DO MANAGEMENT INTERVENTIONS LAST? EVIDENCE FROM INDIA

Nicholas Bloom[a], Aprajit Mahajan[b], David McKenzie[c] and John Roberts[d]

October 15, 2017

Abstract:

We revisit the plants described in Bloom et al. (2013) eight years after the end of the original management intervention, finding three main results. First, while about half of the management practices adopted in the original experimental plants had been dropped, there was still a large significant gap between the treatment and control plants. Likewise, there was a still a significant performance gap between treatment and control plants, suggesting lasting impacts of effective management interventions. Second, many management practice had spread from the experimental plants to the non-experimental plants within the same firm, suggesting large inter-firm spillovers. Third, managerial turnover and the lack of Directors time were two of the biggest drivers of the drop in management practices in experimental plants, highlighting the importance of key employees for management practices.

JEL No. L2, M2, O14, O32, O33.

Keywords: management, organization, productivity, and India.

Acknowledgements: Financial support was provided by SEED at the Graduate School of Business at Stanford and 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.

[a] Stanford Economics; [b] UC Berkeley ARE ; [c] Development Research Group, The World Bank; [e] Stanford GSB

I. INTRODUCTION

After an early focus on management as a driver of differences in firm performance (e.g. Walker, 1887 and Marshall 1887) economists are again paying increasing attention to the role of management in firm and economy-wide performance (Gibbons and Roberts, 2013). Whereas the size and wealth of the management consulting industry is often cited as one revealed preference measure of the importance of management, recent academic work has established a credible causal link between changes in management practices and performance (Bloom et al, 2013; Bruhn et al, 2017). The longer-term persistence of management improvements caused by consulting interventions remains an open question. The received wisdom at a leading global management consulting firm when two of the authors were employed there was that such innovations lasted approximately three years.

Competing views of management offer differing predictions about the persistence in consulting induced improvements in management practices. One view, best exemplified by the "Toyota way" (Liker, 2004) sees management improvements delivering a continuous cycle of improvement as systems put in place for measuring, monitoring, and improving operations and quality, enable the firm to continually improve. A related idea is that management practices are complementary to one another, so that the costs of adding new practices falls as others are put in place. For example, in our context, scientific management of inventory levels will only be possible once the firm has put in place systems to record all yarn transactions and to regularly monitor stock levels. Some evidence for the for the lasting impacts of changes in management practices on firm performance comes from Giorcelli (2016), who finds that Italian firms that received Marshall Plan sponsored management training trips to the U.S. in the 1950s experienced significantly better performance over the next fifteen years (relative to firms that applied for but did not receive the training).

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 (The Economist, 2009; Kiechel, 2012). This may be even harder when changes are introduced externally, with the Boston Consulting Group reporting that two-thirds of transformation initiatives ultimately fail (Sirkin et al, 2005). One reason may be that these practices may be inappropriate and will be abandoned as firms learn that they are not suitable in their setting. Both

Karlan et al. (2015) and Higuchi et al. (2016) find that light consulting engagements in smaller firms than ours led to firms gradually discarding practices over the subsequent three years.

This paper examines the persistence of management practices adopted after an extensive consulting-supported intervention that we undertook in a set of multi-plant Indian textile weaving firms from 2008 to 2010 (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 were given recommendations on improving management practices in several areas, and the treatment plants received additional consulting 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 when we presented our work --- about whether these changes would last. So, in 2017 we returned to the same firms with support from the same consulting firm. We find that both treatment and control experimental plants had in fact dropped some practices, but not as much as 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 (noting the treatment firms had multiple plants and we usually only worked with one plant) had adopted many of the recommendations, so their package of current practices were very close to those of the treatment plants.

We were also able to get some limited performance data and to gather some information on why practices were dropped. We find that practices are more likely to be dropped when the plant manager changes, when the Directors (the CEO and CFO) are busy, and when the practice is 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[1], while the latter emphasizes the importance of beliefs. Despite dropping some practices, we do find treated firms show lasting improvements in worker productivity (as measured by looms per worker), that treated firms have

[1] 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. (2017).

gone on to use more consulting services of their own accord, and that they have supplemented the operational management practices introduced by the consultants with better marketing practices.

II. THE 2008-2010 CONSULTING EXPERIMENT

II.A. The Experimental Design

To investigate the impact of management on firm productivity we initiated a randomized controlled intervention on management practices in a set of large textile plants near Mumbai in 2008. This 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 all 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 $13 million and annual sales of $7.5 million. Compared to US manufacturing firms, these firms would be in the top 1% by employment and the top 4% by sales, and compared to India manufacturing they are in the top 1% by both measures (Hsieh and Klenow, 2010). Hence, these are large manufacturing firms.[2]

These firms are complex organizations, with a median of 2 plants per firm (in addition to a head office in Mumbai) and 4 reporting levels from the shop-floor to the managing director. In all the firms, the managing director was the largest shareholder, and all directors belonged to the same family. Two firms were publicly listed on the Mumbai Stock Exchange, although more than 50% of the equity in each of these was held by the managing family.

The field experiment aimed to improve management practices in the treatment plants and we measured the impact of doing so on firm performance. We contracted a large international management consultancy firm (Accenture) to work with the plants as the easiest way to change plant-level management practices rapidly. The full-time team of (up to) 6 consultants came from leading Indian business and engineering schools, and most of them had prior experience working with U.S. 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

---

2 Note that most international agencies define large firms as those with more than 250 employees.

standard in American, European, and Japanese firms and which can be grouped into five broad areas – factory operations, quality control, inventory control, human-resources management, and sales and orders management (for details see Appendix Table A1). Each practice is measured as a binary indicator of the adoption (1) or non-adoption (0) of the practice. A general pattern at baseline was that plants recorded a variety of information (often in 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 this 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 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 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 firms 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.

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).

II.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 38 percentage points by August 2010 (approximately one year after the start of the intervention). This adoption rate dropped by 3 percentage points in the second year of tracking, showing persistence in practices after the consultants had left. Not all practices were adopted equally, with firms adopting the practices that were (unsurprisingly) the easiest to implement and/or had the largest perceived short-run pay-offs, 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 they 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, increased their adoption of the management practices, but by only 12 percentage points on average. This is substantially less than the increase in adoption in the treatment firms, indicating that the four months of the implementation phase were important in changing management practices. Column 2 shows a statistically significant 20 percentage point treatment effect on management practices in 2011 as a result. However, this change for the control firms is still an increase relative to the rest of the industry around Mumbai (the non-project plants), which did not change their management practices on average between 2008 and 2011.

