

## Do Management Interventions Last? Evidence from India<sup>†</sup>

By NICHOLAS BLOOM, APRAJIT MAHAJAN, DAVID MCKENZIE,  
AND JOHN ROBERTS\*

*We revisited Indian weaving firms nine years after a randomized experiment that changed their management practices. While about half of the practices adopted in the original experimental plants had been dropped, there was still a large and significant gap in practices between the treatment and control plants, suggesting lasting impacts of effective management interventions. Few practices had spread across the firms in the study, but many had spread within firms. Managerial turnover and the lack of director time were two of the most cited reasons for the drop in management practices, highlighting the importance of key employees. (JEL D22, D24, L67, L84, M11, O14)*

After an early recognition of management as a driver of differences in firm performance (e.g., Walker 1887, Marshall 1887), economists are again paying increasing attention to the role of management in firm and economy-wide performance (Roberts 2018). Whereas the size and profitability of the management consulting industry are often cited as a revealed preference measure of the importance of management, recent academic work has also established a credible causal link between changes in management practices and productivity in medium and large firms (Bloom et al. 2013, Bruhn et al. 2018). The longer-term persistence of management improvements caused by consulting interventions, however, remains an open question.<sup>1</sup> The received wisdom at a leading global management consulting

\*Bloom: Stanford University, 579 Serra Mall, Stanford, CA 94305 (email: nbloom@stanford.edu); Mahajan: University of California, Berkeley, 289 University Hall, Berkeley, CA 94720 (email: aprajit@berkeley.edu); McKenzie: Development Research Group, the World Bank, MSN MC3-307, 1818 H Street NW, Washington, DC 20433 (email: dmckenzie@worldbank.org); Roberts: Stanford University, Faculty Building W311, 655 Knight Way, Stanford, CA 94305. Benjamin Olken was coeditor for this article. Financial support was provided by SEED at the Graduate School of Business at Stanford, by the Stanford Center for Poverty and Development, and by the World Bank under the Strategic Research Program (SRP). This research would not have been possible without the consulting team of Saurabh Bhatnagar, Shaleen Chavda, Rahul Dsouza, Sumit Kumar, and Ashutosh Tyagi. We thank three anonymous referees; our formal discussant, Rebecca Henderson; Larry Katz; and seminar participants at Duke, IGC, Maryland, NBER, Stanford, and the World Bank for comments. AEA RCT Registry identifying number AEARCTR-0002808.

<sup>†</sup>Go to <https://doi.org/10.1257/app.20180369> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

<sup>1</sup>To our knowledge, Giorcelli (2019), which uses observational data to examine the effect of management training sponsored by the Marshall Plan on long-term outcomes, is the only other work that examines persistence in a causal framework.

firm when two of the authors were employed there was that such innovations lasted approximately three years.<sup>2</sup>

Competing views of management offer differing predictions about the persistence of consulting-induced improvements in management practices. One view, best exemplified by the “Toyota way” (Liker 2004), sees management improvements as launching a continuous cycle of improvement, as systems put in place for measuring, monitoring, and improving operations and quality enable constant improvement. A related idea is that management practices are complementary with one another, so that the costs of adding new practices fall as others are put in place. For example, in our context of cotton weaving, scientific management of inventory levels will be possible only once the firm has put in place systems to record all yarn transactions and to regularly monitor stock levels.

A countervailing view argues that maintaining good management is difficult, with many of the companies extolled in business books as paragons of good management subsequently failing (*Economist* 2009, Kiechel 2012). This may be even harder when changes are introduced externally, with the Boston Consulting Group (BCG) reporting that two-thirds of transformation initiatives ultimately fail (Sirkin et al. 2005). This finding presumably refers to high-level strategic and organizational change efforts in large firms that would use BCG. But both Karlan et al. (2015) and Higuchi et al. (2019) find that light consulting engagements in smaller firms than the ones we studied led to firms’ gradually discarding practices over the subsequent three years. One reason may be that these practices were inappropriate and were abandoned as firms learned that they were not suitable in their setting.

This paper examines the persistence of management practices adopted after an extensive, consultant-supported intervention that we undertook in a set of multiplant Indian textile weaving firms from 2008 to 2010 (for a more detailed description, see Bloom et al. 2013). The intervention took the form of a randomized controlled trial. Firms were randomly allocated into treatment and control groups, and the intervention was done at the plant level within each firm. Both treatment and control plants, which were never in the same firm, were given recommendations for improving management practices in several areas, and the treatment plants received additional consulting help in implementing the recommendations. The intervention led to a substantial uptake of the recommended practices in the treatment plants and a modest one in the control plants, with corresponding improvements in various measures of performance.

We stopped observing the firms in 2011, but we wondered—as did many in our audiences—about whether these changes would last. As a result, we returned to the study firms in 2017 with the same consulting team and collected data on management practices and basic firm performance. We found that both treatment and control experimental plants had in fact dropped some practices, though fewer than we and the consultants had forecast. Since the control plants also dropped practices, the treatment effect on practices is constant over time, at 20 percentage points. Meanwhile, the *plants* in the treatment *firms* that had not been part of the experiment

<sup>2</sup>This is consistent with a case study described in McNair (n.d.), which recounts quality training for workers as having a half-life of two to three years.

(treatment firms typically had multiple plants) had adopted many of the recommendations, so their packages of current practices were very close to those of the treatment plants.

We were also able to collect information on the reasons for the dropping of management practices. We found that practices were more likely to be dropped when the plant manager changed, when the directors (the CEO and CFO) were busier, and when the practice was one that is not commonly used in many other firms. The first two reasons highlight the importance of key employees within the firm for driving management practices,<sup>3</sup> while the latter suggests it is easier to get more commonplace practices to stick.

Although budgetary constraints and the reluctance of firm owners to reveal financial details rendered us unable to measure long-term impacts on firm profits or overall productivity, we were able to track changes in looms per worker, a simple and commonly used proxy for labor productivity in the industry, and use this to impute worker productivity. Despite their dropping some practices, we found that treated firms show lasting improvements in worker productivity, which is 35 percent higher than in the control group after 8 years. We also found that treated firms are more likely to be exporting, have upgraded the quality of their looms, and are using more consulting services of their own accord, and that they have supplemented the operational management practices introduced by the consultants from our study with better marketing practices.

This paper is related to several literatures, including the literature on the drivers of firm and national productivity (see, e.g., Syverson 2011), the literature on management randomized control trials (see, e.g., Anderson et al. 2016, McKenzie and Woodruff 2014), and the large literature on the importance of management for firm performance (e.g., Osterman 1994, Huselid 1995, Ichniowski et al. 1997, Capelli and Neumark 2001, Braguinsky et al. 2015). Section I of the paper discusses the original consulting experiment, Section II describes the follow-up, and Section III offers concluding remarks.

## I. The 2008–2010 Consulting Experiment

### A. *The Experimental Design*

Our original experiment measured the impact of improving management practices in a set of large textile firms near Mumbai in 2008. The experiment involved 28 plants across 17 firms in the woven cotton fabric industry. These firms had been in operation for 20 years on average and were family owned and managed. They produced fabric for the domestic market (although a few also exported). Table 1 reports summary statistics for the textile manufacturing parts of these firms (a few of the firms had other businesses in textile processing, retail, and real estate). On average, the study firms had about 270 employees, assets of \$8.5 million, and annual sales of \$7.5 million. Compared with US manufacturing firms, these firms would be in

<sup>3</sup>This links to the literature on management and CEOs—for example, Bertrand and Schoar (2003), Bennesden et al. (2007), Lazear et al. (2016), and Bandiera et al. (2018).

