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

Laura Boudreau*

January 22, 2020

Updated version here

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 small, positive effect on indicators of safety committees' effectiveness, 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.

Keywords: Labor Regulation, Enforcement, Corporate Social Responsibility, Private regulation, Supply chain, Working Conditions, Worker voice, Safety, Garment Industry, Bangladesh

*Columbia University (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 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

In developing countries, governments often lack the capacity or the political will to update and to enforce regulation (Dal Bó and Finan, 2016), including labor regulation. For example, in a 2018 global ranking of labor law and enforcement, 65% of developing countries were found to systematically violate or to provide no guarantee of rights to workers (International Trade Union Confederation, 2018).[1] In response, many Western multinationals sourcing from developing countries privately enforce local labor laws on their suppliers through "Corporate Social Responsibility" (CSR) programs (O'Rourke, 2014).

In a globalized production system, such CSR programs could in principle provide an important source of enforcement in countries with weak state capacity. It is an open question, however, whether multinationals have the incentives and the capabilities to improve labor standards in developing countries. On one hand, multinational buyers have incentives to prevent egregious labor violations and industrial disasters in supplier establishments that could pose reputational risks (Tadelis, 2002, McDevitt, 2011, Bachmann et al., 2019). On the other, enforcement of labor laws may increase labor costs, which suggests that without effective monitoring, multinationals' promises to improve standards may not be credible (Besley and Ghatak, 2007). Finally, even if multinationals are motivated to improve labor standards, it's unclear whether they have sufficient bargaining power and monitoring capacity to influence suppliers' practices (Short, Toffel and Hugill, 2016).

Further, if private enforcement of labor law can improve targeted establishments' compliance, it raises the question of whether it generates net benefits or costs to these establishments and to their workers. Existing empirical evidence provides grounds for concern. Botero et al. (2004) examine labor regulation in 85 countries; they show that heavier *de jure* labor regulation is associated with lower labor force participation and higher un-

[1] 28% of high-income countries included in the report were found to systematically violate or to provide no guarantee of rights to workers (International Trade Union Confederation, 2018). Most of these are located in the Middle East.

employment. Besley and Burgess (2004) compare *de jure* labor regulation across Indian states and find that pro-worker regulation is associated with eroded firm competitiveness and increased urban poverty. More recent studies emphasize that a material share of these apparent costs may be due to factors such as corruption and low quality of the written regulation (Amirapu and Gechter, forthcoming, Malesky and Taussig, 2017). Understanding the relationships among labor regulation, firm competitiveness, and worker well-being is of fundamental policy importance in developing countries, but the lack of available causal evidence limits our ability to make informed recommendations.

This paper provides the first experimental evidence on the effects of private enforcement of labor law in a developing country where government enforcement is lacking. I partnered with a set of multinational retail and apparel firms, known as the Alliance for Bangladesh Worker Safety (hereafter, the Alliance), that aimed to improve the safety performance of its shared Bangladeshi supplier base.[2] The 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).[3] In conjunction with the Alliance, I implemented a randomized controlled trial (RCT) in which I randomly assigned supplier factories to the Alliance's enforcement of a local labor law that requires factories to have worker-manager safety committees (SCs). I estimate the intervention's effects on suppliers' compliance with the labor law and their SCs' effectiveness at improving safety. I also assess the intervention's effects on factories' labor productivity, wages, and employment and on their workers' well-being.

The RCT was implemented over 2017-2018 as part of the Alliance's roll-out of its SC

[2] 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.

[3] 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.

Program. It involved 84 garments and garments-related factories in Bangladesh. The Alliance's SC Program is the "treatment" in this RCT. It is a 6-month enforcement intervention that aims to bring factories into meaningful compliance with Bangladesh's SC law.[4] I randomly assigned 41 factories to immediate participation in the Program (treatment group) and 43 factories to deferred participation in the Program approximately 11 months later (control group).[5] A five-member research team made three full-day visits to factories. The team collected a pre-intervention baseline, a post-intervention round about 5 months later, and a second post-intervention round about 9 months after baseline (see Figure 1). For treatment factories, the 5-month data collection visit occurred toward the end of the 6-month enforcement program. I also implemented a retrospective questionnaire to collect production, HR, and other business-related data. Finally, the Alliance provided its own monitoring and administrative datasets. 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 multinationals' enforcement program significantly increases factories' compliance with Bangladesh's SC labor law, which I measure using a pre-specified index of compliance outcomes. The intervention improves factories' compliance-related outcomes by 0.20 standard deviations (sds) on average. Most factories begin with SCs that are formed correctly but largely inactive. The intervention significantly increases their level of activity; for example, they begin to meet more frequently and are nearly four times more likely to conduct risk assessment. This increase in compliance translates into a statistically significant improvement in a pre-specified index of factory safety indicators. Treatment factories outperform control factories on this index by 0.14 sds on average. This

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

[5]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 naturally aligns with the Alliance's standard operating procedures.

improvement is driven by a statistically significant improvement in treatment factories' performance on an independent spotcheck of safety conditions by the research team. The intervention also improves some measures of workers' awareness of the SC.

These findings demonstrate that private enforcement of labor law can significantly improve compliance and contribute to achieving the law's objectives. They provide reason for greater optimism regarding the ability of private enforcement to improve labor standards in developing countries compared to an existing literature in political science on private regulation in global supply chains. Historically, this literature has been largely skeptical (Locke, Qin and Brause, 2007, Locke, Amengual and Mangla, 2009, Locke and Romis, 2010, Locke, 2013, Distelhorst et al., 2015, Toffel, Short and Ouellet, 2015), but as more causal evidence emerges, it is beginning to update its view.[6] It also provides the first experimental evidence that firms' CSR initiatives can successfully generate public goods/curtail public bads. This finding joins Dragusanu and Nunn (2018) in beginning to build a body of empirical evidence on the efficacy of CSR. It also provides empirical justification for theoretical models of CSR that identify CSR with the private creation of public goods or curtailment of public bads, such as those of Besley and Ghatak (2007) and Lai et al. (2017).

Contrary to my hypothesis, the enforcement intervention significantly decreases workers' job satisfaction. Treatment factories' performance on an index of job satisfaction and mental well-being is -0.15 sds lower than controls'. The index includes self-reported measures of job satisfaction and mental well-being as well as revealed preference measures including absenteeism and turnover. The decline is driven by a reduction in self-reported measures related to job satisfaction. The third round of data collection, though, shows that these negative effects on workers are temporary in nature.

Evidence on supplier competitiveness, including labor productivity, wages, and em-

[6] A working paper by Amengual and Distelhorst (2019), for example, uses a regression discontinuity design to study the multinational retailer Gap Inc's supplier code of conduct for labor. They find that a failing audit grade only improves suppliers' future compliance if coupled with the threat of a reduction in Gap's orders.

ployment, do not provide evidence of adverse effects. Estimated treatment effects on labor productivity are actually positive but are small. Estimated treatment effects on wages are negative but are also close to zero and not statistically significant. Finally, estimated treatment effects on employment are close to zero. 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.

To further unpack the results, I analyze pre-specified dimensions of heterogeneity. In particular, I consider the role of suppliers' organizational capacity. The existing research on improving adherence to regulation in developing countries focuses on the effects of strengthening state-supplied regulation and 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 organizational capacity of the private sector also constrains regulatory efficacy in developing countries. An emerging literature on managerial practices and environmental and social performance provides a basis to explore this possibility. On the environmental side, Bloom et al. (2010) and Gosnell, List and Metcalfe (2019) show that better managerial practices reduce firms' environmental impacts. Adhvaryu, Kala and Nyshadham (2019) show that managers who implement active personnel management practices mitigate negative productivity shocks caused by air pollution. On the labor side, Distelhorst, Hainmueller and Locke (2017) show that adoption of lean managerial practices is associated with improved compliance. Tanaka (2019) provides evidence of trade-induced social upgrading in which firms in Myanmar improve working conditions in response to access to foreign markets. In this study, I contribute evidence of whether there is complementarity between enforcement capacity for labor regulation and managerial practices in the private sector.

I find that factories' baseline managerial practices are an important factor in determining the enforcement intervention's effects.[7] The treatment has large, positive effects on compliance and on measures of SC effectiveness in factories with better baseline managerial practices. In contrast, factories with worse practices do not significantly improve their compliance or measures of SC effectiveness. These results suggest that there may be complementarity between labor regulation and managerial practices. Increasing compliance with labor regulation may depend not only on providing firms with appropriate incentives, but also on their capacity to respond to these incentives.

Further, the negative effects on workers is driven by poorly-managed factories. I continue to investigate possible mechanisms for this result. For now, I provide suggestive evidence in favor 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 effect would be consistent with recent experimental evidence on low-wage workers' response to an upgrade in employer-provided housing from Adhvaryu, Nyshadham and Xu (2018). While objective measures of housing quality improve, workers' expectations for the improvement are not met, and workers' job satisfaction declines 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 global trade. Several studies have found that heavier *de jure* labor regulation is associated with worse economic performance and adverse consequences for workers (Fishback 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 quality in many developing countries (Duflo et al., 2013, Dal Bó and Finan, 2016, Amirapu and Gechter, forthcoming). Fisman and Wang (2015), for example, show that workplace deaths at politically-connected Chinese

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

firms are two to three times higher than at unconnected firms and that this relationship is best explained by firms using political connections to bypass safety regulations. Recent experimental research in Ethiopia finds that forms of industrial work common in developing countries have large, negative health impacts on workers with no compensating increases in income (Blattman and Dercon, 2018). Scholars have raised supply chain linkages as a possible mechanism to bring about improved regulation and enforcement. For example, Harrison and Scorse (2010) show that anti-sweatshop campaigns led the Indonesian government to raise minimum wages, which resulted in a large real wage increases with some costs for firms but no significant effects on employment. This is the first study, however, to test the potential for private enforcement of regulation in a context where state enforcement is lacking. I further contribute by identifying the causal effects of enforcement on labor productivity and on workers' well-being. Finally, my results suggest an under-explored constraint on regulatory efficacy: Organizational capacity of the private sector.

Second, it contributes to a burgeoning literature on the economics of CSR. Economists have long espoused the Friedman (1970) view that markets should produce private goods and governments should provide public goods and correct failures. Recent theoretical and empirical work, however, highlights two primary reasons why this dichotomy may blur. First, there is significant evidence that 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 provides the theoretical and empirical bases for its

existence and desirability. A recent exception is Dragusanu and Nunn (2018), who consider its efficacy; they show that Fair Trade certification is associated with higher incomes and improved educational outcomes for Costa Rican coffee farmers' families. I study a group of multinationals' CSR program that aims to improve suppliers' compliance with local labor law by leveraging possible monopsonistic power. I provide the first experimental evidence on an important, largely open question: Does private enforcement of *de jure* regulation achieve its stated objective, which, in this setting, is to improve safety?[8] If so, what are the economic consequences?

Third, this paper contributes to the literature on collective worker voice and intra-firm institutions. Economists have long been interested in how increasing workers' collective voice in firms' decision-making impacts firms' economic performance and workers' welfare. The empirical literature on this topic, however, has generally suffered from selection bias, and available causal evidence is limited to marginal firms (Addison, Schnabel and Wagner, 2001, DiNardo and Lee, 2004, Lee and Mas, 2012, Yao and Zhong, 2013). My experimental setting improves on previous research by enabling identification of causal effects that are not local to marginal firms. In so far as Western multinationals are driving the effort to increase worker voice, this research also shares commonalities with the literature on Western attempts to introduce Western governance institutions in developing countries (e.g., Casey, Glennerster and Miguel (2012) and Humphreys, Sánchez de la Sierra and Van der Windt (2017)).

Finally, this paper contributes to the literature on occupational safety and health (OSH), and in particular, on OSH committees. While most of this literature is limited in its ability to make causal statements, Levine, Toffel and Johnson (2012) is a recent exception.

[8] An interesting literature spanning political science and management science asks related questions. It largely concludes that multinationals' private regulation programs are generally ineffective at improving compliance and that local context is the main predictor (Locke, Qin and Brause, 2007, Locke, Amengual and Mangla, 2009, Locke and Romis, 2010, Locke, 2013, Distelhorst et al., 2015, Toffel, Short and Ouellet, 2015). As acknowledged in this literature, though, it suffers from a lack of causal identification and a lack of data access. It relies on cross-country and cross-supplier comparisons that are subject to various forms of omitted variables bias.

The authors use a natural experiment in OSH inspections by California state regulators to show that inspections significantly reduce injury rates and costs without negatively affecting establishments' competitiveness. The literature on OSH committees generally 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 thorough review). My contribution is to randomize enforcement of OSH committees to provide causal evidence of their effects on factory safety. I also identify complementarity between HR management practices and OSH effectiveness. Further, my main measure of SCs' effectiveness does not rely on injury rates, the reporting of which may be impacted by the treatment. Instead, I use indicators of factory safety, such as performance on spotchecks of factory safety conditions, to measure SCs' effectiveness. I will complement this analysis with an analysis of injury data, which I collect from medical clinic records, workers, and factory management.