Finally, since these are multi plant firms and the consulting firm worked at the plant level, the treatment and control firms also had "non-experimental plants". So, for example, if a treatment Firm has three 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 Non-Experimental plants". Likewise if a control firm had plants D, E and F and the diagnostic intervention was only performed on plant D, then D would be an "Control Experimental plant" while E and F would be "Control Non-Experimental plants". 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 non-experimental plants, it was still able to collect bi-monthly management data and some basic plant data. For these non-experimental plants in the treatment firms, we also saw a substantial increase in the adoption of management practices. In these 5 plants the adoption rates increased by 17.5 percentage points by August 2010, without any drop back in the second year. This increase occurred because the owners of the treatment firms copied the new practices from their experimental plants over to their other (non-experimental) plants. Interestingly, this increase in adoption rates is similar to the control firms' 12 percentage point increase, suggesting that the copying of best practices across

plants within firms can be as least as effective at improving management practices as short (1-month) bursts of external consulting.

II.C. The Initial Experimental Results – Firm Performance

Treatment firms experienced a significant increase in output of 9.4% relative to the control firms, 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% due to both the increase in output and a drop in inputs due to reduced inventory and a reduction in labor inputs for mending defective fabric. These improvements were estimated to have increased profits per plant by about $325,000 per year.

We also evaluated the number of plants per firm in 2011, which was significantly higher in the treatment firms both relative to the control firms and the rest of the industry. In conversations with the directors at the treatment firms we learned that better management practices enabled each family director to oversee more production activity, as the additional data allowed them to delegate more decisions to their plant managers while still closely monitoring them. Expansion occurred through increasing the number of plants (as opposed to expanding current plants). This may have been simply because the design of the plants or availability of large plots of land did not support enlargement, but it is also possible that multiple plants reduced the risk of unionization and regulatory problems.

III. THE 2017 FOLLOW UP

III.A. The Follow-up Process

In January 2017, working with the same consulting firm we had worked with in 2008-2011, we re-contacted the 17 textile firms from our original study. Fortunately, all 17 firms agreed to work with the research team again on a second wave. This 100% uptake was aided by a combination of three factors: (A) the positive impact of the intervention in the first wave on the firms' management and performance; (B) the stability of the firms, which, as large firms, had maintained the same address and contact details, and (C) the engagement of the same two

Accenture partners and project manager as the 2008-2011 intervention.[3] One complication is that one single-plant treatment firm was in the midst of closing down after the death of the owner. Without any close male relatives to continue the business, the owner's wife had decided to sell the business, which, given its location, meant the business would stop trading and the site would be converted into residential housing.[4]

One drawback of this follow-up wave is that our budget allowed us only two months of consulting time, which was sufficient to collect management data for all production sites and a basic set of basic firm performance indicators (e.g. on employment and looms), but not to collect detailed weekly output data (which would have required extracting data on a firm-by-firm basis from log-books and accounting software). As such, our analysis is therefore confined to management practices and basic performance indicators like employment or looms/employee.

This follow-up data collection corresponds to an average period of 9 years since the implementation phase of the consulting intervention started, and 7 years since it ended. It therefore enables us to examine the long-term persistence of these large changes in management practices.

III.B. Results on Management Practices

In Figure 1 we plot the management scores from re-visiting 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 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% of the original 0.35 increase (noting these firms started pre-treatment 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% in the original increase in management practices.

Control Experimental Plants: Second, the 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

---

3 These kinds of 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 the 2008-2011 wave.

4 The firm was over 30 years old, and due to the expansion of Mumbai was now located in a primarily residential area so the land was more valuable as housing than for production.

absolute terms compared to the fall in scores relative to 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% of the intervention gain, implying about a 13% depreciation rate of the original management increase.

Together this indicates that, first, even eight years after the initial intervention the treatment firms still had higher management practices. Table 2 reports the results from running the Ancova specification:

$$Management_{i,t} = a + b_1*Treatment_i*Year{=}2011 + b_2*Treatment_i*Year2017 + c*Management_{i,2008} + e_{i,t}$$

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 statistically significant both using conventional (large sample normality based) inference as well as permutation procedures with exact finite sample size (the corresponding p-values are 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%) versus the treatment plants (40%) suggests that small improvements in management may be less stable that 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.

<u>Non-experimental plants</u>: Third, the non-experimental plants in the treatment firms showed no net change, with their management practice adoption rates remaining constant at 0.47. Indeed, by 2017 their management scores were very similar overall to the treatment experimental plants (indeed slightly higher, although not significantly so). Similarly, in the control firms the non-experimental plants also converged with the experimental plants (again slightly higher but not significantly). This suggests (as we discuss further below) that the practice improvements in the experimental plants spilled over to the non-experimental plants during the eight years after the experiment.

Expectations on durability of the intervention: Finally, before we re-contacted the firms in December 2016, each member of the Accenture team from the 2008-2010 intervention and the academic team provided predictions for the management scores we expected to find on revisiting the firms in 2017.[5]

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 to not only persist, but to continue to improve 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 each of the seven team members are shown for the treatment experimental, treatment non-experimental, control experimental plants and control non-experimental plants with the symbols TE, TN, CE and CN respectively on the graph.[6] 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, particular for the non-experimental 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 continual virtuous improvement and of 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 on each practice on the X-axis against the average scores for the non-experimental plants (in the treatment firms) on the Y-axis, for the years 2008 (pre-intervention), 2011 (post-intervention) and 2017 (long-run follow-up). We observe that initially the experimental and non-experimental plants in the treatment firms had similar practice scores, with a correlation of 0.91. After the intervention, the scores for the experimental plants improved considerably leading to an eastward shift in the points and a drop in the correlation to 0.81 (top-right figure). Finally, in the bottom left figure we

---

[5] 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 (2017).

[6] The predictions of the individual Accenture and Stanford team members were made independently – Bloom estimated first and then the other team members individually e-mailed him their predicted scores. The average Accenture and Stanford predicted scores were not particularly different (hence we present them averaged together).

see the experimental plants and non-experimental plants again have very similar scores (correlation of 0.91), with a reversion of the scores towards the 45 degree line.

Figure 3 complements this by showing the long-difference of management practices in the experimental and non-experimental plants (in the treatment firms) between 2008 and 2017 (left-panel) and 2011-2017 (right panel). What this highlights is, firstly, between 2008 and 2017 both sets of plants adopted similar bundles of management practices. But, secondly, looking at 2011-2017 we see the timing of these practice adoptions were not the same. The experimental plants adopted most of these practices between 2008-2011, so that from 2011 to 2017 they mostly had negative practice changes. The non-experimental 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 by 2011 they have substantially higher management scores. These scores subsequently subside as some practices are dropped. The *non-experimental* plants adopted fewer practices in 2008-2010 but continued to adopt practices, and by 2017 had comparable scores with the experimental plants. Thus, by 2017 the management practice improvements appear to have equalized over across plants within treatment firms.

III.C. What Drives Changes in Management Practices

We next explore the proximate causes for the adoption or non-adoption 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 the “Treatment experimental” column we report the percentage of practices added (top panel) and dropped (bottom panel). In the second, third and fourth panels we report similar figures for the “Treatment non-experimental”, “Control Experimental” and “Control Non-experimental”, while reporting all plants in the final column. A few results are worth noting.