TABLE 1—THE FIELD EXPERIMENT SAMPLE PREINTERVENTION (2008)

	All				Treatment	Control	Diff
	Mean	Median	Min	Max	Mean	Mean	<i>p</i> -value
<b>Sample sizes</b>							
Number of plants	28	NA	NA	NA	19	9	NA
Number of experimental plants	20	NA	NA	NA	14	6	NA
Number of firms	17	NA	NA	NA	11	6	NA
Plants per firm	1.65	2	1	4	1.73	1.5	0.393
<b>Firm/plant sizes</b>							
Employees per firm	273	250	70	500	291	236	0.454
Employees, experimental plants	134	132	60	250	144	114	0.161
Hierarchical levels	4.4	4	3	7	4.4	4.4	0.935
Annual sales (\$M) per firm	7.45	6	1.4	15.6	7.06	8.37	0.598
Current assets (\$M) per firm	8.50	5.21	1.89	29.33	8.83	7.96	0.837
Daily meters, experimental plants	5,560	5,130	2,260	13,000	5,757	5,091	0.602
<b>Management and plant ages</b>							
BVR management score	2.60	2.61	1.89	3.28	2.50	2.75	0.203
Management adoption rates	0.262	0.257	0.079	0.553	0.255	0.288	0.575
Age, experimental plant (years)	19.4	16.5	2	46	20.5	16.8	0.662

*Notes:* Data are provided at the plant and/or firm level depending on availability. Number of plants is the total number of textile plants per firm including the nonexperimental plants. Number of experimental plants is the total number of treatment and control plants. Number of firms is the number of treatment and control firms. Plants per firm reports the total number of other textile plants per firm. Several of these firms have other businesses—for example, retail units and real estate arms—that are not included in any of the figures here. Employees per firm reports the number of employees across all the textile production plants, the corporate headquarters, and the sales office. Employees, experimental plants, reports the number of employees in the experimental plants. Hierarchical levels displays the number of reporting levels in the experimental plants—for example, a firm with workers reporting to foreman, foreman to operations manager, operations manager to general manager, and general manager to managing director would have 4 hierarchical levels. Annual sales (\$M) and current assets (\$M) are both in millions of 2009 US dollar values, exchanged at 50 rupees = 1 US dollar. Daily meters, experimental plants, reports the daily meters of fabric woven in the experiment plants. Note that about 3.5 meters is required for a full suit with jacket and trousers, so the mean plant produces enough for about 1,600 suits daily. BVR management score is the Bloom and Van Reenen (2007) management score for the experimental plants. Management adoption rates are the adoption rates of the management practices listed in online Appendix Table A1 in the experimental plants. Age, experimental plant (years), reports the age of the plant for the experimental plants. NA is not applicable.

the top 1 percent by employment and the top 4 percent by sales, and compared with Indian manufacturing firms, they are in the top 1 percent by both measures (Hsieh and Klenow 2010). Hence, these are large manufacturing firms.<sup>4</sup>

These firms are complex organizations, with a median of two plants per firm (in addition to a head office in Mumbai) and four reporting levels from the shop floor to the managing director. The managing director was the largest shareholder in each firm, and all directors were his close relatives. Two firms were publicly listed on the Mumbai Stock Exchange, although more than 50 percent of the equity in each of these was held by the managing family.

The field experiment aimed at improving management practices in the treatment plants, and we measured the impact of doing so on firm performance. We contracted with a leading international management consultancy firm to work with the plants as the easiest way to change plant-level management practices rapidly. The full-time team of (up to) six consultants had been educated at leading Indian business and

<sup>4</sup>Note that most international agencies define large firms as those with more than 250 employees.

engineering schools, and most of them had prior experience working with US and European multinationals.

The intervention ran from August 2008 until August 2010, with data collection continuing until November 2011. The intervention focused on a set of 38 management practices that are standard in American, European, and Japanese manufacturing firms and that can be grouped into five broad areas: factory operations, quality control, inventory control, human resources management, and sales and orders management (for details, see online Appendix Table A1). Each practice was measured as a binary indicator of the adoption (1) or nonadoption (0) of the practice. A general pattern at baseline was that plants recorded a variety of information (often on paper sheets) but had no systems in place to monitor these records or use them in decisions. For example, 93 percent of the treatment plants recorded quality defects before the intervention, but only 29 percent monitored them daily or by the particular sort of defect, and none of them had any standardized system to analyze and act upon these data.

The consulting intervention had three phases. The first phase, called the *diagnostic* phase, took one month and was given to all treatment and control experimental plants. It involved evaluating the current management practices of each plant and constructing a performance database. At the end of the diagnostic phase the consulting firm provided each plant with a detailed analysis of its current management practices and performance and, crucially, recommendations for change.

The second phase was a four-month *implementation* phase given only to the treatment experimental plants. In this phase, the consulting firm followed up on the diagnostic report to help introduce as many of the 38 management practices as the plants could be persuaded to adopt. The consultant assigned to each plant worked with the plant management to put the procedures into place, fine-tune them, and stabilize them so that employees could readily carry them out. It is on this dimension that treatment and control plants differed.

The third phase was a *measurement* phase, which lasted until November 2011. This involved collection of performance and management data from all treatment and control plants. In return for this continuing data, the consultants provided light consulting advice to the treatment and control plants (primarily to keep them involved).

### B. *The Initial Experimental Results—Management Practices*

The intervention led to increases in the adoption of the 38 management practices in the treatment plants by an average of 37.8 percentage points by August 2010 (approximately one year after the start of the intervention). This adoption rate dropped by only 3 percentage points in the subsequent year, showing considerable persistence in practices after the consultants had exited the firms. Not all practices were adopted equally, with firms (unsurprisingly) adopting the practices that were the easiest to implement and/or had the largest perceived short-run payoffs, e.g., the daily quality, inventory, and efficiency review meetings. This adoption also occurred gradually, in large part reflecting the time taken for the consulting firm to gain the confidence of the firms' directors. Initially many directors were skeptical about the

suggested management changes, and the intervention often started by piloting the easiest changes around quality and inventory in one part of the factory. Once these started to generate improvements, these changes were rolled out, and the firms then began introducing the more complex improvements around operations and human resources.

In contrast, the control plants, which were given only the one-month diagnostic and corresponding recommendations, increased their adoption of the management practices, but by only 12 percentage points on average. This is substantially less than the increase in adoption among the treated plants, indicating that the four months of the implementation phase were important in changing management practices. Table 2, column 2, reflects this and shows a statistically significant 25 percentage point treatment effect on management practices in 2011. We note that the change for the control firms is still an increase relative to the rest of the industry cluster around Mumbai (which had more than 100 nonproject plants), which did not change their management practices on average between 2008 and 2011.

Finally, since these are multiplant firms and the consulting firm worked at the plant level, the treatment and control firms also had plants that were not part of the intervention, which we label “nonexperimental plants.” For example, if a treatment firm has plants A, B, and C and the diagnostic and implementation intervention was performed on plant A, this would be a “treatment experimental plant,” while plants B and C would be “treatment nonexperimental plants.” Likewise, if a control firm had plants D, E, and F and the diagnostic intervention was performed only on plant D, then D would be a “control experimental plant,” while E and F would be “control nonexperimental plants.” Online Appendix Table A2 reports the breakdown of the plant count into these four groups.

Although the consulting firm did not provide consulting services to the nonexperimental plants, it was still able to collect bimonthly management data and some basic data for these plants. The nonexperimental plants in the treatment firms saw a substantial increase in the adoption of management practices. In these five plants the adoption rates increased by 17.5 percentage points by August 2010, without any drop in the second year. This increase occurred because the executives of the treatment firms copied the new practices from their experimental plants over to their other (nonexperimental) plants. Interestingly, this increase in adoption rates is similar to the experimental control firms’ 12 percentage point increase, suggesting that the copying of new practices across plants within firms can be as least as effective at improving management practices as short (one-month) bursts of external consulting advice without implementation support.