The remainder of this paper is organized as follows: Section 2 describes the context, including the Alliance and the SC Program. Section 3 presents the research design. Section 4 presents the preliminary results. Section 5 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). Multinational buyers rely on Bangladesh for its combination of low prices and large production capacity (McKinsey & Company, 2011).[9]

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),

[9] 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."

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.[10] 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, for its lack of enforcement of regulation, and for its low minimum wages for many years.[11] 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 major risks to sourcing from the country (McKinsey & Company, 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 2), 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 some 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 multinational buyers 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 safety and health in Bangladesh's garments sector. The coalition was known as the Accord on Fire and Building Safety in Bangladesh (hereafter, the Accord). Several U.S. retailers re-

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

[11] 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.

fused to sign the Accord due to the participation of labor unions and 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 improve workplace safety and to strengthen some freedom of association rights. One of the amendment's key provisions was the requirement for SCs.[12] The GoB also agreed to a a multi-stakeholder action plan that included strengthening its labor inspection capacity, building and fire safety audits and remediation of the full garments-related sector, safety training, and numerous other actions (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 the remaining 30-40% of the sector (International Labor Organization, 2017).

The GoB published implementation rules for the SC mandate on September 15, 2015. These rules articulate the specific requirements for SCs' formation, operations, and responsibilities. Table 1 summarizes the key aspects. Factories had six months from this date to form and to operationalize their SCs. Despite the *de jure* requirement that establishments form and implement SCs, *de facto*, enforcement of the regulation was low. According to an International Labor Organization (2017) report, from 2015-2017, the GoB's focus was primarily on physical safety remediation.[13] Unsurprisingly, as captured by the headline of a news article from late 2017, compliance with the regulation was low. The article, titled, *"Half of all apparel units flout needs for safety committees"* (Munni, 2017) summarizes an internal government report on the sector's low level of compliance with the SC regulation.

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

[13] 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).

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 3. 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.[14] 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 operates in Bangladesh as of late 2019.

The Alliance required all supplier factories to participate in its building safety audit, building remediation, and worker training and empowerment programs. Failure to comply with one or more of these programs resulted in suspension from all Alliance Members' supplier bases; over the course of its five-and-a-half year term, the Alliance suspended 179 factories.[15]

The Alliance's intervention to enforce the SC law was its SC Program.[16] The SC Program included four phases:

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 Program's central feature was the SC's preparation and fulfillment of its action

[14]More details on the Alliance and its members are available on its website: www.bangladeshworkersafety.org.

[15]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.

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

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 signed off on the plan. Before approving the plan, the Alliance reviewed it and worked with the establishment to make revisions as necessary. The establishment then implemented the plan. To provide evidence of its activities, the establishment used Alliance-provided templates to submit evidence of each activity by e-mail within 2-3 days of the activity. 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 establishment's progress. If found to be insufficient, the factory could be required to repeat parts of the Program, or the factory's 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.[17]

How do SCs contribute to improved safety and health? Figure 4 displays the causal chain from the Alliance's SC training materials. The causal chain highlights the mechanisms through which the SC may directly and indirectly affect safety- and labor-related outcomes. On the worker side, the SC aims to increase workers' knowledge and awareness, resolve issues through discussion, and improve relations between workers and managers. As a result, the SC reduces accidents and occupational diseases, which in turn reduces absenteeism, improves morale, and increases workers' productivity. On the owner side, the SC aims to strengthen the establishment's OSH policies and implementation, which reduces accidents and occupational diseases, which in turn improves job satisfaction, reduces turnover, and improves productivity and social compliance. Evidently, while in Section 1, I discuss the possibility of increased compliance negatively affecting labor productivity, the Alliance highlights the potential beneficial effects.

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

The Alliance implemented the SC Program with factories that did not supply to Accord buyers; this is because the Accord also implemented a SC Program with its suppliers, including those that were covered by the Alliance.

Does the SC Program actually increase monitoring?

In light of the question of whether multinationals' promises to enforce labor laws are credible, one may question whether the Alliance's SC Program results in increased monitoring. I can partially answer this question using the Alliance's own administrative records. The Alliance's most informative but costly form of monitoring was onsite audits. According to administrative records, during these audits, an Alliance staff member reviewed SC-related documentation, interviewed SC members and workers, and visually inspected the factory. The Alliance conducted audits during factories' participation in the SC Program and afterward as part of its regular monitoring program. Among the treatment factories, 10% were audited during the SC Program and 15% were audited in the six months after completing the Program.

3 Research design

3.1 Randomized assignment to the SC Program

This study's goals are first to identify the effects of multinationals' enforcement of local labor regulation on their suppliers, and second to identify how increasing enforcement affects suppliers' competitiveness and workers' well-being. I address these questions through random assignment of 41 Alliance-covered factories to participate in the SC Program, which entails the Alliance's enforcement of Bangladesh's SC regulation, and 43 Alliance-covered factories serve as controls.[18]

The 84 factory sample is drawn from the population of SC Program-eligible supplier

[18] All control factories were required to participate in the SC Program after completing the study period.

factories. In order to be eligible, factories must have a separate committee that is formed in compliance with Bangladeshi labor law.[19] In most factories, this committee is the Participation Committee (PC), and it is responsible for appointing worker representatives to the SC.[20] Worker representatives on the PC must be elected through free, fair, and competitive elections. The Alliance verified a factory's election process as part of determining a factory's 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 for the SC Program.

The RCT was built into the Alliance's roll out of the SC Program.[21] From January through December 2017, every time 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 strata is a batch of factories. In 11 cases in which multiple factories shared ownership and location (shared building or compound), I randomly selected one factory to participate in the RCT.[22] All other factories at the same location were non-experimental but shared the assignment status of the randomly-selected factory.

3.2 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-

[19] 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.

[20] PCs are legally required for all factories with 50 or more workers located 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 called a Workers' Welfare Association (WWA).

[21] 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.

[22] A compound is a plot of land housing multiple factories at the same address.

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 data collection for this project is unique, as the research team had access to factories that likely would not opt into this type of research without the Alliance's request. The Alliance requested factories to cooperate as part of its impact evaluation efforts. Factories were told that the research team was conducting a general impact evaluation and were not told that we were specifically interested in the SC Program.

A five-member research team visited factories three times.[23] The 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 factory documentation. It also included photographing factory records for later digitization by the research team. The spotchecks of safety conditions entailed a trained assessor visiting the factory production floor and checking physical safety conditions against a checklist. They were only conducted at the second and third visits. The team leader was an assessor, who was responsible for managing interactions with management, verifying factory documentation, and implementing the safety spotchecks of the factory. A junior assessor supported the team leader. The junior assessor oversaw the the survey process, photographed factory records, and supported survey implementation. Three enumerators implemented surveys.

The first visit established factories' baselines. The second visit, approximately five months after baseline, aimed to measure outcomes immediately after treatment factories completed the most intensive portion of the SC Program. The third visit, approximately

[23]When invited to participate in these visits, factories were told that the purpose of the visits was general data collection for research on Alliance programs. Factories were not informed that the primary program of interest was the SC Program. More information about communications with factories and the data collection protocol is available upon request.

10 months after baseline, aimed to measure outcomes several months after treatment factories completed the SC Program. To minimize experimenter demand effects specific to the SC Program, the research team undertook extensive effort to design protocols for onsite visits to minimize the risks of non-truthful reporting and manipulation of data collection. For example, the research team verified several types of safety-related documentation, such training documents for other Alliance programs and inspection logs for factory machinery, in addition to documentation related to the SC.

The bottom of Table 2 presents baseline balance checks for characteristics of randomly selected worker survey participants at treatment and control factories. Workers at treatment and control factories are balanced on observable characteristics.

Figure 1 displays the experiment's timeline. As the timeline makes clear, for treatment factories, the onsite data collection is timed such that the intensive 3.5-month treatment phase is completed prior to the second visit. The second visit occurs during the intensive 2.5 month monitoring phase. The third visit is completed approximately four months after treatment factories complete the 6-month SC Program. Control factories do not participate in the SC Program until they have fully completed all activities related to the experiment.

3.2.1 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. SC effectiveness (index);

The first primary outcome is an index variable that summarizes factories' compliance with the SC regulation. I use a standardized 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, for how they operate, and for their responsibilities. Within each of these categories, there are numerous stipulations (see Table 1). The summary index measures the enforcement intervention's general effect on compliance with these requirements. To determine the variables included in the index, I enumerated the regulation's stipulations as variables. Whenever possible and relevant, I measure a factory's compliance with a stipulation using multiple sources information from the factory. For example, to determine how worker representatives to the SC were selected, I combine reports both from the SC President (a member of management) and from the SC worker representatives. I also incorporate variables to measure the consistency in information provided by different sources. Table A1 lists all sub-variables included in the index.

The second primary outcome aims to measure SCs' effectiveness at fulfilling the intent of the SC regulation. The regulation prescribes responsibilities for SCs related to management of physical factory safety, to training workers, and to safety culture. Correspondingly, the effectiveness index includes both physical and cultural indicators of SCs' effectiveness at improving factory safety. The index is comprised of the following sub-indexes or, in some cases, unique variables (see Table A2 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).[24]
- Factory safety culture:
  - Workers' awareness of SC.
  - Workers' safety knowledge.
  - Senior managers' awareness of the SC.

[24]Every Alliance-audited factory has a Corrective Action Plan (CAP) based on violations found in the Alliance's building safety audit. The CAP details the remediation actions that the factory will take to address the safety violations. The Alliance monitors factories' progress with implementing remediation and suspends factories that fail to make sufficient progress.

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 items included in the spotcheck of safety conditions. 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.[25]

Ideally, to measure SCs' effects on safety, I would have directly measured workers' injuries and illnesses. The intervention, however, aimed to empower workers to raise safety issues and concerns. Consequently, it could result in a net increase of reported injuries and illnesses even if it reduces the actual number of occurrences.[26]

Turning to the economic primary outcomes, these include:

3. Workers' job satisfaction and mental well-being (index);
4. Labor productivity;[27]
5. Employment;
6. Wages.

The third primary 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

[25]We excluded items that the SC could not plausibly work with management to change 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 only including these items in the 9-10 month check was twofold: First, the OSH expert identified several items that required more than 3-4 months, but plausibly less than 8-9 months, for the SC and management to address. Second, I wanted to reserve some spaces (e.g., bathrooms and chemical storage rooms) to only be visited during the third visit. I aimed to use this approach to gain insight into the extent to which management was responding to the research team's visits. Unfortunately, due to an administrative error, these additional items were not included in the spotcheck checklist in the third visit for 14 out of 80 factories. As such, in my analysis, I deviate from my PAP by not including the "endline only" checklist items.

[26]I am in the process of transcribing factories' injury and medical clinic records. While these are still subject to the same reporting concerns, they are the most comprehensive source of information available. They will be added to the analysis when they become available.

[27]In the pre-analysis plan, I pre-specified that I would analyze total factor productivity (TFP) or labor productivity. If I determined that I could not measure non-labor inputs to production with sufficiently high quality, I indicated that I would analyze labor productivity. Ultimately, I decided that I could measure labor productivity for more factories and with less measurement error. I will add an appendix that reports results for TFP.

data on worker turnover and absenteeism. Figure A3 lists all sub-variables included in the index.

The final three primary outcomes measure the interventions' 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.[28] 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 lists.[29] 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 use 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). Anderson's 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 nature of the roll-out, I was not able to collect a complete baseline before committing to the construction of my indexes. More broadly, summary index variables have the added benefit of reducing the number of hypotheses being tested. This approach reduces the risk of overrejection of the null hypothesis. It also increases my ability to detect marginally

[28]Table C1 presents estimated treatment effects on the log of the physical quantity of output and average weekly working hours, respectively. In both cases, the estimated treatment effects are close to zero and are not statistically significant.

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

statistically significant effects on multiple outcomes that, aggregated, achieve statistical significance ((Anderson, 2008)).

I also pre-specified secondary outcome variables that will allow me to analyze possible mechanisms underlying the effects on my primary outcome variables. In particular, I have multiple secondary outcome variables for workers, including workers' perception of their SCs' effectiveness, workers' perception of managers' human resource practices, workers' empowerment to raise safety and other issues, and workers' reported awareness of other types of worker organizations. All of these outcomes are index variables; Tables A4 through A7 display index components for secondary outcome variables. I report results on all secondary outcome variables in the Appendix. See Table C5 for baseline balance tests and Table B1 for the main results on these outcomes.

3.2.2 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. In this model, β is the coefficient of interest. I show results with and without controlling for the baseline value of the dependent variable. All of my 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 estimated difference between these two treatment effects. In the heterogeneity analysis, I report β_1 and $\beta_1 + \beta_3$ as well as the p-value for β_3.

For business competitiveness outcomes, which I measure using monthly administrative data, I also show panel regression results in the 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:

For statistical inference, instead of using the traditional sampling-based approach, I use randomization inference. Randomization inference 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.