First, we see that, while a significant fraction of practices remains unchanged from 2011, there is notable churn in management practices across all plants. In particular, 4.1% of practices have been added and 12.4% of practices dropped since the end of the experiment. We are reasonably confident that these are accurately measured, derived as they are from detailed interviews with firm directors and plant managers. Second, in the non-experimental plants in the treatment firms,

spillovers from other plants (in the same firm) is the single largest reason for practice adoption and accounts for 2.4% of improvements (out of a total improvement rate of 6.9%). In the control firms' spillovers (from other firms[7]) were also the most important driver of management improvements, driving 2.4% on average of the practice upgrades (out of a total of 2.6%). These two figures highlight the importance of within and across firm spillovers in improving management practices over the long run.

Third, in the experimental plants (in the treatment firms) the major reason for dropping practices was the introduction of a new manager (9.9% out of a total of 16.7% 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 then many of the practice improvements subsequently collapsed.[8] Another major factor across all the plants was director time – overall 3.6% 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 of Bandiera et al. (2017). Finally, we see that 4.2% of practices were dropped because of "perceived negative benefits" which means the firms decided the practices were 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 (-1, 0 or 1) across all plants present in both years. In column (1) we see the constant term of 0.083 indicates that, on average across all the experimental and non-experimental plants in treatment firms, and control plants, the management score dropped by 8.4%. In column (2) we control for experimental plant status and see this accounts for all the drop, highlighting that management practices scores were roughly constant after 2011 in the non-experimental treatment 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 plants did not change (experimental treatment

[7] 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.

[8] See also Fryer (2017) who argues that principal turnover was the primary reason for declines in school performance improvements following an experimental intervention aimed at changing school management practices in the United States.

plants dropped their management score and non-experimental treatment plants increased their management score). In column (4) we control for having a new-manager,[9] and given the results in Table 3, split this by treatment and control, and see for treatment plants a large significant negative (which is driven by the treatment experimental plants) with nothing significant for control plants. This highlights the role of managerial turnover in the drop in management practices in well managed plants. Moreover, presumably given the management practices will have only recently improved in the experimental plants they are particularly susceptible to managerial turnover as good practices may not have had time to become established norms.

In column (5) we focus instead on the frequency of usage across all plants of the practices in 2008, which is value 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 (6) we add these all together and the results look similar, suggesting these are reasonably independent relationships (we also tried various cross-products with nothing notably significant). Finally, in column (7) we include the management score in 2011 to look for mean reversion, finding a negative but insignificant coefficient.

Finally, we examine the practices that were adopted to see which were the least likely to be retained, and which were the most sticky. Table A3 reports the number of firms which ever adopted a practice (i.e. were not using it in 2008, and then used it in at least one of 2011, 2014, or 2017), the number who after adopting were no longer using the practice 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 are a set of visual displays and written practices that very few firms were using before the intervention and then were discarded afterwards. These include displaying written procedures for warping, drawing, weaving and beam gaiting; displaying standard operating

[9] We test if having a new plant-manager is differential across treatment and control, or experimental and non-experimental and find no significant difference (p-values of 0.824 and 0.701 respectively).

procedures for quality supervisors; and displaying visual aids of daily efficiency by loom and weaver. The second set of practices most likely to be dropped were 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.

In contrast, we see that many of these practices are very sticky. 14 of our 38 practices, once adopted, are not dropped by a single plant, and a further 8 are dropped by at most one-quarter of those adopting. Particularly noticeable among these sticky practices are that those which were adopted by 10 or more plants and then never dropped 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 old stock, and carrying out preventative maintenance. Finally, we note that not all daily activities were susceptible to being dropped, with those most closely tied to keeping machines running quite persistent: firms still maintain daily monitoring of machine downtime and have daily meetings with the production team.

III.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 we could not undertake a detailed analysis of TFP (as we did in the original study) since this involved detailed analysis of firms log books and production records. 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. Hence, in our analysis, we run Intention to Treat (ITT) panel regressions over four years (2008, 2011, 2014 and 2017) at the plant level with plant and year dummies and standard errors clustered at the firm-level:

$$OUTCOME_{i,t} = aTREAT_{i,t} + b_t + c_i + e_{i,t}$$

where OUTCOME is one of the key outcome metrics of looms, looms/employee etc. We report statistical significance using both conventional inferential procedures based on normal approximations as well as using permutation tests that have exact finite sample size to allay sample size concerns.

In our baseline estimation we define the time-varying variable "TREAT" equal to one for any plant in a treatment firm – experimental and non-experimental (given the results in section III.B

showing that by 2017 both had similar management practices) – post 2008. Plants in the control firms (experimental and non-experimental) are all included in the control group given the evidence from Figure 1 and Table 4 of complete spillovers during this period. Since there was plant entry and exit (as indicated in Appendix Table A2) this is an unbalanced panel.

We start in column (1) of Table 5 in the top panel looking at looms (in logs) which is a basic measure of production. We regress this on a dummy for the year being greater or equal to 2011 - a post-intervention dummy - finding a coefficient of 0.296. This implies that on average plants in treated firms_saw a long-run increase in looms of about 34% (= exp(0.296)-1) compared to plants in control firms. In panel B, we break this down into the experimental and the non-experimental plants, and find in fact the coefficient is greater in the non-experimental plants, although because of the small sample size these differences are not significant[10]. Finally, in panel C we allow the joint treatment on the experimental and non-experimental plants to vary by year, and find some variation in the points estimates but nothing suggesting a significant trend.

In column (2) we examine log looms per employee, which is a classic productivity measure in the literature (see, for example, Clark 1987 or Braguinsky et al. 2015). The reason is employees spend much of their time dealing with malfunctioning looms, so that higher numbers of looms per employee indicate fewer breakdowns and higher rates of production uptime (the time the loom is producing output rather than being repaired). As such column (2), panel A, shows that the long-run treatment effect increased looms per employee by 9%. Panel B shows this efficiency improvement occurred in both the experimental and non-experimental plants in the treatment firms. Panel C suggests this impact of the management intervention on increasing efficiency was rising over time in that the coefficients are generally larger for 2017 that 2011, which is consistent with the intervention having introduced feedback systems (like the performance monitoring and daily meetings) generating a gradual improvement in efficiency.

In column (3) we asked the plants if they had used any consultants since 2011, and if so how many days. Many of these firms had, and indeed, as column (3) shows, this use of consultants was weakly 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 our intervention involved. We interpret this as a revealed preference indicator that treatment firms found the

---

10 P-value of the F-test is

intervention useful and were more willing to pay for commercial consulting in future. This was similar across the experimental and non-experimental plants (panel B), and more likely to occur once some time had passed since their previous consulting use in our project (panel C).