### *C. The Initial Experimental Results—Firm Performance*

Experimental treatment plants experienced a significant increase in output of 9.4 percent relative to the experimental control plants, which came about both by decreasing quality defects (so that less output was scrapped) and by undertaking routine maintenance of the looms, collecting and monitoring breakdown data, and keeping the factory clean, which reduced machine downtime. Total factor productivity (TFP) increased by 16.6 percent as a result of both the increase in output and

TABLE 2—SHORT- AND LONG-RUN IMPACT ON MANAGEMENT PRACTICES

Dependent variable: proportion of management practices implemented	(1)	(2)
Treatment $\times$ year = 2011	0.206 (0.042) [0.003]	0.249 (0.038) [0.001]
Treatment $\times$ year = 2017	0.197 (0.062) [0.007]	0.218 (0.057) [0.004]
Year = 2017	-0.122 (0.016) [0.732]	-0.122 (0.016) [0.694]
Baseline 2008 management score	0.668 (0.219) [0.022]	0.878 (0.176) [0.006]
<i>p</i> -value of test of equality of treatment in 2011 and 2017	0.802	0.457
Sample size	37	34

*Notes:* Robust standard errors are in parentheses, and permutation test *p*-values are in brackets. Both are clustered at the firm level. Permutation tests report the *p*-value for testing the null hypothesis that the treatment had no effect by constructing the permutation distribution of the estimator using 4,000 possible permutations of firm-level random assignment. The second column limits the sample from column 1 to plants that were present in both years with no missing management scores.

a reduction in inputs due to reduced inventory and reduced labor inputs for mending defective fabric. These improvements were estimated to have increased profits per plant by about \$325,000 per year. We estimate that this represented, on average, a 130 percent one-year return on the market cost of the consulting services.

## II. The 2017 Follow-up

### A. The Follow-up Process

In January 2017, working with the same consulting firm, we recontacted the 17 textile firms from the original study. Fortunately, all 17 firms agreed to work with the research team again on a follow-up study. This 100 percent uptake was aided by a combination of three factors: (i) the positive impact of the intervention in the first wave on the firms' management and performance; (ii) the stability of the firms, which had maintained the same address and contact details; and (iii) the engagement of the same three consulting company partners and project manager as in the 2008–2011 intervention.<sup>5</sup> One complication is that one single-plant treatment firm was in the midst of closing down after the owner's death. Without any close male relatives to continue the business, the owner's widow had decided to sell the business, which, given its location, meant the firm would go out of business and the site would be converted into residential housing.<sup>6</sup>

<sup>5</sup>These personal contacts are very important in our context. In fact, we delayed the start of this project to ensure we could staff the project with the same senior consulting team as in the 2008–2011 wave.

<sup>6</sup>The firm was over 30 years old, and due to the expansion of Mumbai, it was now located in a residential area, so the land was more valuable as housing than for production.

One weakness of this follow-up wave is that our budget allowed us only two months of the consultants' time, which was sufficient to collect management data for all production sites and a basic set of firm performance indicators (e.g., on employment and looms) but not to collect detailed weekly output data that would allow TFP estimation because that would have required extracting data on a firm-by-firm basis from log books and accounting software.<sup>7</sup> Firms were also more reluctant to share financial and performance data when sharing was not going to be directly accompanied by intensive consulting help. Consequently, our analysis is confined to management practices and basic performance indicators like employment or looms per employee, along with an imputed measure of labor productivity.

This follow-up data collection corresponds to an average period of nine years since the implementation phase of the consulting intervention started and seven years since it ended. It therefore enables us to examine the long-term persistence of these large changes in management practices.

### *B. Results on Management Practices*

In Figure 1 we plot the management scores over time after revisiting the plants in January 2017, evaluated on the same 38-management-practice scoring grid as in the prior experiment. We find substantial persistence of the management intervention, which we summarize below with four main results.

*Treatment Experimental Plants.*—First, the management scores in the treatment experimental plants fell from 0.60 at the end of the last wave to 0.46 eight years later. This drop of 0.14 points in the management score reverses 40 percent of the original 0.35 increase (noting these firms started pretreatment with an average management score of 0.25) over an eight-year period. This fall in the management practice score is equivalent to about an annual depreciation rate of 6 percent in the original increase in management practices.

*Control Experimental Plants.*—Second, these control plants also saw a drop in their management scores, falling by 0.08 points from 0.40 at the end of the last wave to 0.32. This is smaller in absolute terms compared with the fall in scores in the treatment plants, but the increase in management practices in the control plants was only 0.12 points (from an original score of 0.28), so that the drop in practice scores is 66 percent of the intervention gain, implying about a 13 percent depreciation rate of the original management increase.

<sup>7</sup>Consultants typically spent an entire day at the plant. They began with a set of structured discussions with one of the owners (typically the one in charge of plant operations). Subsequent discussions involved one to two managers and one to two supervisors per plant as well. Following the discussions, the consultants collected data from the plant manager (with the help of various supervisors). This process required the production of registers and worksheets to record and verify the numbers provided. The consultants also "shadowed" plant managers through the day, complementing written records with shop-floor inspections to double-check claims.

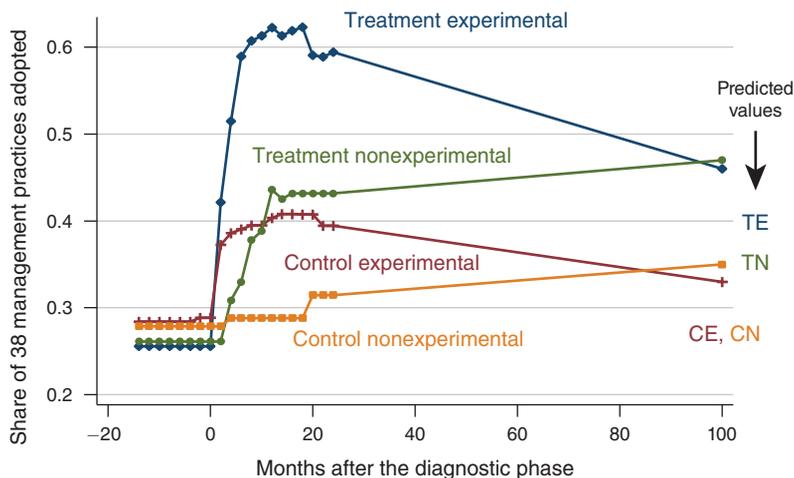


FIGURE 1. MANAGEMENT PRACTICES BY PLANT GROUP

Notes: Sample composed of the balanced panel of plants from 2008 to 2017 (11 treatment experimental, 6 treatment nonexperimental, 6 control experimental, and 2 control nonexperimental). The letters on the right are the average predicted values that the three-person Accenture team and four coauthors made before recontacting the firms for treatment experimental (TE), at 0.4; treatment nonexperimental (TN), at 0.36; control experimental (CE), at 0.29; and control nonexperimental (CN), at 0.29.

Together this indicates that, even eight years after the initial intervention, the treatment firms still had higher management practices. Table 2 reports the results from running the analysis of covariance specification for plants ( $i$ ) at time ( $t$ ):

$$(1) \text{Management}_{i,t} = a + b_1 \times \text{Treat}_i \times (\text{Year} = 2011) \\ + b_2 \times \text{Treat}_i \times (\text{Year} = 2017) + c \times \text{Management}_{i,2008} + e_{i,t}.$$

Indeed, we see that the long-run treatment effect in 2017 of 19.7 percentage points is similar in magnitude to the short-run effect in 2011 (20.6 percentage points), and we cannot reject equality of these treatment effects over time ( $p = 0.802$ ). These effects are individually statistically significant using both conventional (large-sample normality-based) inference and permutation procedures with exact finite sample size (the corresponding  $p$ -values are also reported in Table 2). Thus, the intervention generated persistent impacts on the treatment plants. Moreover, the greater percentage depreciation of the improvements in the control plants (66 percent) versus the treatment plants (40 percent) suggests that small improvements in management may be less stable than large improvements. One possible reason, which we discuss further below, is that bundles of management practices are complementary, so that adopting only parts of them may be less stable than adopting all of them. Of course, given the small sample sizes in this experiment, this could also reflect sampling noise—something that should be remembered when evaluating all our results from this experiment.