3.2.3 Integrity of the Experiment

Baseline Balance

Table 2 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 0.13 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 my results including the outlier and to include an appendix in which I present baseline balance and the main the results after dropping this factory (Appendix part C). 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 treatment and control groups. The treatment group has somewhat more non-sewing factories, such as washing, printing, and accessories factories 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 of absence at the time that the factory was due to begin. The other two factories are located in the Chittagong Region of Bangladesh, where the Alliance implements 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 round data collection visit. All other factories complied with the treatment. I address the non-compliance issue by presenting Intent to treat (ITT) estimates. I will also present a full set of Local Average Treatment Effect (LATE) estimates, or the effect of treatment on the treated, in the next version of this paper. In this draft, Table C2 presents LATE estimates for the primary outcome index variables.

Attrition

Four factories have attrited from the sample. Two are treatment factories, and two are control factories. Three of the four were suspended by the Alliance due to their failure to make progress with physical building safety remediation. One control factory refused to participate in the second onsite visit. I address attrition by reporting Lee (2009) bounds on the treatment effects. In this draft, I report Lee bounds for the first three primary outcome variables (Table C3). For all three variables, there is minimal difference between the upper and lower bounds of the treatment effects, and with the exception of the lower bound for the SC effectiveness index, all estimates are statistically significant at the 5% or 10% level.

4 Results

I present the intervention's effects in four sub-sections. First, I present the effects of multinational enforcement on factories' compliance and on indicators of SC effectiveness. Next, I assess its consequences for workers' job satisfaction and well-being and for factories' business competitiveness. Third, I discuss whether the treatment effects are heterogeneous across factories. Finally, I explore whether the treatment effects persist after multinationals cease intensive enforcement.

4.1 Multinational enforcement , factories' compliance, and SCs' effectiveness

4.1.1 Factories' compliance at baseline

Before a factory begins the SC Program, the Alliance aims to verify that its SC was formed correctly. Specifically, the Alliance conducts verification visits to check whether factories' SCs are formed correctly; the Alliance also works with its members to verify that the bodies responsible for nominating worker representatives to the SC are democratically elected. When a factory begins the SC Program, the Alliance again checks that the SC was formed correctly and reforms 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. According to the labor law, factories were supposed 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.[30] Despite these non-compliances, 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-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 and were not yet central to the factory's safety policies in procedures. 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 those that had met, 89% of SCs maintained legally-required meeting minutes. In 77% of factories, there was no legally-required policy describing the SC's functions and responsibilities. There was also inconsistency in the information about SCs' operations across different sources of information: Presidents' reports matched factory documents and members' reports in about 58% and 57% of cases, respectively. There were some reports of management interference with the SC: In 10% of factories, at least one worker representative reported that management had offered bribes or otherwise attempted to block SC activities. 5% of presidents and 7% of worker representatives reported that they were not considered on duty for SC-related activities.

Many SCs were not implementing their legally-required safety responsibilities. For example, an important 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

[30] 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.

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%).[31]

4.1.2 The effects of multinational enforcement on compliance

Figure 5 and Table 3 present the results for the index of compliance with the SC regulation. Figure 5 compares the performance of treatment and control factories on the compliance index at the first and second data collection visits (pre-treatment phase and treatment phase, respectively). As evident in the figure, both groups start off performing similarly on the compliance index. At the second visit, though, the treatment factories' compliance with the SC regulation markedly improves compared to controls'. Control factories' compliance improves slightly but is mostly unchanged. Panel A of Table 3 shows that the ITT effect on compliance is 0.20-0.21 sds (FDR p=0.004). Evidently, the multinationals' enforcement intervention increases factories' compliance with the SC regulation above and beyond the effects of state-supplied enforcement and of the multinationals' other compliance programs.

Panel B of Table 3 displays the results for the formation, operations, and responsibilities sub-indexes (see Figure A1 for index components).[32] While treatment factories outperform control factories on all sub-indexes, by far the largest treatment effect is on the SC responsibilities sub-index. Treatment factories outperform control factories on this sub-index by 0.44 sds, which is statistically significant according to the RI and the FDR-sharpened p-values. The large, positive effect on this sub-index is consistent with the Alliance requiring factories to complete several legally-required activities during the SC Program. For example, at the second visit, only 15% of control SCs had conducted a risk

[31] 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.

[32] Treatment and control factories are balanced on all sub-indexes at baseline (Table C4).

assessment while 56% of treatment SCs had conducted at least one. According to reports by SC presidents, worker representatives, and senior managers, treatment SCs also made more regular safety reports and recommendations to senior management and followed up on these reports more frequently.

There are not, however, statistically significant treatment effects on the SC formation or operations sub-indexes. The lack of a treatment effect on SC formation is perhaps unsurprising in light of the Alliance's engagement with factory management prior to a factory's becoming eligible for the SC Program. Turning to the operations sub-index, both treatment and control factories improve their performance between the baseline and second visits. The improvement is consistent with SCs' recently becoming active around the baseline visit. It shows that in the absence of the multinationals' enforcement intervention, SCs still would have become more incorporated into factories' safety policies and procedures.[33]

Together, the results suggest that the enforcement intervention is most impactful for more "intensive margin" compliance outcomes that require the greatest amount of effort and engagement by factories. This interpretation of the results is supported by an analysis of the treatment effects on workers' perception of SCs' compliance and effectiveness, which I measure using a pre-specified index. Table B1 presents results on secondary worker outcome variables; column (2) of the first row shows that the intervention improves workers' perception of SC compliance and effectiveness by about 0.20 sds (RI *p*-val=0.097).[34]

These more "intensive margin" compliance outcomes are also arguably the most dif-

[33]One variable in the operations sub-index is whether a SC meets at least once per quarter. Although most SCs meet with the minimum required frequency, 88% of control SCs and 93% of treatment SCs at the second visit, the multinationals' intervention does increase SCs' meeting frequency. SCs' meeting frequency increases by 58%, from an average of 1.27 to 1.95 meetings per three months. This impact may contribute to improving indicators of SCs' effectiveness.

[34]Table C5 presents baseline balance tests for worker secondary outcome variables. While there are no statistically significant differences between the treatment and control groups, the treatment groups' means are uniformly lower than the control groups'. To minimize possible bias, I focus on results controlling for the baseline value of the dependent variable.

ficult for external monitors to observe. And there is evidence that information frictions remain; the multinationals' enforcement intervention does not bring all treatment factories into full compliance. For example, out of 36 treatment SCs that participated in the SC Program before the second data collection visit, 13 had not conducted a risk assessment. In a couple of these cases, the research team determined that the factory had falsified the risk assessment record, and in a few others, managers had conducted risk assessments, but not the SC. In all 13 cases, though, the Alliance's program records show that the SC had conducted a risk assessment before the visit day. Evidently, the multinationals' monitoring of compliance is imperfect, and many factories do not comply with all legal requirements, in particular those that are less easily-observed.

4.1.3 The effects of multinational enforcement on SC effectiveness

The multinationals' enforcement program increases factories' compliance with Bangladesh's SC regulation, in particular increasing SCs' fulfillment of legally-required responsibilities. The next critical question is whether these effects translate into improvements in factory safety. Figure 6 and Table 4 present the results for the index of indicators of SC effectiveness. As can be seen in Figure 6, treatment and control factories perform similarly at baseline. At the second visit, treatment factories again outperform the control factories, in this case, by about 0.14 sds. Table 4 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=0.105). This result provides causal evidence that multinationals' interventions to increase compliance with safety-related labor law can improve safety.

Figure C1 illustrates support for the extremeness of the result on SC effectiveness under the null hypothesis of no average treatment effect. The figure plots the joint distribution of compliance treatment effects and SC effectiveness treatment effects under the null hypothesis. The actual parameter estimates are indicated in red. As is evident in

the figure, the actual parameter estimates are one of the most extreme points on the joint distribution under the null hypothesis. The chance of jointly observing these effect sizes under the null hypothesis is extremely small.

Panel B of Table 4 presents the treatment effects for each sub-index. Baseline balance tests for these sub-indexes are presented in Table C4. There is one baseline imbalance on sub-index variable: Worker awareness of SCs at treatment factories is lower at treatment factories, although this difference lessens and is not significant at the 5% level when the outlier treatment factory is dropped (Table D1). 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 5 shows the treatment effects on each component of the spotcheck index.[35] Treatment factories outperform controls on nearly every sub-component. For example, workers in treatment 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.[36] 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. This effect is consistent with the dramatic increase in SCs' implementation of risk assessment at treatment factories. Table B2 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 enforcement intervention, risk assessment has a large, positive, statistically significant effect on their performance on the safety spotcheck.

[35]Four variables on the spotcheck checklist drop from the analysis because all factories were found to comply with these variables.

[36]PPE includes equipment such as eye guards, finger guards, chain mesh gloves, goggles, boots, among others.

Returning to Panel B of Table 4, the second row shows that the SC Program does not increase factories' progress on completing their CAPs for building safety violations. These violations often require significant financial investment and time to fix, and if the buyers' intervention increases SCs' ability to push management to make 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 (rows 3-5 of Panel B). 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 on worker awareness: At baseline, 81% of workers reported being aware of SCs' general role and responsibilities, and 89% knew that their factory had a SC. As shown in Table B3, even with very high baseline awareness of SCs, the enforcement intervention increases workers' awareness for both of these outcomes and for some other measures of worker awareness.

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

4.2.1 The effects of multinational enforcement on workers

A stated goal of the Alliance's SC Program is to provide workers with a worker-management body with democratically-selected worker representatives that ensures effective identification and resolution of workers' safety concerns. I hypothesized that increasing workers' voice in safety decision-making and improving safety inside the factory would lead workers to feel more satisfied with their jobs, more in control of their safety, and less stressed.

To the contrary, I find that the intervention negatively affects workers' job satisfaction.

Figure 7 shows treatment and control factories' performance on the job satisfaction and mental well-being index. The figure on the left, sub-figure (a), shows the full non-attrited sample, and the figure on the right, sub-figure (b), drops the negative outlier in the treatment group. Although the baseline difference including the outlier is not statistically significant, the figure on the right shows that the outlier does not drive the result. Figure 7 shows a statistically significant decrease in the job satisfaction and mental well-being index at treatment factories relative to controls. Turning to Table 6, the estimated treatment effect is approximately -0.15 sds (RI p= 0.057; FDR p=0.105). The table also shows that the treatment effect remains stable when a control for the baseline value of the index is added. It is also unchanged when the outlier factory is dropped from the analysis (Table D2).

Panel B of Table 6 displays the estimated treatment effects for the job satisfaction, mental well-being, turnover, and absenteeism sub-indexes/variables.[37] Focusing on column (2), the results reveal that the negative effect on the primary index is driven by a large, negative effect on the job satisfaction sub-index (-0.37 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. [38] Consistent with the null effect for turnover, the intervention does not affect workforce composition (Table C6). Together, these results rule out the possibility that changes in workforce composition drive the negative effect on job satisfaction.

To further unpack the negative effect on job satisfaction, Table B4 shows the estimated treatment effect on each sub-variable in the job satisfaction and mental-wellbeing 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. The

[37] Treatment and control factories are balanced on all sub-indexes at baseline (Table C4).

[38] 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 are then multiplied by -1 in order to be unidirectional with other outcomes. A higher value of the sub-variable indicates a lower turnover or absenteeism rate, respectively.

treatment effect is a 79% increase on the control group mean of 2.4% of workers. The proportion of workers who have referred family and friends to their factory at treatment factories also declines.

Why does the Alliance's enforcement intervention negatively affect workers' job satisfaction? In Section 4.3, I show that the negative effect is driven by factories with poor managerial practices where the intervention does not improve compliance or safety. I provide suggestive evidence that the negative effect on job satisfaction may be a result of 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 plausible mechanisms for the negative effect, such as workers learning about unsafe conditions at their factories. The data do not provide evidence in favor of learning about unsafe conditions driving the negative effect on job satisfaction (results not reported). I also continue to explore other possible mechanisms.

4.2.2 The effects of multinational enforcement on factories' business competitiveness

A critical question for this and other forms of labor regulation is what the costs are and who bears them (e.g., Besley and Burgess (2004), Botero et al. (2004)). If multinationals' interventions negatively affect targeted suppliers' productivity, then these suppliers are less able to compete against non-targeted alternatives. Unless multinationals reward compliant suppliers through increased prices or other channels, this dynamic would undermine the long-term viability of multinationals enforcing improved standards, as they would have an incentive to source from lower-cost suppliers. Further, if targeted suppliers' productivity falls, and labor markets are competitive, these suppliers may reduce their level of employment. Wages may also fall, in particular if compensation includes production-based incentives.[39] In this section, I test these possibilities by analyzing the

[39]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

intervention's effects on labor productivity, employment, and gross wages.

There is one control factory that partially or fully suspends production over three months between the first and second data collection visits.[40] 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 sample dropping the factory that partially shuts down.