Finally, in column (4) we look at the adoption of marketing practices. Marketing practices were not part of our initial intervention, and so this enables us to examine whether changes in the specific practices our intervention focused on are accompanied by broader management changes. Our measure is a score given for the adoption of seven practices: (1) does a Director regularly attend trade shows; what is the frequency of systematically analyzing markets, products and prices to assess policies (and make changes wherever necessary) ((2), (3) and (4)); (5) does the firm have a dedicated brand; (6) does the firm have a sales and marketing professional; and (7) does the firm use any e-commerce (for sales) and social media (for advertising). Panel A shows that treatment firms are significantly more likely to adopt these marketing practices. Discussions with firms highlighted two reasons. First, the increased production required a more aggressive sales push. Second, they reported attempts to be more systematic in management across a range of firm activities. So, in this sense, there were cross-practice management spillovers. Improving production and human-resource management practices led firms to value a more systematic management approach, and hence apply this to other areas like marketing. As we see in panel B of column (4) these improvements were similar in the experimental and non-experimental plants. Somewhat oddly the panel C shows these sales and marketing practices were highest in 2011 and 2017 but lower in 2014, which is likely idiosyncratic error.

IV. CONCLUSIONS

In summary, the intervention in 2008-2010 did have lasting effects, but there was not the multiplier effect of on-going further improvements that the "Toyota Model" would have predicted. Indeed, a significant fraction of the induced improvements were dropped, especially if the plant manager changed, the Directors were short of time, or if the practices were not common before the intervention. Still, many of the changes persisted, and they spread throughout the treatment companies. There was also some persistence and some drop in the control plants' set of practices. So, the "inappropriate technologies" view is also not supported. Beyond that, the "three-year life" conventional wisdom described in the introduction fails decisively.

The treatment firms were still much better managed in 2017 than the control, and key practices around quality control and inventory management were kept. Moreover, the treatment firms used more consulting and did more marketing, suggesting that the mores systematic approach to management introduced by the intervention was spreading to areas the intervention had not addressed, and we see long-term benefits in terms of a measure of worker productivity. These lasting impacts highlight the importance of management in explaining persistent productivity differences amongst firms. Understanding why more firms do not invest in improving management, and what types of policies can change this, is therefore an important question for future research.

References

Bandiera, Oriana Renata Lemos, Andrea Prat and Raffaella Sadun, (2017) "Managing the Family Firm: Evidence from CEOs at Work.", *Review of Financial* Studies, forthcoming.

Bennesden, Morten, Kasper Nielsen, Francisco Pérez-Gonzáles and Daniel Wolfenzon, (2007). "Inside the Family Firm: The Role of Families in Succession Decisions and Performance", Quarterly Journal of Economics, 122(2), 647-691.

Bertrand, Marriane and Schoar, Antoinette (2003), "Management with style", *Quarterly Journal of Economics*

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

Braguinsky, Serguey, Atsushis Ohyama, Tetsuji Okazaki and Chad Syverson. 2015. "Acquisition, Productivity and Profitability: Evidence from the Japanese Cotton Spinning Industry." *American Economic Review,* 105(7): 2086-2119.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar (2017), "The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico" *Journal of Political Economy*, forthcoming.

Clark, Greg (1987), "Why Isn't the Whole World Developed? Lessons from the Cotton Mills" *Journal of Economic History*, vol. 47, 141-173.

Dellavigna,Stefano and Devin Pope (2017) "Predicting Experimental Results: Who knows what?", *Journal of Political Economy*, forthcoming.

The Economist (2009) "Good to great to gone", July 7.

Fryer, Roland (2017) "Management and Student Achievement: Evidence from a Randomized Field Experiment", Harvard Working Paper.

Gibbons, Robert and John Roberts (2013) *The Handbook of Organizational Economics*. Princeton University Press, Princeton, NJ.

Giorcelli, Michela (2016) "The long-term effects of management and technology transfer: Evidence from the US Productivity Program", Mimeo. Stanford.

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 Mhede, and Tetsushi Sonobe (2016) "Short- and Longer-Run Impacts of Management Training: An Experiment in Tanzania", Mimeo.National Graduate Institute for Policy Studies, Tokyo.

Hirschleifer, 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): 2115-2146.

Hsieh, Chiang-Tai, and Pete Klenow (2010), "Development Accounting," *American Economic Journal: Macroeconomics,* 2(1), 207-223.

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.

Lazear, Edward, Shaw, Kathryn and Stanton, Christopher (2016), "The value of bosses", *Journal of Labor Economics*

Liker, Jeffrey K. (2004*). The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer.* McGraw-Hill

Marshall, Alfred, (1887), “The Theory of business profits”, *Quarterly Journal of Economics*, 1(4), pp. 477-481.

Sirkin, Harold, Perry Keenan and Alan Jackson (2005) “The Hard side of Change Management”, *Harvard Business Review*, October.

Walker, Francis (1887), “On the sources of business profits”, *Quarterly Journal of Economics,* 1 (3) , 265-288.

Appendix

Plant sample:

Table A2 reports the sample of plants by the four types (treatment and control, experimental and non-experimental). 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 into the same firm – that is all the looms and equipment were moved onto one site for production economies of scale. We count these as a plant closure (since that plant stopped operating) but the output of that plant will still be included at the firm-level. Finally, both treatment and control firms opened some plants over this period due to demand growth.

Table A1: The textile management practices adoption rates

Area	Specific Practice	2008	2011	2017
Factory Operations	Preventive maintenance is carried out for the machines*	0.4	0.7	0.95
	Preventive maintenance is carried out per manufacturer's recommendations*	0.1	0.15	0.15
	The shop floor is marked clearly for where each machine should be	0.1	0.3	0.25
	The shop floor is clear of waste and obstacles	0.05	0.3	0.3
	Machine downtime is recorded*	0.6	0.9	0.9
	Machine downtime reasons are monitored daily*	0.45	0.9	0.85
	Machine downtime analyzed at least fortnightly & action plans implemented to try to reduce this*	0..05	0.65	0.6
	Daily meetings take place that discuss efficiency with the production team*	0.05	0.7	0.8
	Written procedures for warping, drawing, weaving & beam gaiting are displayed	0.1	0.45	0
	Visual aids display daily efficiency loomwise and weaverwise	0.25	0.7	0.4
	These visual aids are updated on a daily basis	0.15	0.6	0.25
	Spares stored in a systematic basis (labeling and demarked locations)	0.1	0.2	0.4
	Spares purchases and consumption are recorded and monitored	0.5	0.55	0.35
	Scientific methods are used to define inventory norms for spares	0	0.05	0.1
Quality Control	Quality defects are recorded*	0.95	1	1
	Quality defects are recorded defect wise	0.25	0.85	0.95
	Quality defects are monitored on a daily basis*	0.3	1	0.5
	There is an analysis and action plan based on defects data*	0.05	0.7	0.3
	There is a fabric gradation system	0.55	0.85	1
	The gradation system is well defined	0.45	0.85	0.45
	Daily meetings take place that discuss defects and gradation*	0.15	0.75	0.3
	Standard operating procedures are displayed for quality supervisors & checkers	0.05	0.6	0
Inventory Control	Yarn transactions (receipt, issues, returns) are recorded daily*	0.89	1	1
	The closing stock is monitored at least weekly*	0.28	0.83	0.56
	Scientific methods are used to define inventory norms for yarn	0	0	0
	There is a process for monitoring the aging of yarn stock	0.28	0.538	0.72
	There is a system for using and disposing of old stock*	0.05	0.78	0.56
	There is location wise entry maintained for yarn storage*	0.28	0.61	0.5
Loom Planning	Advance loom planning is undertaken	0.35	0.55	0.1
	There is a regular meeting between sales and operational management	0.5	0.6	0.45
Human Resources	There is a reward system for non-managerial staff based on performance*	0.6	0.7	0.6
	There is a reward system for managerial staff based on performance*	0.3	0.45	0.2
	There is a reward system for non-managerial staff based on attendance	0.35	0.5	0.5
	Top performers among factory staff are publicly identified each month	0.15	0.25	0.2
	Roles & responsibilities are displayed for managers and supervisors	0.05	0.5	0.5
Sales and Orders	Customers are segmented for order prioritization	0	0	0.11
	Orderwise production planning is undertaken	0.67	0.89	1
	Historical efficiency data is analyzed for business decisions regarding designs	0	0.1	0.08
All	Average of all practices	0.271	0.576	0.466