*Nonexperimental Plants.*—Third, the nonexperimental plants in the treatment firms experienced a slight improvement in their management practice adoption rates, from 0.43 in 2011 to 0.47 in 2017. Indeed, by 2017 their management scores were very similar overall to the treatment experimental plants (in fact slightly higher, although not significantly so). Similarly, in the control firms, the nonexperimental plants also converged with the experimental plants (again slightly but not significantly higher). This suggests (as we discuss further below) that the practice improvements in the experimental plants spilled over to the nonexperimental plants during the seven years after the intervention.

*Expectations on Durability of the Intervention.*—Finally, before we recontacted the firms in December 2016, each member of the consulting team from the original intervention and the academic team provided predictions for the management scores we expected to find on revisiting the firms in 2017.<sup>8</sup>

These expectations were informed by the contrasting views of management improvements noted in the introduction: under the “Toyota way” of continuous improvement, we would expect the management practices not only to persist but to continue to spread in treatment plants so that the gap with the control plants would widen, whereas under the “inappropriate technology” view, we would expect many practices to be dropped and the treatment group to converge back to the control group. The average values of the estimates of the seven team members are shown for the treatment experimental, treatment nonexperimental, control experimental, and control nonexperimental plants with the symbols TE, TN, CE, and CN, respectively, on the graph.<sup>9</sup> These predicted values are all below the actual outcomes, indicating that the project team expected steeper declines in management practices relative to what actually occurred, particularly for the nonexperimental plants. While some of the practices were dropped, the majority of the interventions remained in place eight years later, and the gap with the control group remained steady. The results therefore lie between these two extreme views of constant improvement and no long-run impact.

To delve further into the management changes, we also analyzed the 38 individual practices as highlighted in Figure 2, which plots the average score for the experimental plants in the treatment firms on each practice on the x-axis against the average scores for the nonexperimental plants (in the same firms) on the y-axis, for the years 2008 (preintervention), 2011 (postintervention), and 2017 (long-run follow-up). We observe that initially the experimental and nonexperimental plants in the treatment firms had similar practice scores, with a correlation of 0.91. After the intervention, the scores for the experimental plants increased considerably, leading to an eastward shift in the points and a drop in the correlation to 0.81 (panel B). Finally, in panel C, we see the experimental plants and nonexperimental plants again

<sup>8</sup>Other examples of getting experts to provide ex ante predictions of the results of an experiment can be found in Hirschleifer et al. (2016), Groh et al. (2016), and DellaVigna and Pope (2018).

<sup>9</sup>The predictions of the individual consultant and academic team members were made independently—Bloom estimated first, and then the other team members individually emailed him their predicted scores. The average predicted scores were not particularly different across the two groups (hence, we present them averaged together).

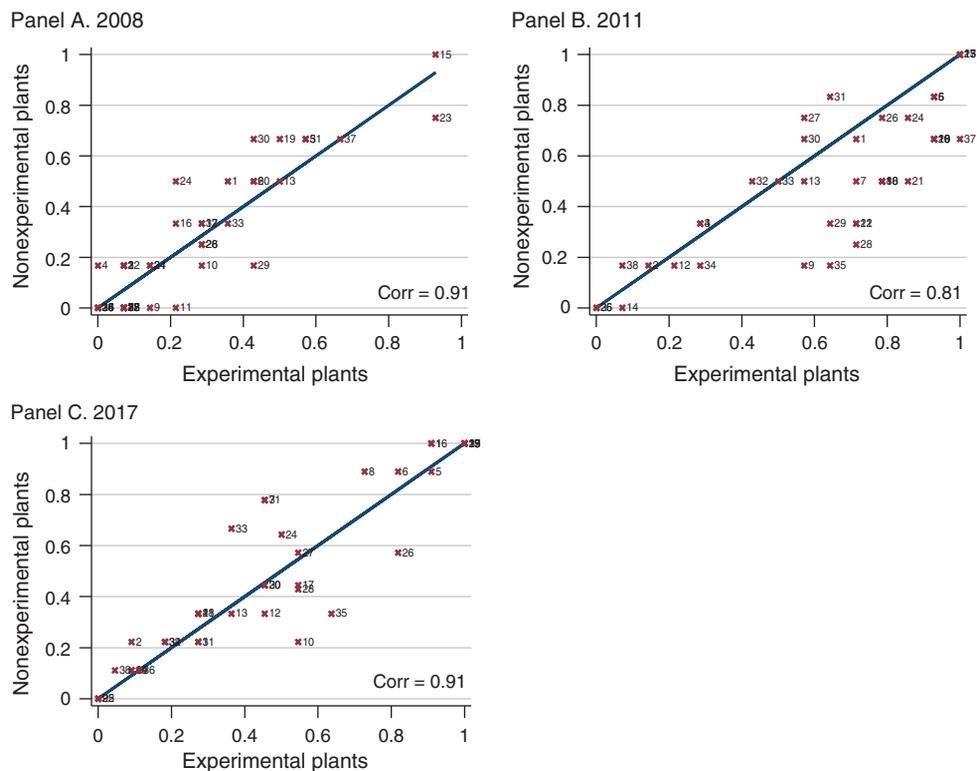


FIGURE 2. PRACTICES APPEAR TO SPREAD OUT FULLY IN TREATMENT FIRMS

Notes: The three graphs plot the average scores for each of the 38 questions for the 14 (11 in 2017) treatment experimental plants (on the x-axis) and the 6 treatment nonexperimental plants (on the y-axis) in 2008 (panel A), 2011 (panel B), and 2017 (panel C). The correlations between these scores for the 38 practices are reported as well on the graphs.

have very similar scores (correlation of 0.91), with a reversion of the scores toward the 45-degree line.

Figure 3 complements this by showing the long difference of management practices in the experimental and nonexperimental plants in the treatment firms between 2008 and 2017 (panel A) and between 2011 and 2017 (panel B). This shows, first, that between 2008 and 2017 both sets of plants adopted similar bundles of management practices. But, second, looking at 2011–2017 we see the timing of these practice adoptions was not the same. The experimental plants adopted most of these practices between 2008 and 2011, so that from 2011 to 2017 they mostly had negative practice changes. The nonexperimental plants, in contrast, were still heavily adopting a number of practices post-2011, so they show a balanced mix of drops and additions post-2011.

So, in summary, Figures 1 to 3 paint a picture of the treatment (and to a lesser extent the control) experimental plants adopting a slew of management practices during the initial intervention phase in 2008–2010, so that by 2011 they had substantially higher management scores. These scores subsequently subsided as some practices were dropped. The nonexperimental plants adopted fewer practices in

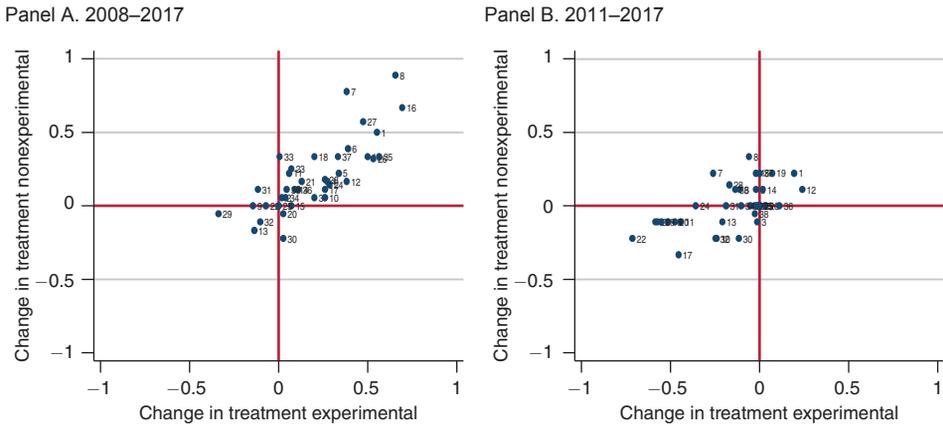


FIGURE 3. CHANGES IN EXPERIMENTAL AND NONEXPERIMENTAL PLANTS IN THE TREATMENT FIRMS BETWEEN 2008 AND 2017 AND BETWEEN 2011 AND 2017

*Note:* The figure plots the change in the share of practices of each of the 38 questions for the 14 (11 in 2011) treatment experimental plants (on the x-axis) and the 6 treatment nonexperimental plants (on the y-axis) between 2008 and 2017 (panel A) and between 2011 and 2017 (panel B).