Panel A of Table 7 shows the estimated treatment effects on labor productivity using the main regression model (equation 1). All three specifications include product-type fixed effects. In column (1), which includes the full sample, the estimated treatment effect is actually positive, a 11.5% increase. In column (2), which includes the trimmed sample, the estimated effect remains positive but is now smaller, an 8.2% increase. Finally, in column (3), which includes all observations except those for the factory that partially shuts down, the estimated effect is now a 3.6% increase. None of these estimated effects is statistically significant. The dramatic decrease in the estimated effect when the distribution is trimmed and the partial shutdown factory is dropped supports the interpretation that the intervention did not affect labor productivity. The same pattern of estimated effects are found using the panel regression model (Table C7). In light of the null results, Table B5 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 moderately large effects.

Panel B of Table 7 presents the estimated treatment effect on employment and on gross wages. Column (1) shows the estimated treatment on employment, which is a 0.3% increase in employment (RI $p = 0.872$). Table C7 presents the panel regression results. The

impacted.

[40]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.

estimated effect is a 1.2-1.3% decline in employment, which is not statistically significant. Turning to wages, both the main regression model and the panel find small, negative estimated treatment effects between 1.7-2.5% declines in gross wages. The effects are not statistically significant in either model. Again, Table B5 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. The results support the potential for multinationals' enforcement interventions to improve labor standards without coming at significant costs in terms of suppliers' efficiency. In the summer of 2019, I will incorporate the final three months of the 13 months of administrative data. These data may help me to rule out the possibility of smaller treatment effect sizes and to explore potential treatment effect dynamics.

4.3 Heterogeneous treatment effects by managerial capacity

In this sub-section, I explore heterogeneity in the intervention's effects depending on factories' baseline managerial capacity.[41] In my pre-analysis plan, I also 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 Appendix (see Tables C8, C9, and C10).[42]

My measure of managerial practices is a variable that summarizes senior and lower-level managers' reported frequency of holding production-related meetings with work-

[41] In this draft, I explore treatment effect heterogeneity for the first three primary outcomes. I will add an analysis for the latter three outcome variables.

[42] As can be seen in Table C8, there are large differences between the seven treatment and the seven control factories located in EPZs. For this reason, I depart from the PAP and do not analyze this dimension of heterogeneity.

ers.[43] 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).[44] The question measures one specific managerial practice; it was not feasible to conduct a complete management diagnostic. As such, it is reasonable to question whether this measure reflects broader managerial capacity. Figure B1 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 the meeting question).[45] Evidently, this question captures meaningful information about firms' overall managerial practices.

I partition the sample into above/below median groups using baseline values of the managerial practices variable. I refer to the below median group as poorly-managed establishments and to the above median group as better-managed establishments. Panel A of Table C8 shows baseline balance within each interaction-term subgroup for primary outcome variables for non-attrited establishments. There are no statistically significant differences between subgroups.

Panel A of Table 8 shows the 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

[43]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.

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

[45]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.

subgroups. The regression specification is equation 2.

Beginning with column (1), the estimated treatment effect on SC compliance for poorly-managed establishments is 0.11 sds (not statistically significant). In contrast, the estimated effect on better-managed establishments is nearly three times as large (0.31 sds, RI $p \approx 0.000$). The difference between these estimates approaches marginal statistical significance (RI p=0.121). For poorly-manged establishments, the small improvement in compliance translates into a very small improvement in SC effectiveness (0.05 sd effect, RI p=0.563). In contrast, consistent with the large effect on compliance, better-managed factories improve SC effectiveness 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 establishments. The estimated treatment effect on job satisfaction and mental well-being for poorly-managed factories is -0.27 sds (RI p=0.029). The estimated effect for better-managed factories, though, is close to zero (-0.034 sd effect, RI p=0.721)). The RI p-value for the difference in the treatment effects for these groups is p=0.168. Evidently, the decline in job satisfaction found in Section 4.2.1 is driven by poorly-managed establishments, for which the intervention does not result in meaningful improvements in compliance or in SC effectiveness.

I can try to increase statistical power to detect differences between the treatment effects for poorly- and better-managed establishments 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. But if the differences between groups are stable, it will help. Panel C of Table 8 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 reject the null hypothesis of equality of treatment effects for above and below median groups at RI p=0.084. Turning to column (2), the estimated treatment effects on SC effectiveness, although attenuated, exhibit the same pattern. I remain unable, however, to reject that the effects for both groups are equal (RI p=0.239). Finally, in column (3), the difference in effects is fairly stable, and I reject the null of equality with RI p=0.064. In sum, the results of the pooled analysis support the interpretation that multinationals' enforcement has differential effects on poorly- and better-managed establishments.

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 B6 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=0.075). In column (2), in which SC effectiveness is outcome, the above median managerial practices interaction term is again largest in magnitude (RI p=0.178). 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=0.096). Together, the results show that managerial practices are an important dimension of heterogeneity after controlling for factories' other characteristics and their interaction with the treatment.

Another concern about the heterogeneity results is that multinationals may more intensively monitor less compliant factories and that this generates the heterogeneous ef-

fects. In this case, one would expect the Alliance to be more likely to audit factories that are less compliant with the SC law at baseline. 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 of SCs could not drive the heterogeneous effect patterns in Table 8.

Finally, I implement the analysis using an alternative measure of managerial 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 variable to measure worker-manager relations (see Table A5 for index components)). I find a qualitatively similar pattern of heterogeneous treatment effects using this variable as with my main managerial practices measure. See the bottom panels of Tables C8- C10.

To summarize the heterogeneity analysis, the results show that organizational capacity plays an important role in determining the effect of labor regulation enforcement on factories' and workers' outcomes. The multinationals' enforcement intervention improves compliance and SC effectiveness only in factories with better managerial capacity. The improvements at these factories 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 workers.

Why does job satisfaction decline at poorly-managed factories?

Why does job satisfaction appear to decline at poorly-managed factories when the intervention is having little to no effect on compliance and SC effectiveness? One plausible mechanism is that the Alliance's intervention raises workers' expectations about what SCs will deliver, but in poorly-run factories, SCs' performance fall short of workers' expectations, and they are disappointed. This effect would be consistent with recent

findings from an experiment with low-skill workers in India by Adhvaryu, Nyshadham and Xu (2018). The authors find that a randomly-assigned upgrade in employer-provided housing that improved objective quality measures reduced workers' job satisfaction and increased turnover. The authors provide evidence that the negative effects were due to the improvement in quality falling short of workers' expectations.

I cannot directly test that the negative effect on job satisfaction in poorly-managed factories is due to unmet expectations, as I did not collect data on workers' expectations for SCs. I find support for an important role for workers' expectations and learning about the SCs' role from qualitative evidence gathered from interviews with compliance managers from eight treatment factories. 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). In these cases, managers sometimes indicated that the SC relayed the concern to a separate committee responsible for these issues. If these concerns go unresolved, though, it is easy to see why workers may be disappointed, even if their factories' SCs are fulfilling their legal responsibilities. It is also unsurprising that workers may not initially understand the scope of SC authority, as the SC Program is possibly the first time that workers have been informed that there is an institution inside the factory responsible for addressing any type of worker concern. I do not have data with which I can directly test this possibility; while I have access to records of issues raised to the SC, SCs only recorded the safety-related issues in these documents. We did not ask workers about the specific issues that they raised to the SC. If this mechanism contributes to lower job satisfaction at treatment factories, though, it suggests that the negative effect on job satisfaction may be temporary. I test this possibly using the third data collection round.

4.4 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 SC effectiveness after the multinationals cease intensive enforcement? Do the adverse effects on workers' job satisfaction in poorly-managed factories persist?

Table 9 presents the estimated treatment effects on the first three primary outcome variables measured 3-4 months after the end of the intensive enforcement period for treatment factories. Column (3) of the first row shows that the estimated treatment effect on compliance persists. The estimated treatment effect remains around 0.21 sds (FDR p=0.078). Treatment factories continue to outperform controls on the responsibilities sub-index, and also begin to outperform controls on the operations sub-index (results not reported).

While treatment factories continue to surpass controls on the compliance index, they no longer do on the SC effectiveness index. Treatment factories outperform controls on this index by 0.06 sds (RI p=0.371). The difference in treatment and control factories' performance on the research team's spotcheck of safety conditions is no longer statistically significant. Further, to the extent that workers' awareness of SCs at treatment factories improved during the treatment phase, these differences fade.

While treatment factories' performance on the SC effectiveness index attenuates slightly, control factories also improve their performance on this index and on the compliance index. I argue that this convergence is partially 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 of compliance and HR managers of factories under shared-ownership with experimental factories conducted as part of this research, 46% of

managers whose factories had participated in the SC Program report taking actions in advance to prepare.[46]

Interestingly, the negative effect on the job satisfaction and mental well-being index disappears, and the estimated treatment effect is actually positive (0.11 sd effect, RI p=0.201).

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

I do not find evidence of delayed adverse effects on labor productivity, employment, or wages. The estimated treatment effects for all variables remain close to zero. Appendix Table C12 presents the results.

5 Conclusion

In this paper, I analyze the effects of a coalition of multinationals' 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. 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 in-

[46]Further, my agreement with the Alliance was 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 in the SC Program.

crease collective worker voice inside the firm. In addition, through my collaboration with some of the world's largest multinationals, the study has provided unique evidence from a population of factories that would otherwise be unlikely to participate in academic research.

I find that the multinationals' 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 has a small, positive effect on SCs' effectiveness at improving safety. In particular, it improves factories' performance on independent checks of physically-measurable safety conditions and increases workers' knowledge and awareness of SCs. These findings demonstrate that private enforcement of labor law can significantly improve compliance and contribute to achieving the law's objectives. They provide reason for greater optimism regarding the ability of private enforcement to improve labor standards in developing countries compared to existing research on this topic in political science, which is largely skeptical but which lacks causal evidence. It also provides the first experimental evidence that firms' CSR initiatives can successfully generate public goods/curtail public bads. Finally, it provides empirical justification for theoretical models of CSR that identify CSR with the private creation of public goods or curtailment of public bads, such as those of Besley and Ghatak (2007) and Lai et al. (2017).

Pre-specified subgroup analysis reveals that the multinationals' enforcement intervention is only effective at improving compliance and safety in better-managed factories. The estimated treatment effects on these factories are large. In contrast, the intervention does not improve compliance or safety in poorly-managed factories. Further, workers in poorly-managed factories respond more negatively to survey questions related to job satisfaction. I provide suggestive evidence that this result may be due to workers' disappointment with their factories' SCs.

These results have important implications for economic theory, multinational firm

strategy, and policymaking. Specifically, they show that organizational capacity in the private sector matters for the efficacy of labor law enforcement in developing countries. Multinationals that aim to enforce local or industry standards on their suppliers need to take into account their suppliers' organizational capacity. Their interventions can have large, beneficial effects when suppliers have capacity to meet higher labor standards. But by raising workers' expectations regarding improvements in factories' that do not have capacity to implement them, their intervention adversely impacts indicators of workers' job satisfaction in the short-run.

My analysis of the intervention's effects on labor productivity, employment, and wages does not provide evidence of negative effects. Estimated effects on labor productivity, employment, and wages are either positive or negative and close to zero. They are not statistically significant. 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, there is evidence that the treatment effect on compliance persists beyond the period of intensive enforcement by multinational buyers. Treatment factories also continue to outperform controls on an index of indicators of SC effectiveness, although the difference is not statistically significant. The evidence suggests that treatment factories may maintain, in particular, those improvements that are more easily observed by the multinationals. In contrast, the intervention's negative effect on workers' job satisfaction dissipates, consistent with the short-term negative effect possibly being due to a disappointment mechanism.

My findings raise several important directions for future research. First, this research highlights an important constraint on the efficacy of labor regulation in developing countries, which is the organizational capacity of the private sector. Future research can more fully investigate how firms' organizational capacity supports compliance with labor laws.

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 (2018) 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 make trade-offs between wages and workplace risks. 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*.

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2019. "Management and Shocks to Worker Productivity." *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.

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.

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

Bachmann, Ruediger, Gabriel Ehrlich, Ying Fan, and Dimitrije Ruzic. 2019. "Firms and Collective Reputation: A Study of the Volkswagen Emissions Scandal." *Working Paper*.

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*.

Blattman, Christopher, and Stefan Dercon. 2018. "The Impacts of Industrial and Entrepreneurial Work on Income and Health: Experimental Evidence from Ethiopia." *American Economic Journal: Applied Economics*, 10: 1–38.

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.

Bloom, Nicholas, Genakos. Christos, Ralf Martin, and Raffaella Sadun. 2010. "Modern Management: Good for the Environment or Just Hot Air?" *The Economic Journal*, 120: 551–572.

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. 2018. "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.

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*.

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.

Distelhorst, Greg, Jens Hainmueller, and Richard M Locke. 2017. "Does Lean Improve Labor Standards? Management and Social Performance in the Nike Supply Chain." *Management Science*, 63(3): 707–728.

Distelhorst, Greg, Richard M Locke, Timea Pal, and Hiram Samel. 2015. "Production goes global, compliance stays local: Private regulation in the global electronics industry." *Regulation and Governance*.

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.

Gosnell, Greer K., John A. List, and Robert D. Metcalfe. 2019. "The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains." *Working Paper*.

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*.

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.

Humphreys, Macartan, Rául Sánchez de la Sierra, and Peter Van der Windt. 2017. "Institutional Change by Imitation: Introducing Western Governance Practices in Congolese Villages." *Working Paper*.