Notes: Reports the 38 individual management practices for all treatment plants (both experimental and non-experimental, unbalanced panel) in 2008, 2011 and 2017.

Table A2: Plant count

	2008	2011	2014	2017
Treatment – experimental	14	14	11	11
Treatment – non experimental	6	9	9	9
Control – experimental	6	6	6	6
Control – non-experimental	2	2	4	4
Total	28	31	30	30

Notes: Lists the total number of plants in 2008 to 2017 plus ever (adds all dead and alive plants). One firm closed in 2014, so the total number of firms was 17, 17, 16 and 16 across the first four columns.

Table A3: Practice stickiness

		Adopted	Dropped	Share Dropped
9	Written procedures for warping, drawing, weaving & beam gaiting are displayed	7	7	1.00
22	Standard operating procedures are displayed for quality supervisors & checkers	11	10	0.91
11	These visual aids are updated on a daily basis	11	7	0.64
10	Visual aids display daily efficiency loomwise and weaverwise	11	6	0.55
21	Daily meetings take place that discuss defects and gradation*	13	7	0.54
18	There is an analysis and action plan based on defects data*	14	7	0.50
17	Quality defects are monitored on a daily basis*	16	6	0.38
4	The shop floor is clear of waste and obstacles	6	2	0.33
33	There is a reward system for non-managerial staff based on attendance	9	3	0.33
20	The gradation system is well defined	8	2	0.25
24	The closing stock is monitored at least weekly	13	3	0.23
7	Machine downtime analyzed at least fortnightly & action plans implemented to try to reduce this	15	3	0.20
8	Daily meetings take place that discuss efficiency with the production team*	19	3	0.16
5	Machine downtime is recorded*	9	1	0.11
6	Machine downtime reasons are monitored daily*	13	1	0.08
27	There is a system for using and disposing of old stock*	15	1	0.07
1	Preventive maintenance is carried out for the machines*	10	0	0.00
12	Spares stored in a systematic basis (labeling and demarked locations)	6	0	0.00
16	Quality defects are recorded defect wise	20	0	0.00
19	There is a fabric gradation system	9	0	0.00
26	There is a process for monitoring the aging of yarn stock	11	0	0.00
28	There is location wise entry maintained for yarn storage*	7	0	0.00
35	Roles & responsibilities are displayed for managers and supervisors	9	0	0.00
37	Orderwise production planning is undertaken	6	0	0.00

Notes: Lists the practices ordered by the share of adopters between 2008 and 2011 that subsequently dropped them by 2017.

Table 1: The field experiment sample pre-intervention (2008)

	All				Treatment	Control	Diff
	Mean	Median	Min	Max	Mean	Mean	p-value
<u>Sample sizes:</u>							
Number of plants	28	n/a	n/a	n/a	19	9	n/a
Number of experimental plants	20	n/a	n/a	n/a	14	6	n/a
Number of firms	17	n/a	n/a	n/a	11	6	n/a
Plants per firm	1.65	2	1	4	1.73	1.5	0.393
<u>Firm/plant sizes:</u>							
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 mtrs, experimental plants	5560	5130	2260	13000	5,757	5,091	0.602
<u>Management and plant ages:</u>							
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 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 non-experimental 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 textiles plants per firm. Several of these firms have other businesses – for example retail units and real-estate arms – which 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 sales office. **Employees per experiment plant** reports the number of employees in the experiment 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 the general manager and general manager to the managing director would have 4 hierarchical levels. **BVR Management score** is the Bloom and Van Reenen (2007) management score for the experiment plants. **Management adoption rates** are the adoption rates of the management practices listed in Table A1 in the experimental plants. **Annual sales ($m)** and **Current assets ($m)** are both in 2009 US $million values, exchanged at 50 rupees = 1 US Dollar. **Daily mtrs, 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 1600 suits daily. **Age of experimental plant (years)** reports the age of the plant for the experimental plants.

Table 2: Short and long run impact on management practices

Dep Var: Proportion of management practices implemented	(1)	(2)
Treatment*Year=2011	0.206***	0.249***
	(0.042)	(0.038)
Treatment*Year=2017	0.197**	0.218**
	(0.062)	(0.057)
Year=2017	-0.122***	-0.122***
	(0.016)	(0.016)
Baseline Management Score	0.668**	0.878***
	(0.219)	(0.176)
P-value of test of equality of treatment in 2011 and 2017	0.802	0.457
Sample Size	37	34

Notes: Notes: Robust standard errors in parentheses, clustered at the firm level. *, **, and *** denote significance at the 10, 5, and 1 percent levels respectively.

Table 3: Reasons for the change in management practices

	Treatment Experimental	Treatment Non-Experimental	Control Experimental	Control Non-Experimental	All
Added Practices (%)					
New manager	1.2	0.6	0.4	0	0.8
Product, customer or equipment change	0.7	1.8	0	0	0.9
Spillovers from other firms	0.7	0.3	2.2	2.7	1.1
Spillovers from other plants in the same firm	0	4.2	0	0	1.3
Total	2.6	6.9	2.6	2.7	4.1
Dropped Practices (%)					
New Manager	9.9	0.6	1.8	1.4	4.6
Perceived negative benefit	2.9	3.0	5.3	1.4	4.2
Reduced directors time	3.9	3.0	3.6	4.1	3.6
Total	16.7	6.6	10.7	6.9	12.4
No Change (%)	80.7	86.4	86.7	90.4	83.5
Total	100	100	100	100	100

Notes: Lists the shares of practice by plant cells in terms of reasons for change between 2011 and 2017 in terms of practices added, dropped or left unchanged. Calculated as a share of 1,042 practices, which are comprised of the 38 practices across the 28 plants (11 treatment experimental, 9 treatment non-experimental, 6 control experimental and 2 control non-experimental) in operation in both 2011 and 2017, except for the inventory practices which are missing in plants which hold no inventory because they make to order.