2008–2010 but continued to adopt practices, and by 2017 they had scores comparable to those of the experimental plants. Thus, by 2017 the management practice improvements appear to have equalized across plants within treatment firms.

### C. What Drives Changes in Management Practices

We next explore the proximate causes for the adoption or nonadoption of management practices on a practice-by-practice basis in Table 3 using directors' and plant managers' stated reasons for adding or dropping practices. In column 1 we report the percentage of practices added (top panel) and dropped (bottom panel) at treatment experimental plants. In columns 2, 3, and 4 we report similar figures for the treatment nonexperimental, control experimental, and control nonexperimental plants, respectively, and in column 5 we report all plants. A few results are worth noting.

First, we see that, while a substantial fraction of practices remain unchanged from 2011, there is notable churn in management practices across all plants. In particular, 4.1 percent of practices have been added and 12.4 percent of practices dropped since the end of the experiment (see column 5). We are reasonably confident that these are accurately measured, derived as they are from detailed interviews with firm directors and plant managers combined with lengthy firm visits by the consulting team. Second, the reasons why practices change differ between treatment and control plants. In the nonexperimental plants in the treatment firms, spillovers from other plants (in the same firm) are the single largest reason for practice adoption and account for 4.2 percent of improvements (out of a total improvement rate of 6.9 percent). There are no such spillovers in any of the other three types of plant. In

TABLE 3—REASONS FOR THE CHANGE IN MANAGEMENT PRACTICES

	Treatment experimental (1)	Treatment nonexperimental (2)	Control experimental (3)	Control nonexperimental (4)	All (5)
Added practices (percent)					
New manager	1.2 (0.72)	0.6 (0.58)	0.4 (0.41)	0	0.8 (0.36)
Product, customer, or equipment change	0.7 (0.50)	1.8 (1.17)	0	0	0.9 (0.44)
Spillovers from other firms	0.7 (0.50)	0.3 (0.29)	2.2 (0.98)	2.7 (1.89)	1.1 (0.37)
Spillovers from other plants in the same firm	0	4.2 (2.39)	0	0	1.3 (0.83)
Total	2.6 (0.98)	6.9 (2.91)	2.6 (0.88)	2.7 (1.89)	4.1 (1.08)
Dropped practices (percent)					
New manager	9.9 (3.23)	0.6 (0.58)	1.8 (1.65)	1.4 (1.02)	4.6 (1.57)
Negative perceived benefit	2.9 (1.45)	3.0 (1.47)	9.3 (2.06)	1.4 (0.94)	4.2 (1.01)
Reduced director time	3.9 (1.21)	3.0 (1.26)	3.6 (0.51)	4.1 (0.86)	3.6 (0.64)
Total	16.7 (2.90)	6.6 (2.30)	14.7 (1.39)	6.9 (0.78)	12.4 (1.65)
No change (percent)					
Total	80.7 (2.75) 100	86.4 (3.98) 100	82.7 (1.81) 100	90.4 (2.67) 100	83.5 (1.82) 100

Test [treatment experimental changes = control experimental changes], using all cells,  $p$ -value = 0.03

Test [treatment experimental changes = control experimental changes], changed practice cells only,  $p$ -value = 0.04

*Notes:* The table lists the shares of practice-by-plant cells in terms of reasons for change between 2011 and 2017 by practices added, dropped, or left unchanged. These are calculated as a share of 1,042 practices, which are composed of the 38 practices across the 28 plants (11 treatment experimental, 9 treatment nonexperimental, 6 control experimental, and 2 control nonexperimental) in operation in both 2011 and 2017, except for the inventory practices, which are missing in plants that hold no inventory because they make to order. Robust standard errors are in parentheses (clustered at the plant level). The bottom two rows present  $p$ -values that test the joint hypothesis that the reasons provided for adding or dropping practices differed between treatment experimental and control experimental plants conditional on the practice having changed. The second-to-last row tests the same hypothesis but without conditioning on practices changing.

the control firms, spillovers from other firms outside the experimental group<sup>10</sup> were the most important driver of management changes, driving 2.2 percent on average of the practice upgrades (out of a total of 2.6 percent). These two figures highlight the importance of within- and across-firm spillovers in improving management practices over the long run.

In the experimental plants (in the treatment firms), the major reason for dropping practices was the introduction of a new plant manager (9.9 percent out of a total of 16.7 percent, so well over a half). The plant manager was evidently a critical part of the management improvement in the intervention plants, and if he left the firm

<sup>10</sup>Qualitatively, these improvements appear to be from copying other firms in the industry, outside of those in our experimental sample. We did not come across cases of the control firms saying they had learned from the treated firms.

then many of the practice improvements subsequently collapsed. Moreover, presumably, given that management practices will have only recently been improved in the experimental plants, they were particularly susceptible to managerial turnover as good practices may not have had time to become established norms. Another major factor across all the plants was director time—overall 3.6 percent of practices were dropped when directors had to reduce the time they spent managing the plant, often because of other business commitments (e.g., finance, marketing, or other businesses, like retail or real estate). This highlights the importance of CEO time for firm management, consistent with the work in Bandiera et al. (2018). Finally, we see that 4.2 percent of practices were dropped because of “perceived negative benefits,” which means the firms decided the practices were actually not worth adopting and decided to drop them.

Table 4 analyzes the drivers of the changes in management practices by looking at each practice-by-plant cell between 2011 and 2017 in a regression format. Hence, we examine the change in each practice (−1, 0, or 1) for each plant between 2011 and 2017 (for plants present in both years). In column 1 we see the constant term of −0.083 indicates that, on average across plants (experimental and nonexperimental plants in treatment and control firms) and practices, the average practice dropped by 8.3 percent over this period. In column 2 we control for experimental plant status and see this accounts for all of the drop, highlighting that management practice scores were roughly constant after 2011 in each of the treatment and control nonexperimental plants. In column 3 we instead add a treatment dummy and find this is completely insignificant—as can be seen from Figure 1, on average, treatment firms experienced a change in practices similar to that of control firms.

In column 4 we focus instead on the correlation of changes in practices with the frequency of usage across all plants of the practices in 2008, which is valued from 0 to 1, measuring the share of plants in the pre-experimental period that had adopted this practice. This proxies for how widespread their adoption was prior to the intervention, and the positive coefficient indicates that common practices were more likely to be maintained (so uncommon practices were more likely to be dropped). This highlights that the intervention was more successful at getting badly managed plants to adopt relatively standard practices—such as basic measurement systems—than getting plants to adopt more advanced practices like data review meetings and performance rewards. In column 5 we add these all together, and the results look similar, suggesting these are reasonably independent relationships.