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

International Trade Union Confederation. 2018. "2018 ITUC Global Rights Index: The World's Worst Countries for Workers." International Trade Union Confederation, Brussels, Belgium.

Lai, Chien-Yu, Andreas Lange, John A List, and Michael K Price. 2017. "The Business of Business is Business: Why (Some) Firms Should Provide Public Goods when they Sell Private Goods." *Working Paper*.

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.

Levine, David I, Michael W Toffel, and Matthew S Johnson. 2012. "Randomized Government Safety Inspections Reduce Worker Injuries with No Detectable Job Loss." *Science*, 336: 907–911.

List, John A, Azeem M Shaikh, and Yang Xu. 2016. "Multiple hypothesis testing in experimental economics." *NBER Working Paper*.

Locke, Richard M. 2013. *The Promise and Limits of Private Power: Promoting Labor Standards in a Global Economy.* Cambridge, United Kingdom:Cambridge University Press.

Locke, Richard M, and Monica Romis. 2010. "The promise and perils of private voluntary regulation: Labor standards and work organization in two Mexican garment factories." *Review of International Political Economy*, 17(1): 45–74.

Locke, Richard M, Fei Qin, and Alberto Brause. 2007. "Does monitoring improve standards? Lessons from Nike." *Industrial and Labor Relations Review*, 61(1): 3–31.

Locke, Richard M, Matthew Amengual, and Akshay Mangla. 2009. "Virtue out of Necessity? Compliance, Commitment, and the Improvement of Labor Conditions in Global Supply Chains." *Politics and Society*, 37(3): 319–351.

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*.

Malesky, Edmund, and Markus Taussig. 2017. "The Danger of Not Listening to Firms: Government Responsiveness and the Goal of Regulatory Compliance." *Academy of Management Journal*, 60(5): 1741–1770.

McDevitt, Ryan C. 2011. "Names and Reputations: An Empirical Analysis." *American Economic Journal: Microeconomics*, 3(3): 193–209.

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.

Short, Jodi L, Michael W Toffel, and Andrea R Hugill. 2016. "Monitoring global supply chains." *Strategic Management Journal*, 37: 1878–1897.

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

Tadelis, Steven. 2002. "The Market for Reputations as an Incentive Mechanism." *Journal of Political Economy*, 110(4): 854–882.

Tanaka, Mari. 2019. "Exporting Sweatshops? Evidence from Myanmar." *Working Paper*.

Toffel, Michael W, Jodi L Short, and Melissa Ouellet. 2015. "Codes in context: How states, markets, and civil society shape adherence to global labor standards." *Regulation and Governance*, 9: 206–223.

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*.

6 Figures and Tables

Figure 1: RCT timeline

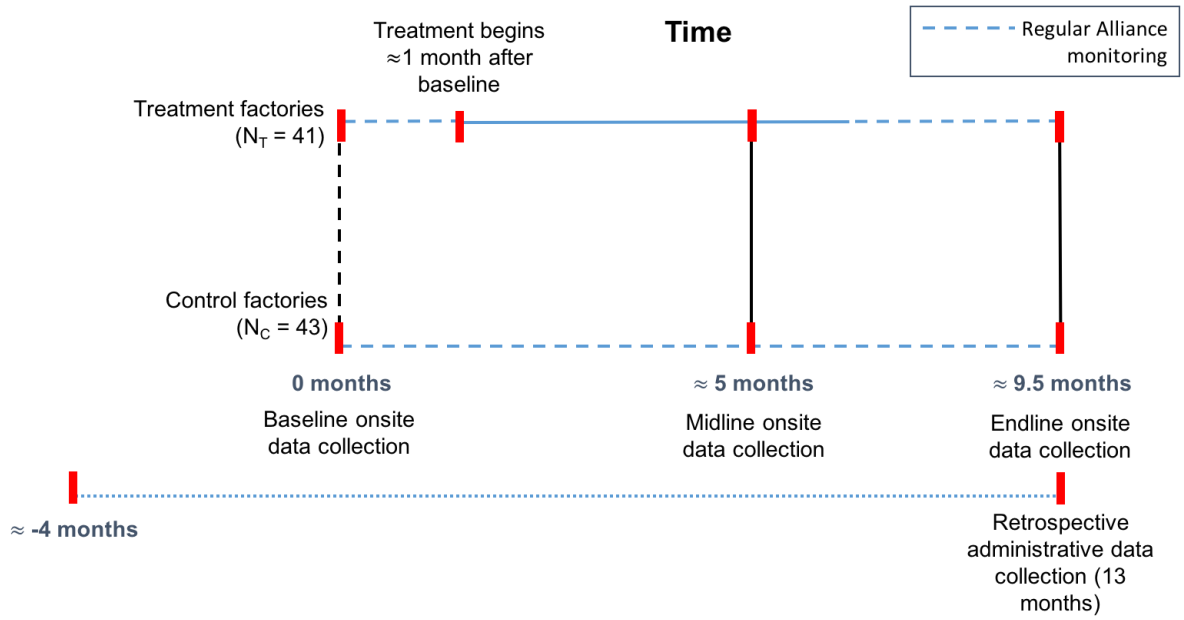
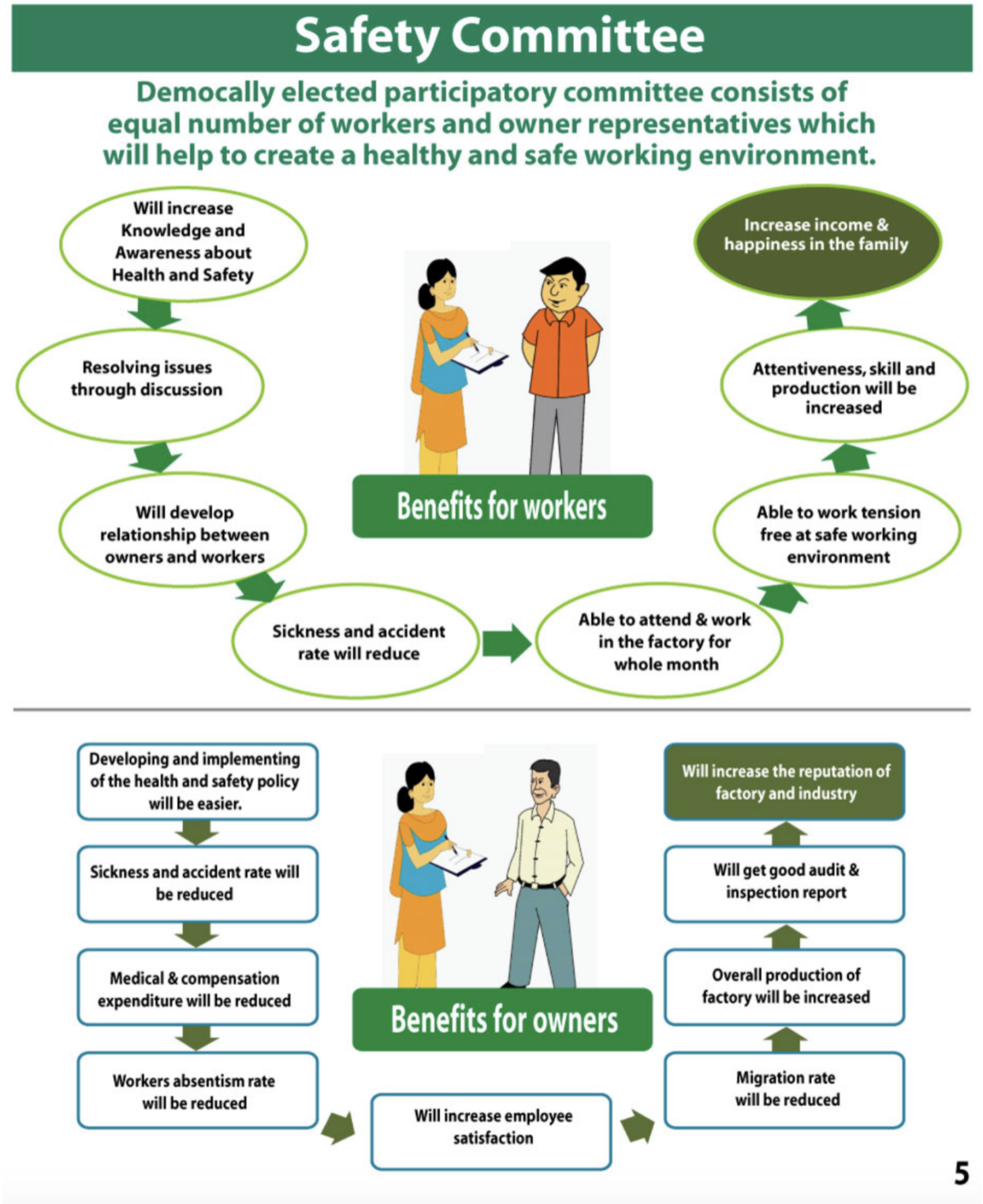


Figure 2: Rana Plaza building collapse

Source: Wikipedia.

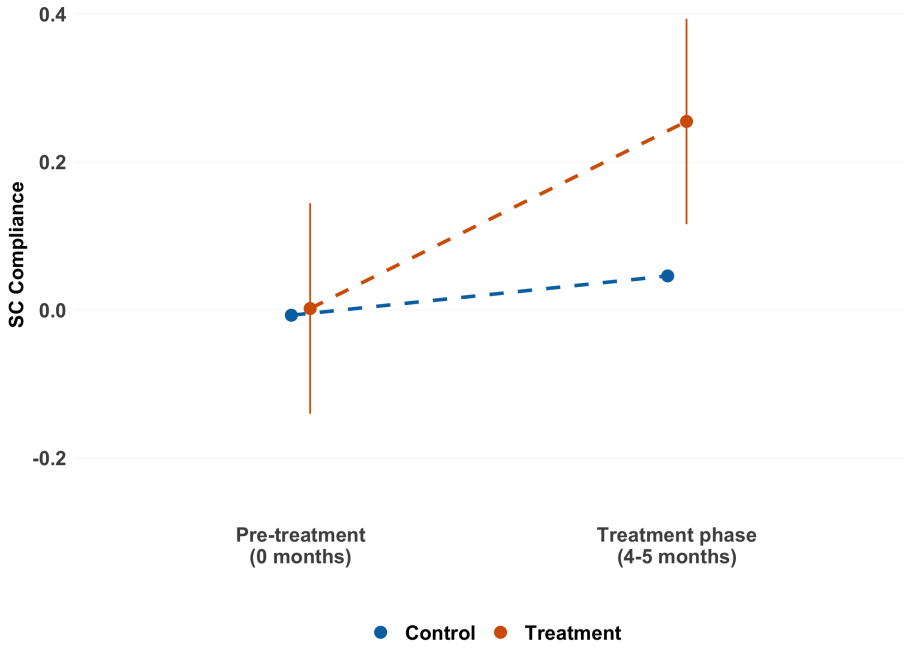
Figure 3: Alliance member companies

Figure 4: Alliance SC Theory of Change



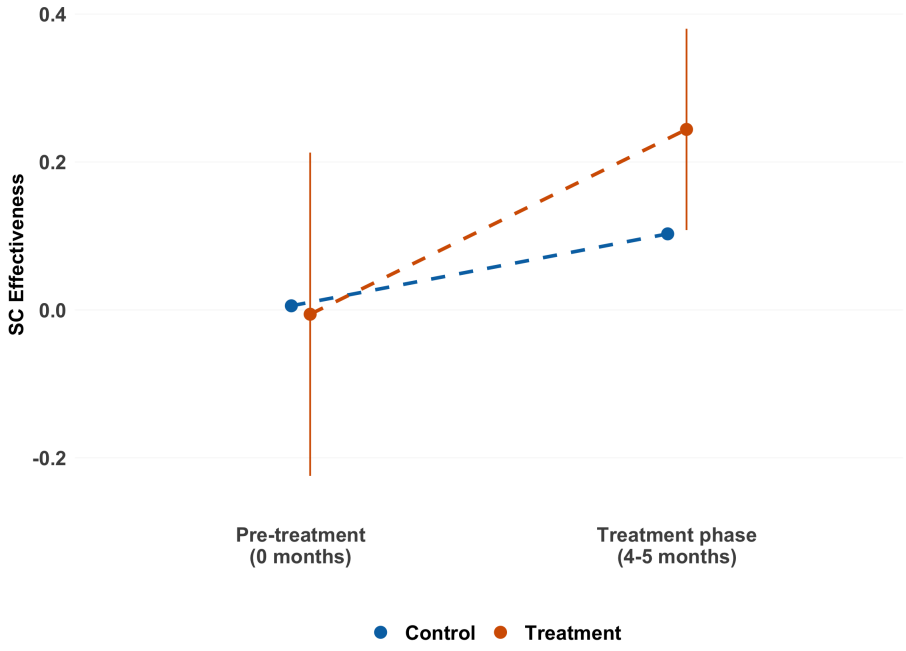
Source: Alliance training materials for SC members (English translation).