Table 4: Determinants of changes in management from 2011 to 2017

DV=0/1/-1 management score change	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Experimental plant		-0.128**				-0.098***	-0.097***
		(0.046)				(0.021)	(0.022)
Treatment plant			0.020			0.047	0.043
			(0.037)			(0.029)	(0.023)
New plant manager*treated				-0.103**		-0.096**	-0.075*
				(0.047)		(0.038)	(0.045)
New plant manager*control				-0.035		-0.007	-0.010
				(0.029)		(0.027)	(0.036)
Frequency of practice usage in 2008					0.095**	0.095**	0.095**
					(0.037)	(0.037)	(0.037)
Management score in 2011							-0.132
							(0.160)
Constant	-0.083***	0.050	-0.101***	-0.048**	-0.111***	-0.052*	-0.052*
	(0.027)	(0.046)	(0.015)	(0.023)	(0.028)	(0.027)	(0.027)
Observations	1,042	1,042	1,042	1,042	1,042	1,042	1,042

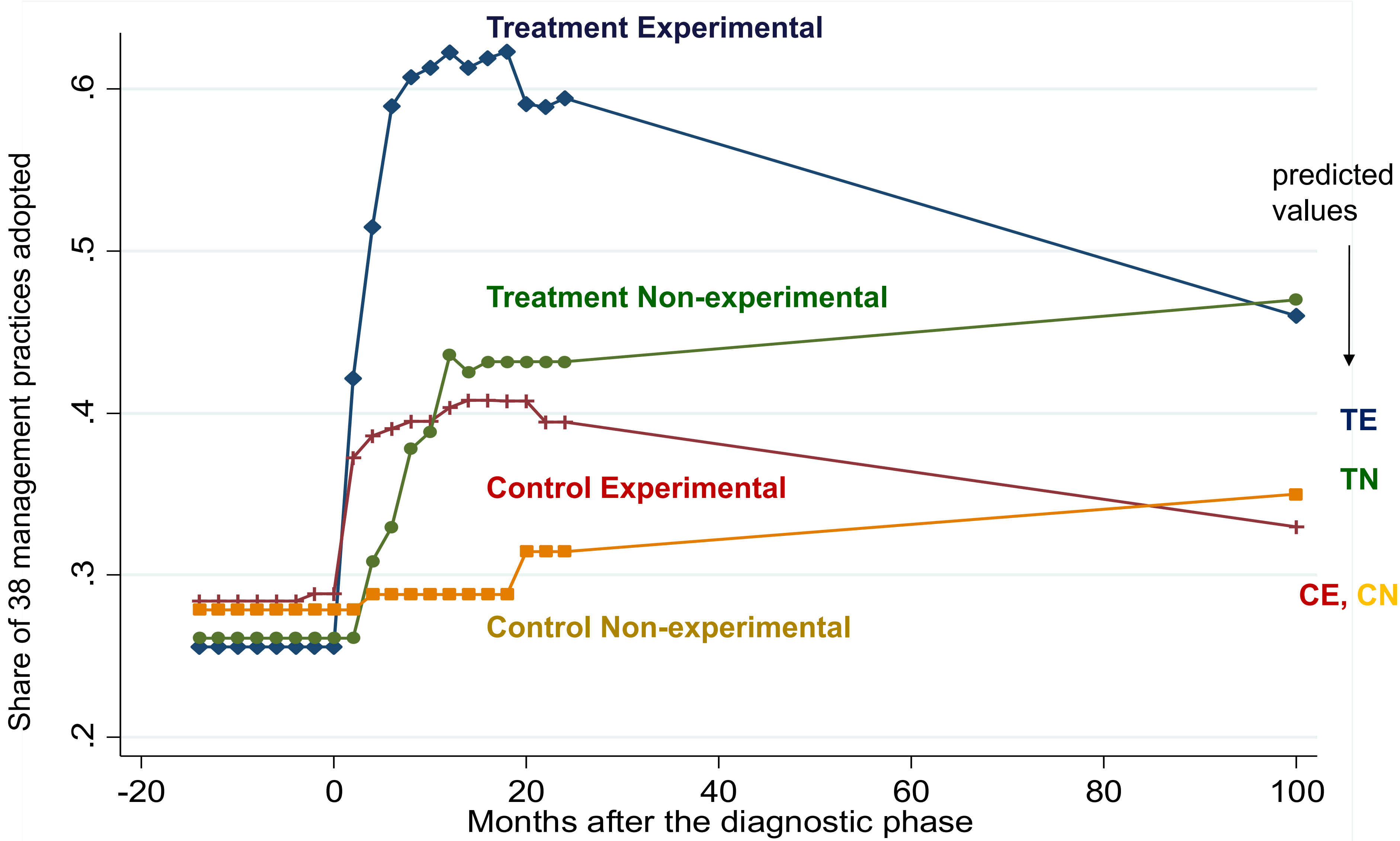
Notes: Dependent variable is the change in the -1,0,1 indicator for the change in management practice between 2011 and 2017. The sample is the 38 practices across the 28 plants (11 treatment experimental, 9 treatment non-experimental, 6 control experimental and 2 control non-experimental) in operation across both periods, except for the inventory practices which are missing in plants which hold no inventory because they make to order. Regressions clustered at the firm level. *** denotes 1%, ** denotes 5%, * denotes 10%

Table 5: Longer-run performance and management changes

Dep Var	Looms (in logs) (1)	Looms per employee (in logs) (2)	Consulting days (in logs) (3)	Marketing practices (4)
Panel A: Long-run performance				
$Treatment_i$*$(Year>=2011)_t$	0.296**	0.088**	1.414*	1.405**
	(0.120)	(0.038)	(0.666)	(0.514)
Panel B: Experimental and non-Experimental plants				
Experimental*$Treatment_i$*$(Year>=2011)_t$	0.171*	0.300**	1.21**	1.31**
	(0.074)	(0.139)	(0.53)	(0.55)
Non-Experimental*$Treatment_i$*$(Year>=2011)_t$	0.511***	0.300*	2.08	1.70***
	(0.067)	(0.137)	(1.39)	(0.52)
Panel C: Treatment impact by period				
$Treatment_i$*$(Year==2011)_t$	0.123	0.163	-0.073	1.149**
	(0.076)	(0.101)	(0.080)	(0.450)
$Treatment_i$*$(Year==2014)_t$	0.100	0.289*	1.859*	-1.494**
	(0.082)	(0.147)	(0.943)	(0.518)
$Treatment_i$*$(Year==2017)_t$	0.296*	0.451**	2.77**	2.294**
	(0.138)	(0.168)	(1.120)	(0.884)
F-test $Treatment_i$*$(Year==2014)_t$ & $Treatment_i$*$(Year==2017)_t$	0.123	0.047	0.073	0.015
Control group mean (all in levels)	57.6	0.509	0.114	0.486
Years	2008, 11, 14, 17	2008, 11, 14, 17	2008, 11, 14, 17	2008, 11, 14, 17
Firms	17	17	17	17
Plants	31	31	31	31
Observations	109	109	109	109