In column 6 we include the management score in 2011 to look for mean reversion, finding a negative but insignificant coefficient. This is confirmed in Figure 4, which shows that both the initial treatment increase in management practices from 2008 to 2011 and the subsequent drop are uncorrelated with initial levels of management practices. So changes in management practices appear not to be strongly correlated with their initial levels, implying that, like TFP, a highly persistent autoregressive (or random-walk) form of stochastic evolution. Figure 4 is also useful in showing the distribution of changes in management practices among treated plants. We see that every single treated experimental plant improved its practices between 2008 and 2011, and every one of these plants subsequently saw a drop in its management practice score between 2011 and 2017. It is therefore not the case that there were

TABLE 4—DETERMINANTS OF CHANGES IN MANAGEMENT FROM 2011 TO 2017

DV = 0/1/-1 management score change	(1)	(2)	(3)	(4)	(5)	(6)
Experimental plant		-0.128 (0.034) [0.00]			-0.128 (0.035) [0.00]	-0.104 (0.022) [0.00]
Treatment plant			0.020 (0.037) [0.56]		-0.009 (0.032) [0.98]	0.027 (0.023) [0.27]
Frequency of practice usage in 2008				0.095 (0.037) [0.02]	0.095 (0.038) [0.03]	0.095 (0.037) [0.02]
Management score in 2011						-0.207 (0.128) [0.13]
Constant	-0.083 (0.027) [0.00]	-0.004 (0.033) [0.90]	-0.101 (0.015) [0.00]	-0.111 (0.028) [0.00]	-0.032 (0.029) [0.34]	0.041 (0.055) [0.55]
Observations	1,042	1,042	1,042	1,042	1,042	1,042

Notes: Dependent variable (DV) is the change in the -1, 0, 1 indicator for the change in management practice between 2011 and 2017. The sample is the 38 practices across the 28 plants (11 treatment experimental, 9 treatment nonexperimental, 6 control experimental, and 2 control nonexperimental) in operation across both periods, except for the inventory practices, which are missing in plants that hold no inventory because they make to order. Regressions are clustered at the firm level. Numbers in brackets are the *p*-values for testing the null hypothesis that the relevant coefficient is equal to zero, computed using the (firm) clustered wild bootstrap (999 replications using Rademacher weights).

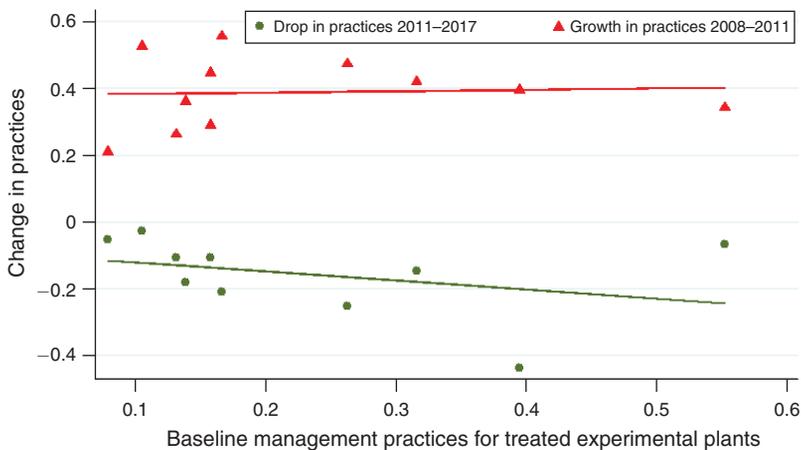


FIGURE 4. INITIAL TREATMENT INCREASES AND SUBSEQUENT POSTTREATMENT DROPS IN MANAGEMENT WERE UNCORRELATED WITH INITIAL LEVELS

Notes: Results are plotted for the sample of experimental treatment plants. Baseline management practices are the proportion of management practices employed in the plant in 2008. The red line is a fitted ordinary least squares regression of the change in practices between 2008 and 2011 against baseline practices, and the green line is the change between 2011 and 2017 against baseline practices. Neither slope is statistically significant.

some treated experimental plants in which a “Toyota way” virtuous cycle of continuous improvement occurred.

Finally, we examine the practices that were adopted to see which were the least likely to be retained and which were the stickiest. Online Appendix Table A3 reports

the number of firms that ever adopted a practice (i.e., that were not using it in 2008 and then used it in at least one of 2011 or 2017), the number that after adopting a practice were no longer using it in 2017, and the proportion of adopters who dropped the practice. We see two types of practices that were most likely to be dropped. The first is a set of visual displays and written practices that very few firms were using before the intervention and that were discarded afterward. These include displaying written procedures for warping, drawing, weaving, and beam gaiting; displaying standard operating procedures for quality supervisors; and displaying visual reports of daily efficiency by loom and weaver. The second set of practices most likely to be dropped included ones that required daily attention from management: monitoring defects on a daily basis, meeting daily to discuss quality defects and gradation, and updating visual aids of efficiency on a daily basis. They were thus costly and presumably were seen as not very valuable.

In contrast, we see that many of the practices are very sticky. Of our 38 practices, once adopted, 14 were not dropped by a single plant and a further 8 were dropped by at most one-quarter of adopters. Particularly noticeable among these sticky practices are that those that were adopted by ten or more plants and then never dropped. These relate very closely to the most immediate improvements in quality and inventory levels that we saw from the original consulting intervention: recording quality defects in a systematic manner (defect-wise), having a system for monitoring and disposing of old stock, and carrying out preventative maintenance. Finally, we note that not all daily activities were susceptible to being dropped; those most closely tied to keeping machines running were quite persistent. Firms still maintained daily monitoring of machine downtime and had daily meetings with the production team.<sup>11</sup>

Why do we see these correlations? Our preferred interpretation is one of learning. This is most plausible in the early period, when the nonexperimental plants adopted some of the practices that had been implemented in the experimental plants. There could also have been learning in the later period, when experimental plants dropped practices because the management saw that the nonexperimental plants were performing well absent these practices. This is consistent with the negative impact of a new manager in the treatment experimental plants: the new manager is not wedded to the practices and drops those that are not very useful.

An alternative explanation is there are complementarities across plants in the choice of practices. There are certainly complementarities across practices within a plant; for example, acting on machine downtime (practices 6 and 7) cannot happen if downtime is not monitored (practice 5). However, it is not obvious that similar operational complementarities exist across plant boundaries. Considering the actual practices and the nature of textile production, the one place where there might be returns to doing the same practices across plants would be at the top management level, where it would allow comparative performance evaluation of plant managers. However, our data cover only evaluation on overall performance, so we cannot address this issue.

<sup>11</sup> Breaking down the adoption status by the treatment and experimental status (e.g., “treatment nonexperimental plant”) reveals that control nonexperimental plants were the least likely to adopt any practices but conditional on adoption did not drop them subsequently.

It is worth noting that the senior consultant on our team, when asked about the drivers of practice transfer across plants, identified learning rather than other possible causes.

#### D. Results on Long-Run Performance

The other question we investigated when returning to the plants was the long-run performance impact of the original management interventions. Because of budget limitations and the reluctance of firms to share financial data, we are not able to undertake a detailed analysis of TFP.<sup>12</sup> We were able, however, to collect basic information on plant size and looms in 2014 and 2017 to supplement our original data for 2008 and 2011. Because there were changes over time in the number of plants per firm, and because the management practices have converged across plants within firms, we examine looms, employees, and management practices at the firm level.

We run intention-to-treat panel regressions over four years (2008, 2011, 2014, and 2017) at the firm level with firm and year fixed effects and standard errors clustered at the firm level:

$$(2) \quad Outcome_{i,t} = aTreat_{i,t} + b_t + c_i + e_{i,t},$$

where *Outcome* is one of the key outcome metrics of looms, looms/employee, etc. We report statistical significance both using conventional inferential procedures based on normal approximations and using permutation tests that have exact finite sample size to allay sample size concerns.<sup>13</sup> The treatment variable is a postintervention dummy taking the value 1 for 2011 onward.

In column 1 of Table 5 we regress export status (a 1/0 dummy indicating the plant exports) on the treatment dummy and find a weakly significant positive coefficient of 0.189. Both the textile firms and the management consultants reported that the improved management practices had allowed firms to raise their quality to more easily export their products. For example, one firm reported exporting fabric for tablecloths to Walmart in the United States.<sup>14</sup> In column 2 we look at the intensive margin of exporting—the log of exports—and again find a significant positive coefficient.

In column 3 we examine the number of looms the firms upgraded and find a weakly significant positive impact. The improved management practices led the firms to focus on expanding output by upgrading looms. Most of the looms they operated were 30 or more years old, purchased secondhand from US and European factories (indeed, several of them had Italian, French, or US labeling from their original owners). These machines are basic and produce simple textiles, so they are well suited to poor management practices since they need limited maintenance and

<sup>12</sup>In our original study the consulting firm spent many months extracting production data from firms' log books and production records, which were used to construct a measure of TFP. We were not able to extract these data in our longer-term follow-up. Even in our original study, where firms were getting months of advice from the consultants, they would not reveal profit data.