Figure 5: Pre-specified index: SC Compliance



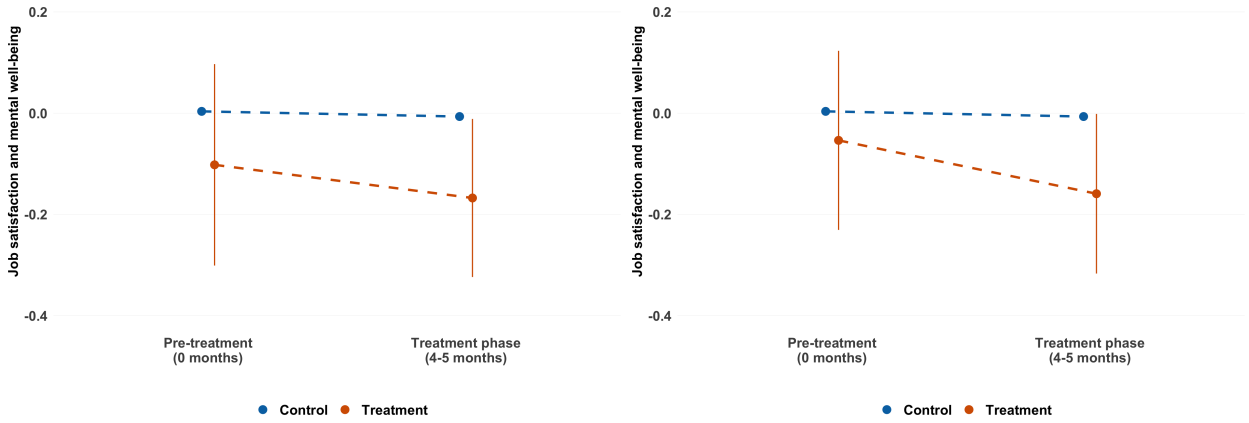
Notes: FDR p-val=0.020 for difference in post-treatment means. Whiskers show the the 95% confidence interval calculated from regressions of the outcome variable on a treatment indicator and stratification variables separately for pre-treatment and post-treatment rounds using robust standard errors. Summary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

Figure 6: Pre-specified index: SC Effectiveness



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 7: Worker job satisfaction and mental well-being



(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 1: Key Safety Committee 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

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

Table 2: 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	0.465	84
Effectiveness index	0.002	-0.053	0.627	84
Job satisfaction & well-being index	-0.001	-0.130	0.209	84
Number of employees	1190	-156	0.628	80
Gross wages (log)	15.82	-0.196	0.456	72
Labor productivity (log)	1.575	0.398	0.271	75
Labor productivity (log)†	1.575	-0.049	0.741	75
Labor productivity (log)†, trimmed sample	1.538	-0.053	0.726	75
Panel B: Factory characteristics				
Trade union at factory (1=Yes)	0.047	-0.045	0.334	84
EPZ(1=Yes)	0.163	0.036	0.687	84
Sewing (only)	0.47	-0.13	0.250	84
Number product types	1.16	0.05	0.663	84
Monthly absenteeism (%)	4.85	-0.66	0.468	80
Monthly turnover (%)	3.98	-0.69	0.490	80
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	27.18	0.19	0.824	84
Proportion female	0.57	-0.11	0.084*	84
Education (yrs)	6.22	-0.43	0.248	84
Tenure (yrs)	3.85	-0.18	0.717	84
Prior industry experience (yrs)	1.54	0.041	0.861	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 factories in the 1st and 99th percentiles of labor productivity. *p<0.1; **p<0.05; ***p<0.01.

Table 3: Treatment effects: Compliance with SC regulation

	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	0.080
		[0.298]
		{0.424}
Operations sub-index	0.184	0.057
		[0.522]
		{0.534}
Responsibilities sub-index	-0.149	0.442
		[0.000]***
		{0.004}***
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 4: Treatment effects: Indicators of SC effectiveness

	Control mean	ITT Effect
	(1)	(2)
Panel A: Primary outcome		
SC Effectiveness	0.103	0.143
		[0.046]**
		{0.105}
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	0.378	-0.065
		[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 index of SC effectiveness 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 5: 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 [0.621]
Cutting: Machines have knife guards *and* workers wear PPE for their task	0.792	0.071 [0.561]
Dyeing and jobs handling chemicals: Safety masks, goggles, gloves, aprons, and boots worn by workers handling chemicals	0.545	0.102 [0.674]
All PPE appropriate size, functional, and well-maintained	0.951	0.050 [0.260]
Aisles clearly marked and markings visible	0.780	0.052 [0.565]
Aisles clear of sewing scrapes and debris	0.951	0.048 [0.300]
Aisles clear of obstruction	0.854	0.014 [0.867]
Machines in good working order & dangerous parts properly covered	0.927	0.070 [0.153]
Work stations maintained in tidy condition (no loose materials close to electrical appliances)	0.976	0.022 [0.726]
One or more easily accessible first aid kit in section	0.976	0.022 [0.726]
Physical separation between storage & production areas	0.976	-0.005 [0.997]
Drinking water easily accessible for all workers	1.000	-0.025 [0.568]
Drinking water provided appears clean (visual check)	1.000	-0.025 [0.568]
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 A2). 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 6: 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)	-0.007	-0.149 [0.057]* {0.105}
Panel B: Sub-indexes and sub-variables		
Job satisfaction sub-index	-0.100	-0.373 [0.017]** {0.075}*
Mental well-being sub-index	0.003	-0.058 [0.712] {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 7: Treatment effects: Business competitiveness outcomes

	(1)	(2)	(3)
Panel A	Log(Labor productivity)		
Treatment effect	0.115	0.082	0.036
	[0.148]	[0.189]	[0.392]
			{0.418}
Factories	75	75	74
Observations	375	370	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(Employment)	Log(Gross wages)	
Treatment effect	-0.011	-0.015	
	[0.635]	[0.612]	
	{0.466}	{0.466}	
Factories	80	72	
Observations	400	360	
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-product type. In column (2), the sample is trimmed at the 1st and 99th percentile of 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 8: Heterogeneous treatment effects by managerial practices

	SC Compliance	SC Effectiveness	Job Satisfaction & Mental Well-being
	(1)	(2)	(3)
Panel A: Treatment Phase			
Below median	0.105	0.050	-0.268
	[0.225]	[0.563]	[0.029]**
Above Median	0.312	0.235	-0.034
	[0.000]***	[0.049]**	[0.7212]
p-val, diff	[0.121]	[0.236]	[0.168]
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	0.008
	[0.327]	[0.783]	[0.950]
Above Median	0.292	0.090	0.238
	[0.014]**	[0.301]	[0.067]*
p-val, diff	[0.356]	[0.693]	[0.210]
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	-0.130
	[0.129]	[0.583]	[0.149]
Above Median	0.302	0.162	0.101
	[0.000]***	[0.023]**	[0.243]
p-val, diff	[0.084]*	[0.239]	[0.064]*
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 9: Persistence of treatment effects after end of intensive enforcement: Primary Outcome Index Variables

	Control mean	ITT Effect
	(1)	(2)
SC Compliance	0.149	0.208
		[0.012]**
		{0.078}*
SC Effectiveness	0.156	0.063
		[0.371]
		{1.000}
Job satisfaction & mental well-being	-0.096	0.113
		[0.201]
		{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 A1: SC Compliance Index (primary outcome)

	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	President is management member and Vice President is worker member	Factory Namelist	
4	Formation	Compliant worker representative selection process: CBA, PC, or WWA as required	SC President Survey	SC Worker Rep. Survey
5	Formation	Management does not select worker representatives on SC	SC President Survey	SC Worker Rep. Survey
6	Formation	In factories with >= (1/3) female workforce, at least (1/3) worker representatives are female	Factory documentation	
7	Formation	Factory maintains list of SC Members	Factory documentation	
8	Formation	Correlation between SC President's reports and factory documentation	Factory documentation	SC President Survey
9	Operations	Factory maintains description of SC Members' roles and responsibilities	Factory documentation	
10	Operations	Factory Safety Policy includes a section on the Safety Committee's role and responsibilities	Factory documentation	
11	Operations	Safety Committee meets at least 1 time per 3 months	Factory documentation	
12	Operations	Frequency of meetings per 3 months	Factory documentation	
13	Operations	Meeting minutes are available for all Safety Committee meetings in past three months	Factory documentation	
14	Operations	Meeting attendance lists are available for all Safety Committee meetings in past three months	Factory documentation	
15	Operations	Safety Committee members have received training in their role on SC	SC President Survey	SC Worker Rep. Survey
16	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
17	Operations	Safety Committee uses compliant decision rule (unanimous or majority vote)	SC President Survey	SC Worker Rep. Survey
18	Operations	Correlation between SC President's reports and factory documentation	Factory documentation	SC President Survey
19	Operations	Correlation between SC President's reports and SC worker member reports	SC President Survey	SC Worker Rep. Survey
20	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	
21	Responsibilities	Safety Committee has completed a risk assessment of the factory	Factory documentation	
22	Responsibilities	Safety Committee has developed an action plan for safety improvements	Factory documentation	
23	Responsibilities	Safety Committee makes regular safety reports/recommendations to management	SC President Survey	SC Worker Rep. Survey
24	Responsibilities	Frequency of follow-up: Regular reports and recommendations to management	SC President Survey	SC Worker Rep. Survey

25	Responsibilities	Senior management frequency of reports from SC (should be minimum 1x quarter)	Senior Manager Survey	
26	Responsibilities	Safety Committee organizes training and fire drills	SC President Survey	SC Worker Rep. Survey
27	Responsibilities	Number of fire drills, previous 3 months	Factory documentation	
28	Responsibilities	Proportion of workers who report participation in safety-related training	Worker Survey	
29	Responsibilities	Proportion of workers who report participation in fire drill	Worker Survey	
30	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)
31	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)
32	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)
33	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)
34	Responsibilities	Safety Committee investigates accidents and make recommendations to prevent future accidents	SC President Survey	SC Worker Rep. Survey
35	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

Notes: In my analysis, I deviate from my PAP by not including the third sub-variable in the formation sub-index ("President is management member and Vice President is worker member"). Due to my own 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.

Table A2: SC Effectiveness Index (primary outcome)

	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	Floor Spotcheck	Machines are appropriatly placed and spaced (1 meter from wall with 1 meter aisles between) (machines in section)	Floor Spotcheck (endline only)
20	Floor Spotcheck	Fire doors are installed, unlocked, and without obstruction (in section)	Floor Spotcheck (endline only)
21	Floor Spotcheck	Toilet facilities in section are clean, functional (clean running water and soap), and provide privacy (stalls with doors)	Floor Spotcheck (endline only)
22	Floor Spotcheck	Chemicals are stored separately from production activites in a well-ventilated room	Floor Spotcheck (endline only)
23	Floor Spotcheck	Chemicals are stored in appropriate containers and containers are stored in an orderly fashion	Floor Spotcheck (endline only)

24	Floor Spotcheck	Material safety data sheets (MSDS) are prominently posted in both storage and use zones, and maintained in languages understood by workers	Floor Spotcheck (endline only)
25	Floor Spotcheck	Chemicals and hazardous substances are properly labelled as per label instructions/MSDS	Floor Spotcheck (endline only)
26	Floor Spotcheck	No loose wiring visible in production area (in section)	Floor Spotcheck (endline only)
27	CAP	Percent compliant	Alliance CAP data
28	Worker Awareness	Proportion of workers aware that factory has a SC	Worker Survey
29	Worker Awareness	Proportion of workers aware of SC's function and responsibilities	Worker Survey
30	Worker Awareness	Proportion of workers aware of how to contact SC member with issue	Worker Survey
31	Worker Awareness	Proportion of workers aware of SC topic-specific responsibilities	Worker Survey (midline, endline only)
32	Worker Safety Knowledge	Proportion of workers correctly answer fire question	Worker Survey
33	Worker Safety Knowledge	Proportion of workers correctly answer earthquake question	Worker Survey
34	Senior Manager Awareness	Senior management can provide at least one example of one issue identified by SC that has been resolved	Senior Manager Survey

Notes: In my analysis, I deviate from my PAP by not including the ”endline only” checklist items. Due to an administrative error, these additional items were not included in the checklist for 14 out of 80 factories. Thus, I do not include these items in my main analysis.

