.086	0.379	83
Safety Indicators	0.002	-0.032	0.765	83
Job satisfaction & well-being index	-0.005	-0.061	0.479	83
Number employees	1166	-242	0.432	84
Gross wages (log)	15.820	-0.263	0.303	71
Labor productivity (log)	1.575	0.456	0.202	74
Labor productivity (log)†	2.588	-0.036	0.817	74
Total factor productivity (log)†, trimmed sample	2.849	0-0.028	0.855	74
Panel B: Compliance sub-components				
Formation sub-index	0.000	-0.189	0.254	83
Operations sub-index	0.000	0.028	0.840	83
Responsibilities sub-index	0.000	-0.113	0.286	83
Panel C: SC Effectiveness sub-components				
CAP completion sub-variable	0.017	0.095	0.659	83
Worker awareness sub-index	0.000	-0.424	0.051*	83
Worker knowledge sub-index	0.000	-0.089	0.646	83
Senior manager awareness sub-variable	0.000	0.321	0.191	83
Panel D: Worker job satisfaction and mental well-being sub-components				
Job satisfaction sub-index	0.000	-0.133	0.396	83
Mental well-being sub-index	0.000	-0.118	0.424	83
Turnover sub-variable	0.000	-0.003	0.989	83
Absenteeism sub-variable	-0.023	0.143	0.448	83

Notes: This table reports OLS estimates of baseline differences between control and treatment groups. For each outcome or covariate, I report the baseline control group mean in column (1). In column (2), I report the estimated coefficient for the treatment indicator from a regression of the outcome or covariate on the treatment indicator and stratification variables. In column (3), I report the randomization inference (RI) p-value for the coefficient reported in column (2) based on 5000 draws. The regression sample remains the same in all rows unless otherwise indicated. *p<0.1; **p<0.05; ***p<0.01.

Table EII: Treatment effects: Compliance with SC regulation, dropping outlier on worker outcomes

	Control mean	ITT Effect
	(1)	(2)
Panel A: Primary outcome		
SC Compliance Index	0.046	0.204
		[0.001]***
		{0.004}***
Panel B: Sub-indexes		
Formation sub-index	0.118	0.079
		[0.332]
		{0.498}
Operations sub-index	0.184	0.056
		[0.534]
		{0.553}
Responsibilities sub-index	-0.149	0.437
		[0.001]***
		{0.002}***
Observations		79
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on the index of compliance with the SC regulation and its sub-indexes. Outcome variables are listed on the left. In all cases, higher values of the index correspond to "positive" outcomes. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator and stratification variables. Column (3) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. In Panel A, *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. In Panel B, *p*-values adjusted to control the False Discovery Rate (FDR) across this primary outcome's sub-indexes are reported in curly brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table EIII: Treatment effects: Safety Indicators Index, dropping outlier on worker outcomes

	Control mean	ITT Effect
	(1)	(2)
Panel A: Primary outcome		
Safety Indicators	0.103	0.139
		[0.052]*
		{0.142}
Panel B: Sub-indexes and sub-variables		
Factory safety spotcheck index	-0.000	0.218
		[0.014]**
		{0.077}*
CAP completion sub-variable	0.345	0.025
		[0.780]
		{1.000}
Worker SC awareness sub-index	0.049	0.202
		[0.159]
		{0.466}
Worker safety knowledge sub-index	0.551	-0.077
		[0.575]
		{1.000}
Senior manager awareness sub-variable	0.086	0.053
		[0.838]
		{1.000}
Observations		79
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on the index of Safety Indicators Index and its sub-indexes. Outcome variables are listed on the left. In all cases, higher values of the index correspond to "positive" outcomes. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator and stratification variables. Column (3) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. In Panel A, *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. In Panel B, *p*-values adjusted to control the False Discovery Rate (FDR) across this primary outcome's sub-indexes are reported in curly brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table EIV: Treatment effects: Workers' job satisfaction and mental well-being, dropping outlier on worker outcomes

	Control mean	ITT Effect
	(1)	(2)
Panel A: Primary outcome		
Worker job satisfaction & mental well-being (well-being index)	-0.013	-0.147 [0.075]* {0.142}
Panel B: Sub-indexes and sub-variables		
Job satisfaction sub-index	-0.130	-0.387 [0.021]** {0.092}*
Mental well-being sub-index	0.011	-0.053 [0.746] {0.781}
Turnover sub-variable	0.115	-0.011 [0.877] {0.781}
Absenteeism sub-variable	0.088	-0.084 [0.173] {0.352}
Observations		79
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on the workers' job satisfaction and mental well-being index and its sub-indexes. Outcome variables are listed on the left. In all cases, higher values of the index correspond to "positive" outcomes. Column (1) reports the control group mean of the outcome variable. Column (2) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator and stratification variables. Column (3) reports the estimated ITT effect from a regression of the outcome variable on the treatment indicator, stratification variables, and a control for the baseline value of the outcome variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. In Panel A, *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. In Panel B, *p*-values adjusted to control the False Discovery Rate (FDR) across this primary outcome's sub-indexes are reported in curly brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table EV: Treatment effects: Business competitiveness outcomes, dropping outlier on worker outcomes

	(1)	(2)	(3)
Panel A	Log(Labor productivity)		
Treatment effect	0.118	0.091	0.041
	[0.140]	[0.169]	[0.343]
			{0.347}
Factories	74	74	73
Observations	370	363	365
Stratification variables	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y
Product type FE	Y	Y	Y
Trimmed sample	N	Y	N
Dropping outlier	N	N	Y
Panel B	Log(Gross wages)	Log(Employment)	
Treatment effect	-0.015	-0.009	
	[0.647]	[0.718]	
	{0.560}	{0.560}	
Factories	71	79	
Observations	355	395	
Stratification variables	Y	Y	
Control, baseline dep. var.	Y	Y	

Notes: This table reports OLS estimates of treatment effects on labor productivity, employment, and gross wages. Each column in the table reports the estimated coefficient from a separate regression. Panel A reports results for labor productivity. In column (1), the regression includes five post-treatment observations per factory. In column (2), the sample is trimmed at the 1st and 99th percentile of all factory-month labor productivity observations. In column (3), a factory in the control that partially shut down during the study is dropped. In Panel B, each regression includes five post-treatment observations per factory, where each observation is one month. The regression sample changes across columns due to differential data availability. The dependent variable in each column is regressed on the treatment indicator, stratification variables, and a control for the baseline value of the dependent variable. Randomization inference (RI) *p*-values based on 5000 draws are reported in square brackets. *p*-values adjusted to control the False Discovery Rate (FDR) across primary outcomes are reported in curly brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.