<sup>13</sup>We also estimate the regression at the plant level, and the results are qualitatively similar.

<sup>14</sup>Walmart is not usually seen as a "high-quality" retailer in the United States, but its quality standards for products are significantly above those of domestic Indian retailers (e.g., tablecloths sold from Walmart would not be expected to have loose threads, pattern blemishes, small holes, or frays).

TABLE 5—LONGER-RUN FIRM PERFORMANCE AND MANAGEMENT CHANGES

Dependent variable	Exporter dummy (1)	Exports (in logs) (2)	Looms upgraded (3)	Looms (in logs) (4)	Employees (in logs) (5)	Looms per employee (in logs) (6)	Any consultant days (7)	Marketing practices (score) (8)
$Treat_i \times (Year \geq 2011)_i$	0.189 (0.106) [0.24]	0.416 (0.110) [0.02]	10.275 (5.106) [0.19]	-0.032 (0.226) [0.87]	-0.269 (0.277) [0.28]	0.237 (0.090) [0.03]	0.206 (0.109) [0.18]	1.361 (0.618) [0.07]
Control group mean	0.514	3.09	1.875	4.271	5.021	-0.750	0.067	0.583
Observations	109	66	28	66	66	66	66	66

Notes: Data are from pretreatment (2008) and posttreatment (2011, 2014, and 2017) years, except for plants/firms for which basic performance data were missing or zero, and column 3, which is just for 2017. Export data are collected at the plant level, while all other variables are measured at the firm level because of the changing number of plants per firm. Marketing practices is a discrete variable from 0 to 10 defined as the count of ten 0/1 sales and marketing practice indicators like “attending trade shows,” “hiring sales and marketing professionals,” “analyzing product portfolios,” and “setting up a firm brand.” Any consultant days is a binary variable equal to 1 if the firm hired any consultants in the relevant period. Regressions are clustered at the firm level, and standard errors are in parentheses. Permutation tests, in brackets, report the  $p$ -value for testing the sharp null hypothesis of no treatment effect by constructing the permutation distribution of the estimator, using 4,000 possible permutations of firm-level random assignment.

care. But they produce far lower volumes per machine and lower-quality fabric (more frequent defects, simpler patterns and stitching, etc.). After the implementation of the management intervention the owners felt able to upgrade the looms—double-width looms, faster air- or water-jet looms, or enhanced-function looms that could perform embroidery, embossing, or Jacquard stitching. Column 4 shows, however, that the total number of looms did not change, with a statistically insignificant coefficient of  $-0.032$ , so firms focused on increasing output by upgrading looms rather than increasing the loom count.

In column 5 we examine employment. The point estimates suggest a relatively large drop in employment, of 23 to 24 percent on average over the full period, and in 2017. However, this drop is also not statistically significant. There are two reasons why employment may have fallen. The first is that, at baseline, firms employed many workers fixing quality defects and would need less of this sort of labor as quality improved. Second, production process improvements and fewer breakdowns can enable the same worker to be in charge of more looms.

Column 6 combines these measures to focus on our main measure of long-term firm productivity, which is log looms per employee. This is a classic productivity measure in the literature (see, e.g., Clark 1987, Braguinsky et al. 2015). One reason is that employees spend much of their time dealing with malfunctioning looms, so that a higher number of looms per employee indicates fewer breakdowns and higher rates of production uptime (the time the loom is producing output rather than being repaired). We find that the average treatment effect over the full postintervention period was to increase looms per employee by a statistically significant 26.7 percent.<sup>15</sup>

<sup>15</sup>Breaking the postintervention dummy into three dummies (one each for 2011, 2014, and 2017), we find that improvements were rising over time and the coefficients are generally larger for 2017 than for 2011. We also address the concern that outliers may be driving the results by winsorizing the top and bottom 10 percent of the data (each year) and find that the results do not change substantively.

We next investigate the impact on labor productivity. While we did not collect direct information on labor productivity in 2017, we can use the survey data from the initial wave to impute a labor productivity impact. In particular, we use data from a survey we ran in 2011 of 113 firms in the broader textile industry cluster around Mumbai (see details in online Appendix A2), in which we collected data on physical production, employment, and looms. Using this, we show in the online Appendix (Table A4 and Figure A1) that there is a strong correlation between labor productivity (output per worker) and looms per worker in both the cross section and the panel. Taking the fitted coefficient of 0.734 from column 4 of Table A4, we impute labor productivity from looms per employee for our experimental firms. The average imputed increase in labor productivity after 2011 is then 19.0 percent ( $\exp(0.237 \times 0.734) - 1$ ), and the long-run impact is 35.3 percent.<sup>16</sup> These impact figures are remarkably similar to the respective 15.3 percent and 31.2 percent one-year and ten-year productivity impacts reported for management interventions in postwar Italy in table 3 of Giorcelli (2019).<sup>17</sup>

In column 7 we asked the plants whether they had used any consultants since 2011, and if so, for how many days. Many of these firms had, and indeed, as column 7 shows, this use of consultants was significantly higher in the treatment plants. These consultants were local firms offering very practical advice on loom-changing practices, fabrics, human resources, or textile marketing, rather than the types of expensive international-firm management consulting provided by our intervention. We interpret this as a revealed preference indicator that treatment firms found the intervention useful and were more willing to pay for outside expert advice in the future.

Finally, in column 8 we look at the adoption of marketing practices. Marketing practices were not targeted by our initial intervention, and this enables us to examine whether changes in the specific practices on which our intervention focused are accompanied by broader management changes in untargeted areas. Our measure is a score given for the adoption of seven practices: (i) Does a director regularly attend trade shows? What is the frequency of systematically analyzing (ii) markets, (iii) products, and (iv) prices to assess policies (and make changes wherever necessary)? (v) Does the firm have a dedicated brand? (vi) Does the firm have a sales and marketing professional? (vii) Does the firm use any e-commerce (for sales) and social media (for advertising)? Treatment firms are significantly more likely to adopt these marketing practices. Discussions with firms highlighted their attempts to be more systematic in management across a range of activities. In this sense, there were cross-practice management spillovers. This is evidence consistent with the idea that improving production and human resource management practices led firms to value a more data-driven, systematic management approach and to apply this to other areas like marketing.

<sup>16</sup>The long-run impact is estimated from a regression that breaks down the postintervention dummy into three separate treatment dummies (one for each year). The coefficient on the treatment dummy interacted with a dummy for 2017 is 0.412 and  $\exp(0.412 \times 0.734) - 1 = 35.3$ .

<sup>17</sup>The results are also similar to the one-year impact of 17 percent reported in Bloom et al. (2013).

### III. Conclusions

In summary, the intervention in 2008–2010 did have lasting effects, but not the multiplier effect of ongoing further improvements that the “Toyota way” theory would have predicted. Indeed, a significant fraction of the induced improvements were dropped, especially if the plant manager changed, if the directors were short of time, or if the practices were not common before the intervention. Still, many of the changes persisted—particularly those involving quality control and inventory management—and spread throughout the treatment firms, resulting in long-run improvement in worker productivity. Thus, the “inappropriate technologies” view does not find much support. Beyond that, the “three-year life” conventional wisdom ascribed to management change programs described in the introduction is also decisively rejected, at least for the sort of practice changes our intervention induced.

Interestingly, the treatment firms also used more consulting and did more marketing, suggesting that the more systematic approach to management introduced by the intervention was spreading to other areas the intervention had not addressed. These broader lasting impacts highlight the importance of management in explaining persistent productivity differences among firms. Understanding why more firms do not invest in improving management, and what types of policies can change this, is therefore an important area for future research.