Table A3: Worker Job Satisfaction and Mental Well-being Index (primary outcome)

	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
#	Turnover		Factory Questionnaire

Table A4: Perceived SC Compliance & Effectiveness Index (secondary outcome)

	Variable	Variable Source
1	Proportion of workers reporting compliant worker representative selection process: CBA, PC, or WWA as required	Worker Survey
2	Proportion of workers reporting management selects worker representatives on SC	Worker Survey
3	Proportion of workers reporting participation in activities organized by the SC	Worker Survey
4	Proportion of workers reporting SC provides reports of health and safety issues and/or recommendations on these issues to senior management	Worker Survey
5	Mean reported responsiveness of SC to workers' concerns	Worker Survey
6	Proportion of workers that think SC would be able to affect the factory's safety policies if learn their concerns	Worker Survey
7	Mean reported extent to which SC helps to improve safety at the factory	Worker Survey
8	Proportion of workers that report that factory management implemented one or more of the SC's recommendations	Worker Survey

Table A5: Perceived Human Resource Practices/Relations Index (secondary outcome)

	Variable	Variable Source
1	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
2	Can workers at your factory make anonymous reports/recommendations?	Worker Survey
3	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
4	To what extent does management care about workers' safety?	Worker Survey
5	To what extent do workers and management discuss and make plans to improve safety together at your factory?	Worker Survey
6	If you got hurt at work, would you be concerned that management may punish you because of reporting your injury to management?	Worker Survey
7	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
8	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
9	Mean reported comfort sharing safety concern with senior managers	Worker Survey
10	Mean reported comfort sharing safety concern with direct supervisor	Worker Survey
11	I have seen supervisors verbally or physically abuse me or my coworkers.	Worker Survey
12	Worker unrest	Factory Questionnaire

Notes: In my analysis, I deviate from my PAP by not including the worker unrest variable.

Table A6: Worker Empowerment Index (secondary outcome)

	Variable	Variable Source
1	How confident are you in your ability to identify unsafe conditions at your factory?	Worker Survey
2	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
3	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
4	If you were to have a safety concern, would you report it?	Worker Survey
5	If you were to get hurt at work, would you report the incident?	Worker Survey
6	Mean reported comfort sharing safety concern with member of SC	Worker Survey
7	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
8	Do you feel that if you wanted to change jobs, you could?	Worker Survey
9	Do you have a goal for job promotion at your factory (e.g., operator level, process supervisor or production line manager)?	Worker Survey

Table A7: Worker Organization Awareness Index (secondary outcome)

	Variable	Variable Source
1	Familiar with the function and responsibilities: Trade union	Worker Survey
2	Familiar with the function and responsibilities: Participation Committee (non-EPZ only)	Worker Survey
3	Familiar with the function and responsibilities: Workers' Welfare Association (EPZ only)	Worker Survey
4	Awareness of factory's PC (non-EPZ only)	Worker Survey
5	Awareness of factory's WWA (EPZ only)	Worker Survey

B: Figures and Tables

Figure B1: Correlation between WMS Management Index (excluding meeting question) and WMS Meeting-related Question, apparel firms in all countries

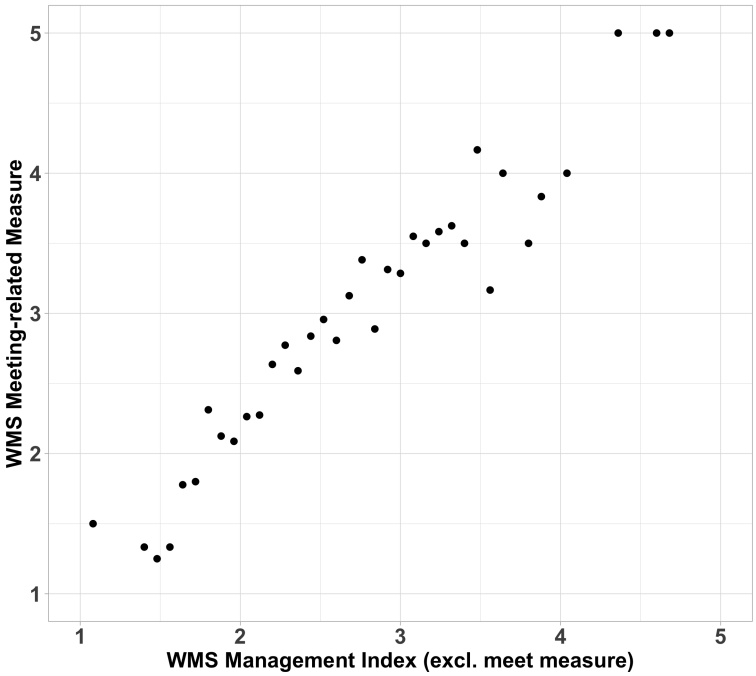


Table B1: Treatment effects: Worker Secondary Outcome Variables

	Control mean	ITT Effect
	(1)	(2)
Perceived SC compliance & effectiveness index	-0.109	0.195
		[0.097]*
Perceived worker-manager relations index	0.072	-0.039
		[0.632]
Worker empowerment index	-0.178	0.067
		[0.488]
Worker organization awareness index	0.073	0.016
		[0.851]
Observations		80
Stratification variables		Y
Control, base. dep. var.		Y

Notes: This table reports OLS estimates of treatment effects on secondary outcome index variables. 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. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01

Table B2: 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 B3: Treatment effects: Worker awareness outcome variables

	Control mean	ITT Effect
	(1)	(2)
SC Effectiveness 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 SC Effectiveness 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 B4: 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 B5: Ex post minimum detectable effect sizes (MDEs): Business competitiveness outcomes

	Control mean (sd)	MDE
	(1)	(2)
Log(Labor productivity)†	1.527	0.185
	(1.313)	
Log(Employment)	6.665	0.042
	(1.038)	
Log(Gross wages)	15.865	0.070
	(1.080)	

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 labor productivity.

Table B6: Heterogeneous treatment effects: Testing the importance of each dimension of heterogeneity

	Dependent variable:		
	Compliance	SC Effectiveness	Job satisfaction & mental well-being
	(1)	(2)	(3)
Treat	0.167	0.023	-0.550
	[0.275]	[0.866]	[0.014]**
Treat*Abv med Mgmt	0.242	0.193	0.306
	[0.075]*	[0.178]	[0.096]*
Treat*Abv med Size	-0.140	0.110	0.119
	[0.271]	[0.441]	[0.461]
Treat*Abv med Compli	-0.037	-0.092	0.387
	[0.792]	[0.522]	[0.037]**
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.

Table B7: Heterogeneous treatment effects by managerial practices, business competitiveness outcomes

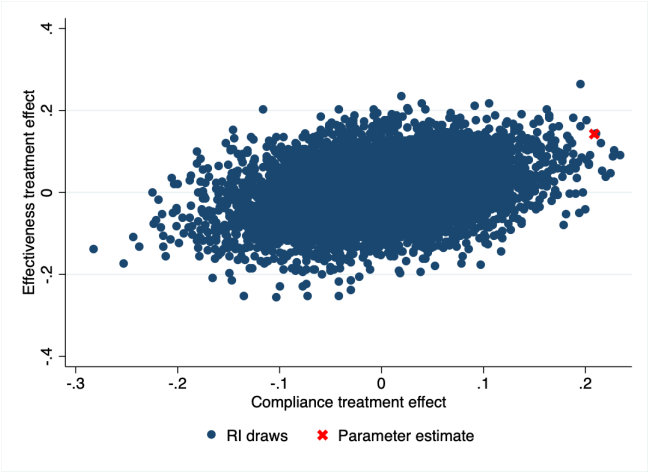
	Log(Labor productivity)		Log(Employment)	Log(Gross wages)
	(1)	(2)	(3)	(4)
Panel A: Treatment Phase				
Below median	0.161	0.049	-0.01	-0.015
	[0.127]	[0.403]	[0.754]	[0.673]
Above Median	0.017	0.03	-0.007	-0.018
	[0.830]	[0.662]	[0.858]	[0.731]
p-val, diff	[0.281]	[0.828]	[0.941]	[0.968]
Factories	75	74	80	72
Observations	368	370	400	360
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
Panel B: Post-treatment phase				
Below median	0.005	0.029	-0.009	-0.004
	[0.957]	[0.712]	[0.850]	[0.929]
Above Median	-0.027	-0.038	0.023	-0.015
	[0.636]	[0.504]	[0.671]	[0.808]
p-val, diff	[0.758]	[0.5]	[0.624]	[0.877]
Factories	75	74	80	72
Observations	218	222	240	216
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
Panel C: Pooled				
Below median	0.101	0.042	-0.010	-0.011
	[0.166]	[0.499]	[0.801]	[0.759]
Above Median	0.001	0.004	0.004	-0.017
	[0.987]	[0.939]	[0.922]	[0.727]
p-val, diff	[0.279]	[0.661]	[0.784]	[0.928]
Observations	586	592	640	576
Factories	75	74	80	72
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

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. *p<0.1; **p<0.05; ***p<0.01.

Online Appendices

C: Online Figures and Tables

Figure C1: Joint distribution of compliance and SC effectiveness 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 SC effectiveness under the null hypothesis of no treatment effects. The actual parameter estimates are indicated in red.

Table C1: Treatment effects: Output and Working Hours

	(1)	(2)
	Log(Output)	Average Weekly Working Hours
Treatment effect	0.026	-0.550
	[0.788]	[0.368]
Factories	75	79
Observations	374	395
Stratification variables	Y	Y
Control, baseline dep. var.	Y	Y

Notes: This table reports OLS estimates of treatment effects on physical output and average weekly working hours. 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 C2: Local Average Treatment Effects: Primary Outcome Index Variables

	Control mean	Local Average Treatment Effect	
	(1)	(2)	(3)
SC Compliance	0.046	0.226	0.221
		(0.070)***	(0.059)***
SC Effectiveness	0.103	0.153	0.154
		(0.071)**	(0.069)**
Worker job satisfaction & mental well-being	-0.007	-0.174	-0.161
		(0.080)**	(0.078)**
Observations		80	80
Stratification variables		Y	Y
Control, base. dep. var.		N	Y

Notes: This table reports two stage least squares (2SLS) estimates of treatment effects on primary outcome index variables. 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 local average treatment effect (L.A.T.E.) for factories that receive the treatment, using the randomized assignment to treatment to instrument for receiving the treatment before the second data collection visit. Column (3) reports the L.A.T.E. for factories that receive the treatment, using the randomized assignment to treatment to instrument for receiving the treatment before the second data collection visit and including a control for the baseline value of the dependent variable. Robust standard errors are reported in parentheses. Index variables constructed using Anderson (2008) variance-covariance weighted index. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table C3: Lee (2009) bounds for primary outcome index variables

	Lower bound	Upper bound
SC compliance index	0.206	0.218
	(0.076)***	(0.069)***
SC effectiveness index	0.140	0.141
	(0.091)	(0.072)**
Job satisfaction & mental well-being index	-0.173	-0.158
	(0.080)**	(0.083)*

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 C4: Baseline balance tests, sub-index components

	(1) Control mean	(2) T-C diff	(3) RI *p*-value	(4) N
Panel A: Compliance				
Formation sub-index	0.000	-0.171	0.298	84
Operations sub-index	0.000	0.035	0.798	84
Responsibilities sub-index	0.000	-0.096	0.360	84
Panel B: SC Effectiveness				
CAP completion sub-variable	0.017	0.092	0.671	84
Worker awareness sub-index	0.000	-0.547	0.028**	84
Worker knowledge sub-index	0.000	-0.082	0.674	84
Senior manager awareness sub-variable	0.000	0.348	0.134	84
Panel C: Worker job satisfaction and mental well-being				
Job satisfaction sub-index	0.000	-0.198	0.218	84
Mental well-being sub-index	0.000	-0.241	0.220	84
Turnover sub-variable	0.000	-0.002	0.991	84
Absenteeism sub-variable	0.000	0.148	0.466	80

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 C5: Baseline balance tests, secondary outcome variables

	(1) Control mean	(2) T-C diff	(3) RI *p*-value	(4) N
Panel A: Full sample				
Perceived SC compliance & effectiveness index	0.000	-0.170	0.206	84
Perceived worker-manager relations index	0.000	-0.187	0.137	84
Worker empowerment index	0.000	-0.192	0.137	84
Worker organization awareness index	0.000	-0.165	0.315	84
Panel B: Dropping outlier on worker outcomes				
Perceived SC compliance & effectiveness index	0.000	-0.126	0.320	83
Perceived worker-manager relations index	0.020	-0.147	0.232	83
Worker empowerment index	0.022	-0.117	0.268	83
Worker organization awareness index	-0.025	-0.121	0.460	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 C6: 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 C7: 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-product type. In Panel A columns (3) and (4), the sample is trimmed at the 1st and 99th percentile of 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 C8: Baseline balance tests within subgroups for heterogeneity analysis, primary outcome variables