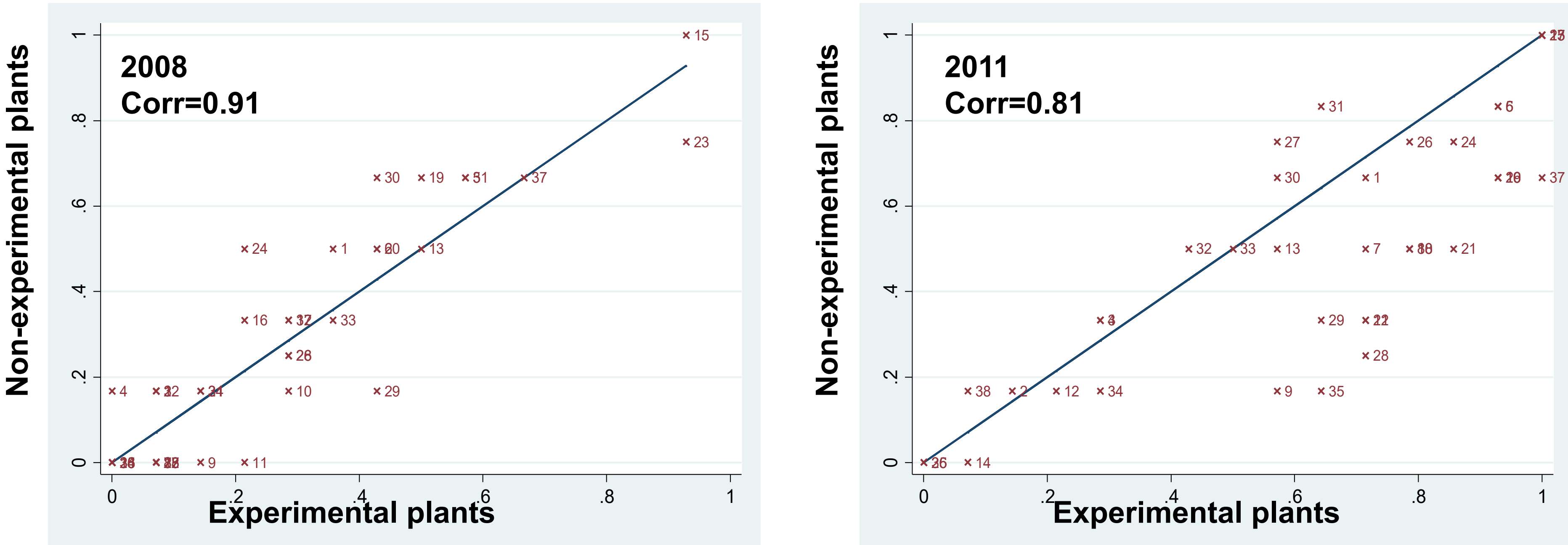
Notes: Data for pre-treatment (2008) and post-treatment (2011, 2014 and 2017) years, except plants for which basic performance data was missing. Sales and marketing practices is an indicator from 0 to 10 defined as the count of ten 0/1 Sales and Marketing practices like “Attending trade shows”, “Hiring sales and marketing professionals”, “Analyzing product portfolios”, “Setting up a firm brand”. Regressions clustered at the firm level. *** denotes 1%, ** denotes 5%, * denotes 10%. F-test reports p-value of the joint test.

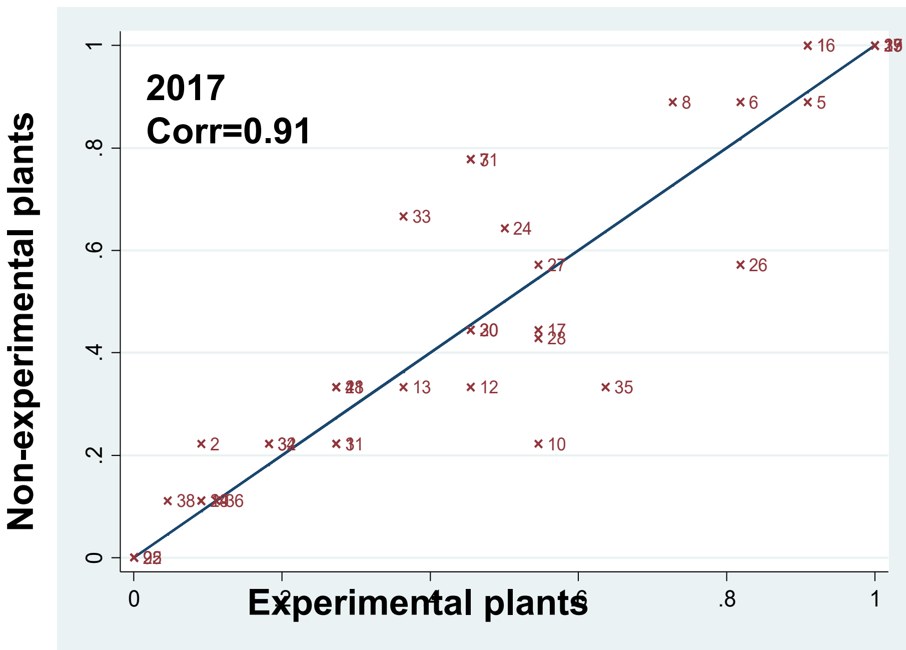
Figure 1: Management practices by plant group



Notes: Sample comprised of the balanced panel of plants from 2008 to 2017 (11 treatment experimental, 6 treatment non-experimental, 6 control experimental and 2 control non-experimental. The letters on the right are the average predicted values from the 3-person Accenture team and 4 co-authors made before re-contacting the firms for the Treatment Experimental (TE) at 0.4, Treatment Non-Experimental (TN) at 0.36, Control Experimental and Control Non-Experimental (CE and CN) both at 0.29 respectively.