## APPENDIX A

### I. Plant Sample

Online Appendix Table A2 reports the sample of plants by the four types (treatment and control, experimental and nonexperimental). As noted in the text, one treatment firm exited because of the death of the owner without any male heirs, which led to the closure of one plant. Two more treatment plants closed because they were amalgamated into other plants within the same firm—that is, all the looms and equipment were moved onto one site for production economies of scale. We count these as plant closures (since those plants stopped operating). Finally, both treatment and control firms opened some plants over this period due to demand growth.

### II. Management Survey in 2011 and Imputing Labor Productivity

Between November 2011 and January 2012 we ran an in-person survey of textile firms around Mumbai with 100 to 1,000 employees, using the Ministry of Commercial Affairs registry of firms plus a combination of industry lists, internet searches, and referrals as a sample frame (for more sampling details, see online appendix A2 of Bloom et al. 2013). We identified 172 such firms and were able to interview 113 of them (17 project firms and 96 nonproject firms). The main purpose of this survey was to benchmark the management practices of our experimental sample against the industry as a whole, and we found that our project firms did not differ significantly in management practices from the nonproject firms interviewed.

The interview followed a relatively standardized script, asking background questions about the firm (age, ownership, family involvement, markets, etc.), followed by questions about plant size (employees, output, plant numbers, production quantity), management practices, organizational structure, computerization, prior consulting, prior knowledge of the Stanford–World Bank project (we skipped this question for firms involved in the experiment), and any potential interest in future consulting waves. The full survey is available at <https://nbloom.people.stanford.edu/research>.

In this paper, we use the data collected in this survey on the annual physical output of the firm (in meters or production picks), the number of employees (permanent plus contract), and the number of looms in the firm. We attempted to collect this information for four years (2008–2011); we were able to do so for all four years for 87 firms, and for two or three years for a further 7 firms. Using these data, we construct labor productivity as the log of physical production units per worker. This is similar to the sales per worker term often used to measure labor productivity, but it has the advantage of not incorporating price effects.

Online Appendix Figure A1 shows the strong correlation (0.561) between labor productivity and looms per employee. Online Appendix Table A4 presents the corresponding regression relationship. Column 1 shows the strong cross-sectional relationship, which persists after adding year fixed effects (column 2), firm fixed effects (column 3), and both year and firm fixed effects (column 4). Column 4 then shows that annual changes in looms per employee are associated with changes in labor productivity. This yields the fitted relationship

$$\log \text{ production per worker} = \underset{\text{(SE 0.114)}}{0.734} \times \log \text{ looms per worker} \\ + \text{ year effect} + \text{ firm fixed effect}.$$

We use this fitted relationship to impute labor productivity impacts from our impact on looms per worker in Table 5.<sup>18</sup>

## REFERENCES

- Anderson, Steven J., Rajesh Chandy, and Bilal Zia.** 2016. “Pathways to Profits: Identifying Separate Channels of Small Firm Growth through Business Training.” World Bank Group Policy Research Working Paper 7774.
- Bandiera, Oriana, Renata Lemos, Andrea Prat, and Raffaella Sadun.** 2018. “Managing the Family Firm: Evidence from CEOs at Work.” *Review of Financial Studies* 31 (5): 1605–53.
- Bennesden, Morten, Kasper Meisner Nielsen, Francisco Perez-Gonzales, and Daniel Wolfenzon.** 2007. “Inside the Family Firm: The Role of Families in Succession Decisions and Performance.” *Quarterly Journal of Economics* 122 (2): 647–91.
- Bertrand, Marianne, and Antoinette Schoar.** 2003. “Managing with Style: The Effect of Managers on Firm Policies.” *Quarterly Journal of Economics* 118 (4): 1169–208.

<sup>18</sup>If we just use the baseline cross-sectional association in the project firms, the coefficient is 1.16 (SE 0.30). This is not statistically different from the estimate using all firms, and using the coefficient estimated on the basis of all firms is more conservative given the lower point estimate, as well as allowing us to rely more on time variation for identification.

- Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts.** 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics* 128 (1): 1–51.
- Bloom, Nicholas, and John Van Reenen.** 2007. "Measuring and Explaining Management Practices across Firms and Countries." *Quarterly Journal of Economics* 122 (4): 1351–408.
- Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson.** 2015. "Acquisitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry." *American Economic Review* 105 (7): 2086–119.
- Bruhn, Miriam, Dean Karlan, and Antoinette Schoar.** 2018. "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico." *Journal of Political Economy* 126 (2): 635–87.
- Capelli, Peter, and David Neumark.** 2001. "Do 'High-Performance' Work Practices Improve Establishment-Level Outcomes?" *ILR Review* 54 (4): 737–75.
- Clark, Gregory.** 1987. "Why Isn't the Whole World Developed? Lessons from the Cotton Mills." *Journal of Economic History* 47 (1): 141–73.
- DellaVigna, Stefano, and Devin Pope.** 2018. "Predicting Experimental Results: Who Knows What?" *Journal of Political Economy* 126 (6): 2410–56.
- Economist.** 2009. "Good to Great to Gone." July 7, 2009. <https://www.economist.com/business/2009/07/07/good-to-great-to-gone>.
- Giorcelli, Michela.** 2019. "The Long-Term Effects of Management and Technology Transfers." *American Economic Review* 109 (1): 121–52.
- Groh, Matthew, Nandini Krishnan, David McKenzie, and Tara Vishwanath.** 2016. "The Impact of Soft Skills Training on Female Youth Employment: Evidence from a Randomized Experiment in Jordan." *IZA Journal of Labor and Development* 5: 9.
- Higuchi, Yuki, Edwin P. Mhede, and Tetsushi Sonobe.** 2019. "Short- and Medium-Run Impacts of Management Training: An Experiment in Tanzania." *World Development* 114: 220–36.
- Hirscheifer, Sarojini, David McKenzie, Rita Almeida, and Cristobal Ridao-Cano.** 2016. "The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey." *Economic Journal* 126 (597): 2115–46.
- Hsieh, Chiang-Tai, and Peter J. Klenow.** 2010. "Development Accounting." *American Economic Journal: Macroeconomics* 2 (1): 207–23.
- Huselid, Mark A.** 1995. "The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance." *Academy of Management Journal* 38 (3): 635–72.
- Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi.** 1997. "The Effects of Human Resource Management Practices on Productivity." *American Economic Review* 87 (3): 291–313.
- Karlan, Dean, Ryan Knight, and Christopher Udry.** 2015. "Consulting and Capital Experiments with MicroEnterprise Tailors in Ghana." *Journal of Economic Behavior and Organization* 118: 281–302.
- Kiechel, Walter.** 2012. "The Management Century." *Harvard Business Review*, November. <https://hbr.org/2012/11/the-management-century>.
- Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton.** 2015. "The Value of Bosses." *Journal of Labor Economics* 33 (4): 823–61.
- Liker, Jeffrey K.** 2004. *The Toyota Way—14 Management Principles from the World's Greatest Manufacturer*. New York: McGraw-Hill.
- Marshall, Alfred.** 1887. "The Theory of Business Profits." *Quarterly Journal of Economics* 1 (4): 477–81.
- McKenzie, David, and Christopher Woodruff.** 2014. "What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *World Bank Research Observer* 29 (1): 48–82.
- McNair, Sam.** n.d. "Training Has a Half-Life." <https://www.lce.com/Training-Has-a-HalfLife-1299.html>.
- Osterman, Paul.** 1994. "How Common Is Workplace Transformation and Who Adopts It?" *ILR Review* 47 (2): 173–88.
- Roberts, John.** 2018. "Needed: More Economic Analyses of Management." *International Journal of the Economics of Business* 25 (1): 3–10.
- Sirkin, Harold L., Perry Keenan, and Alan Jackson.** 2005. "The Hard Side of Change Management." *Harvard Business Review*, October. <https://hbr.org/2005/10/the-hard-side-of-change-management>.
- Syverson, Chad.** 2011. "What Determines Productivity?" *Journal of Economic Literature* 49 (2): 326–65.
- Walker, Francis A.** 1887. "The Sources of Business Profits." *Quarterly Journal of Economics* 1 (3): 265–88.