	(1) Control mean	(2) T-C diff	(3) RI *p*-value	(4) N
Panel A: Managerial Practices				
Below median subgroup:				
SC Compliance	0.024	0.113	0.260	40
SC Effectiveness	0.099	-0.067	0.669	40
Job Satisfaction & Mental Well-being	0.067	-0.205	0.286	40
Above median subgroup:				
SC Compliance	-0.032	-0.114	0.284	40
SC Effectiveness	-0.067	0.013	0.932	40
Job Satisfaction & Mental Well-being	-0.046	-0.037	0.733	40
Panel B: Factory Size				
Below median subgroup:				
SC Compliance	-0.024	0.059	0.591	40
SC Effectiveness	0.078	0.043	0.807	40
Job Satisfaction & Mental Well-being	0.017	-0.048	0.760	40
Above median subgroup:				
SC Compliance	0.006	-0.079	0.497	40
SC Effectiveness	-0.051	-0.097	0.545	40
Job Satisfaction & Mental Well-being	-0.007	-0.115	0.499	40
Panel C: SC Compliance				
Below median subgroup:				
SC Compliance	-0.224	-0.061	0.401	40
SC Effectiveness	-0.029	-0.059	0.707	40
Job Satisfaction & Mental Well-being	-0.064	0.023	0.851	40
Above median subgroup:				
SC Compliance	0.221	0.027	0.557	40
SC Effectiveness	0.042	0.162	0.315	40
Job Satisfaction & Mental Well-being	0.075	-0.177	0.316	40
Panel D: Location in EPZ				
EPZ subgroup:				
SC Compliance	-0.190	0.385	0.199	14
SC Effectiveness	-0.051	0.220	0.579	14
Job Satisfaction & Mental Well-being	-0.107	0.503	0.126	14
Non-EPZ subgroup:				
SC Compliance	0.031	-0.013	0.870	66
SC Effectiveness	0.017	-0.066	0.580	66
Job Satisfaction & Mental Well-being	0.026	-0.185	0.114	66
Panel E: HR Managerial Practices				
Below median subgroup:				
SC Compliance	-0.051	0.080	0.469	40
SC Effectiveness	-0.181	0.045	0.815	40
Job Satisfaction & Mental Well-being	-0.013	-0.264	0.150	40
Above median subgroup:				
SC Compliance	0.031	-0.095	0.379	40
SC Effectiveness	0.167	-0.055	0.697	40
Job Satisfaction & Mental Well-being	0.018	0.032	0.804	40

Notes: This table reports OLS estimates of baseline differences between control and treatment groups within each pre-specified subgroup for treatment effect heterogeneity analysis. For the first three dimensions of heterogeneity, compliance, size, and managerial practices, I partition the sample into above/below median subgroups using the baseline value of the variable. For the final dimension of heterogeneity, 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, 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 on the treatment indicator and stratification variables within that subgroup. In column (3), I report the randomization inference (RI) *p*-value for the coefficient reported in column (2) based on 5000 draws. In column (4), I report the number of observations in that subgroup. *p<0.1; **p<0.05; ***p<0.01.

Table C9: Other heterogeneous treatment effects: Primary Outcome Index Variables

	SC Compliance	SC Effectiveness	Job Satisfaction & Mental Well-being
	(1)	(2)	(3)
Panel A: Baseline Size			
Below median	0.230	0.068	-0.232
	[0.005]***	[0.433]	[0.089]*
Above Median	0.159	0.189	-0.061
	[0.088]*	[0.125]	[0.527]
p-val, diff	[0.570]	[0.422]	[0.314]
Panel B: Baseline SC Compliance			
Below median	0.237	0.199	-0.296
	[0.029]**	[0.128]	[0.025]**
Above Median	0.169	0.074	0.001
	[0.014]**	[0.325]	[0.991]
p-val, diff	[0.607]	[0.405]	[0.075]*
Panel C: Baseline HR Management Practices			
Below median	0.132	0.093	-0.172
	[0.127]	[0.389]	[0.186]
Above Median	0.281	0.189	-0.113
	[0.003]***	[0.091]*	[0.280]
p-val, diff	[0.246]	[0.535]	[0.729]
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 C10: Other heterogeneous treatment effects: Primary Outcome Index Variables (pooled 4-5- and 9-10 month rounds)

	SC Compliance	SC Effectiveness	Job Satisfaction & Mental Well-being
	(1)	(2)	(3)
Panel A: Baseline Size			
Below median	0.179	0.071	-0.085
	[0.007]***	[0.283]	[0.360]
Above Median	0.193	0.105	0.059
	[0.011]**	[0.142]	[0.500]
p-val, diff	[0.858]	[0.722]	[0.264]
Panel B: Baseline SC Compliance			
Below median	0.231	0.098	-0.049
	[0.006]***	[0.194]	[0.632]
Above Median	0.180	0.099	0.009
	[0.003]***	[0.139]	[0.914]
p-val, diff	[0.639]	[0.988]	[0.646]
Panel C: Baseline HR Management Practices			
Below median	0.128	0.024	-0.093
	[0.076]*	[0.728]	[0.335]
Above Median	0.284	0.180	0.059
	[0.000]***	[0.017]**	[0.481]
p-val, diff	[0.158]	[0.115]	[0.228]
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 using the pooled 4-5- and 9-10-month observations. 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 C11: Baseline balance tests within subgroups for heterogeneity analysis, business primary outcome variables

	(1) Control mean	(2) T-C diff	(3) RI *p*-value	(4) N
Panel A: Log(Labor productivity)†*, trimmed sample*				
Below median Compliance	1.828	-0.278	0.369	35
Above median Compliance	1.248	0.060	0.762	40
Below median Size	2.313	0.090	0.705	39
Above median Size	0.867	-0.012	0.943	36
Below median Mgmt	2.069	-0.208	0.381	40
Above median Mgmt	1.105	-0.112	0.677	35
Panel B: Log(Employment)				
Below median Compliance	6.545	0.095	0.773	40
Above median Compliance	6.761	-0.726	0.060*	40
Below median Size	5.802	-0.316	0.195	40
Above median Size	7.376	-0.173	0.356	40
Below median Mgmt	6.297	0.060	0.873	40
Above median Mgmt	6.923	-0.521	0.165	40
Panel C: Log(Wages)				
Below median Compliance	15.559	0.294	0.445	35
Above median Compliance	16.066	-0.706	0.066*	37
Below median Size	14.923	-0.136	0.606	36
Above median Size	16.575	-0.094	0.681	36
Below median Mgmt	15.625	0.007	0.984	38
Above median Mgmt	16.004	-0.382	0.378	34

Note: This table reports OLS estimates of baseline differences between control and treatment groups within each pre-specified subgroup for treatment effect heterogeneity analysis. For the first three dimensions of heterogeneity, compliance, size, and managerial practices, I partition the sample into above/below median subgroups using the baseline value of the variable. For each outcome, within subgroup, 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 on the treatment indicator and stratification variables within that subgroup. In column (3), I report the randomization inference (RI) *p*-value for the coefficient reported in column (2) based on 5000 draws. In column (4), I report the number of observations in that subgroup. † The regression also includes product-type fixed effects. The trimmed sample drops factories in the 1st and 99th percentiles of labor productivity. *p<0.1; **p<0.05; ***p<0.01.

Table C12: Treatment effects after end of intensive enforcement: Business competitiveness outcomes

	(1)	(2)	(3)
Panel A	Log(Labor productivity)		
Treatment effect	0.115	0.082	0.036
	[0.652]	[0.733]	[0.837]
			{1.000}
Factories	75	75	74
Observations	375	370	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(Employment)	Log(Gross wages)	
Treatment effect	0.003	-0.009	
	[0.932]	[0.811]	
	{1.000}	{1.000}	
Factories	80	72	
Observations	240	216	
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-product type. In column (2), the sample is trimmed at the 1st and 99th percentile of 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.

D: Main Results Dropping Outlier on Worker Variables

Table D1: Baseline balance tests, dropping outlier on worker outcomes

	(1) Control mean	(2) T-C diff	(3) RI *p*-value	(4) Number of factories
Primary outcome variables				
Compliance index	0.000	-0.086	0.378	83
Effectiveness index	0.002	-0.032	0.774	83
Job satisfaction & well-being index	-0.001	-0.078	0.377	83
Number employees	1190	-242	0.419	79
Gross wages (log)	15.82	-0.263	0.303	71
Labor productivity (log)	2.639	0.010	0.962	74
Labor productivity (log), trimmed sample	2.588	0.021	0.926	74
Panel B: Compliance sub-components				
Formation sub-index	0.000	-0.189	0.277	83
Operations sub-index	0.000	0.028	0.841	83
Responsibilities sub-index	0.000	-0.113	0.296	83
Panel C: SC Effectiveness sub-components				
CAP completion sub-variable	0.017	0.095	0.656	83
Worker awareness sub-index	0.000	-0.424	0.058*	83
Worker knowledge sub-index	0.000	-0.089	0.651	83
Senior manager awareness sub-variable	0.000	0.321	0.177	83
Panel D: Worker job satisfaction and mental well-being sub-components				
Job satisfaction sub-index	0.000	-0.154	0.326	83
Mental well-being sub-index	0.000	-0.122	0.411	83
Turnover sub-variable	0.000	-0.003	0.990	83
Absenteeism sub-variable	0.000	0.141	0.487	79

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 D2: Treatment effects: Compliance with SC regulation, dropping outlier on worker outcomes

	Control mean	ITT Effect	
	(1)	(2)	(3)
Panel A: Primary outcome			
SC Compliance Index	0.046	0.203	0.205
		[0.004]***	[0.001]***
		{0.012}**	{0.002}***
Panel B: Sub-indexes			
Formation sub-index	0.118	0.055	0.079
		[0.541]	[0.338]
		⟨0.524⟩	⟨0.511⟩
Operations sub-index	0.184	0.119	0.057
		[0.229]	[0.512]
		⟨0.298⟩	⟨0.519⟩
Responsibilities sub-index	-0.149	0.409	0.437
		[0.002]***	[0.001]***
		⟨0.007⟩***	⟨0.004⟩***
Observations		79	79
Stratification variables		Y	Y
Control, base. dep. var.		N	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 for multiple hypothesis testing using the method of List, Shaikh and Xu (2016) are reported in curly brackets. In Panel B, FDR-sharpened p-values are reported in pointy brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table D3: Treatment effects: Indicators of SC effectiveness, dropping outlier on worker outcomes

	Control mean	ITT Effect	
	(1)	(2)	(3)
Panel A: Primary outcome			
SC Effectiveness	0.103	0.140	0.139
		[0.058]*	[0.053]*
		{0.058}*	{0.108}
Panel B: Sub-indexes and sub-variables			
Factory safety spotcheck index	-0.000	0.218	
		[0.026]**	
		⟨0.150⟩	
CAP completion sub-variable	0.345	0.111	0.025
		[0.593]	[0.801]
		⟨1.000⟩	⟨1.000⟩
Worker SC awareness sub-index	0.049	0.076	0.202
		[0.629]	[0.203]
		⟨1.000⟩	⟨0.684⟩
Worker safety knowledge sub-index	0.378	-0.056	-0.053
		[0.489]	[0.512]
		⟨1.000⟩	⟨1.000⟩
Senior manager awareness sub-variable	0.086	0.077	0.053
		[0.669]	[0.765]
		⟨1.000⟩	⟨1.000⟩
Observations		79	79
Stratification variables		Y	Y
Control, base. dep. var.		N	Y

Notes: This table reports OLS estimates of treatment effects on the index of SC effectiveness 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 for multiple hypothesis testing using the method of List, Shaikh and Xu (2016) are reported in curly brackets. In Panel B, FDR-sharpened p-values are reported in pointy brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table D4: Treatment effects: Workers' job satisfaction and mental well-being, dropping outlier on worker outcomes

	Control mean	ITT Effect	
	(1)	(2)	(3)
Panel A: Primary outcome			
Worker job satisfaction & mental well-being (well-being index)	-0.007	-0.153	-0.147
		[0.052]*	[0.072]*
		{0.103}	{0.062}*
Panel B: Sub-indexes and sub-variables			
Job satisfaction sub-index	-0.100	-0.386	-0.374
		[0.014]**	[0.021]**
		⟨0.060⟩*	⟨0.092⟩*
Mental well-being sub-index	0.003	-0.059	-0.052
		[0.588]	[0.722]
		⟨1.000⟩	⟨0.804⟩
Turnover sub-variable	0.115	0.073	-0.011
		[0.574]	[0.891]
		⟨1.000⟩	⟨0.804⟩
Absenteeism sub-variable	0.088	0.025	-0.084
		[0.842]	[0.188]
		⟨1.000⟩	⟨0.393⟩
Observations		79	79
Stratification variables		Y	Y
Control, base. dep. var.		N	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 for multiple hypothesis testing using the method of List, Shaikh and Xu (2016) are reported in curly brackets. In Panel B, FDR-sharpened *p*-values are reported in pointy brackets. Index variables constructed using Anderson (2008) variance-covariance weighted index. *p<0.1; **p<0.05; ***p<0.01.

Table D5: Treatment effects: Business competitiveness outcomes, dropping outlier on worker outcomes

	Log(Labor productivity)			Log(Employment)	Log(Gross wages)
	(1)	(2)	(3)	(4)	(5)
Treatment effect	0.118	0.086	0.041	0.003	-0.027
	[0.149]	[0.201]	[0.354]	[0.848]	[0.283]
Factories	74	74	73	79	71
Observations	370	365	365	395	355
Stratification variables	Y	Y	Y	Y	Y
Control, baseline dep. var.	Y	Y	Y	Y	Y
Product type FE	Y	Y	Y	N	N
Trimmed sample	N	Y	N	N	N

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. The regression sample changes across columns due to differential data availability. In column (1), the regression includes five post-treatment observations per factory-product type. In column (2), the sample is trimmed at the 1st and 99th percentile of labor productivity observations. In column (3), a factory in the control that partially shut down during the study is dropped. In columns (4) and (5), 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.