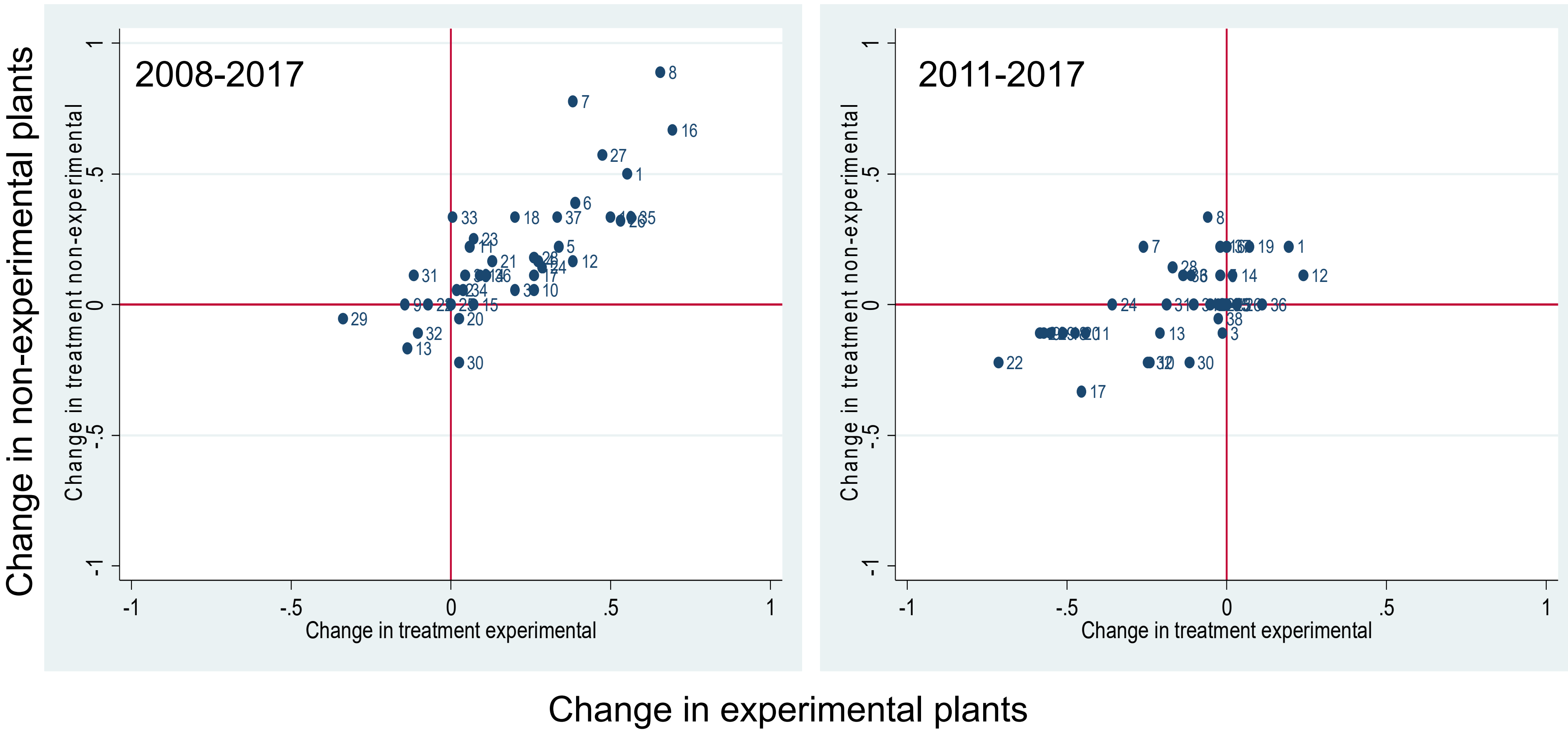
Figure 2: Practices appear to spread out fully in treatment firms





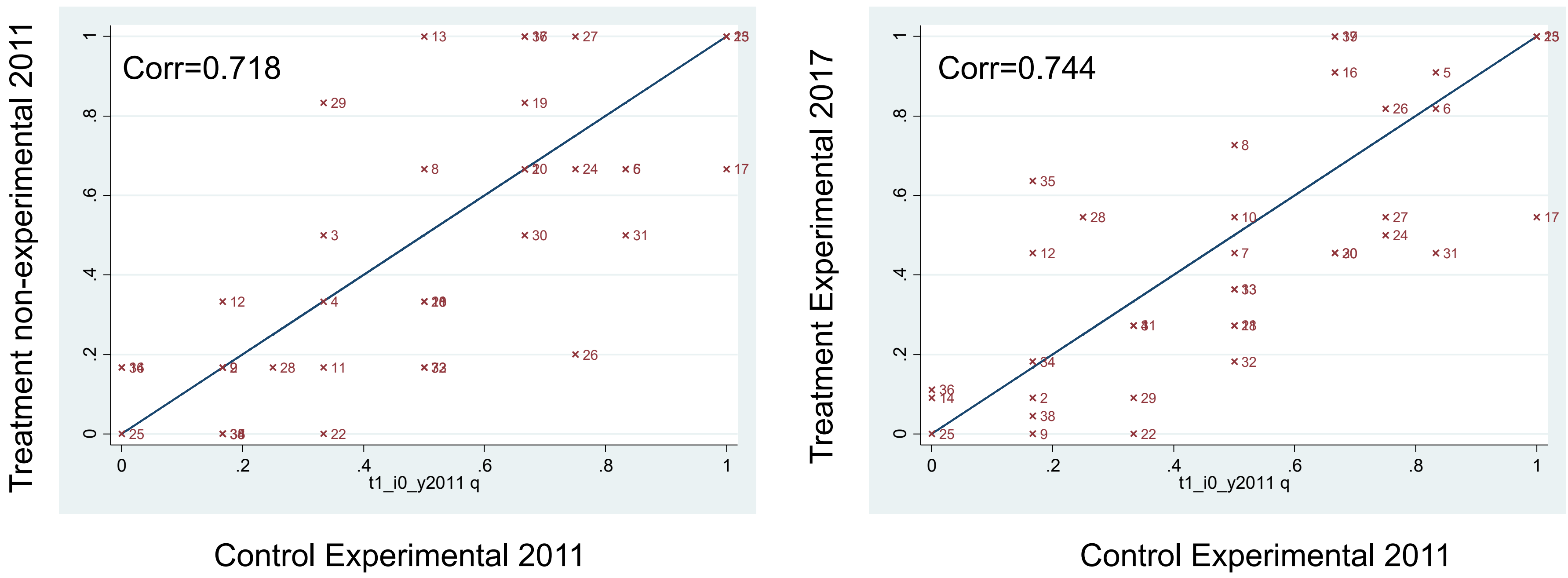
Note: The three graphs plot the average scores for each of the 38 questions for the 14 (11 in 2017) treatment experimental plants (on the x-axis) and the 6 treatment non-experimental plants (on the y-axis) in 2008 (top-left), 2011 (top-right) and 2017 (bottom-left). The correlations between these scores for the 38 practices are reported as well on the graphs.

Figure 3: Changes in experimental and non-experimental plants in the treatment firms between 2008-2017 and 2011-2017



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 non-experimental plants (on the y-axis) between 2008 and 2017 (left panel) and 2011-2017 (right panel).

Figure A1: Control plants in 2011 had similar scores to treatment non-experimental firms in 2011 and treatment experimental firms in 2017, but a different practice mix



Note: Plots the average scores for each of the 38 questions for the 6 control plants (x-axis) in 2011 vs 6 treatment non-experimental plants in 2011 (left plot) and 11 treatment experimental plants in 2017 (right plot) on the y-axis