DOES MANAGEMENT MATTER? EVIDENCE FROM INDIA

Nicholas Bloom[a] , Benn Eifert[b], Aprajit Mahajan[c],

David McKenzie[d] and John Roberts[e]

<u>Preliminary draft</u>**: September 11, 2010**

<u>Abstract:</u>

A long-standing question in social science is the extent to which differences in management cause differences in firm performance. To investigate this we ran a management field experiment on large Indian textile firms. We provided free consulting on modern management practices to a randomly chosen set of treatment plants and compared their performance to the control plants. We find that adopting modern management practices had three main effects. First, it raised average productivity by about 10%, through improved quality and efficiency and reduced inventory. Second, it increased the decentralization of decision making, as the better flow of information enabled owners to delegate more decisions to plant managers. Third, it increased the use of computers, necessitated by the data collection and analysis involved in modern management. Since these practices were profitable this raises the question of why firms had not adopted these before. Our results suggest that informational barriers were a primary factor in explaining this lack of adoption. Modern management is a type of technology that diffuses slowly between firms, with many Indian firms unaware of its existence or impact. And since competition was limited by constraints on firm entry and growth, badly managed firms were not rapidly driven from the market.

<u>JEL No</u>. L2, M2, O14, O32, O33.

<u>Keywords:</u> management, organization, IT, productivity and India.

<u>Acknowledgements:</u> We would like to thank the Alfred Sloan Foundation, the Freeman Spogli Institute and the International Initiative at Stanford, the Graduate School of Business at Stanford, the International Growth Centre, the Kauffman Foundation, the Murty Family, the Knowledge for Change Trust Fund, the Technology Network for Information Technology, and the World Bank for their substantial financial support. This research would not have been possible without our partnership with Kay Adams, James Benton and Breck Marshall, the dedicated work of the consulting team of Asif Abbas, Saurabh Bhatnagar, Shaleen Chavda, Karl Gheewalla, Shruti Rangarajan, Jitendra Satpute, Shreyan Sarkar, and Ashutosh Tyagi, and the research support of Troy Smith. We thank our formal discussants Susantu Basu, Naushad Forbes, Ramada Nada and Paul Romer, as well as seminar audiences at the AEA, Barcelona GSE, Boston University, Chicago, Columbia, the EBRD, Harvard Business School, IESE, Katholieke Universiteit Leuven, Kellogg, , the LSE, the NBER, PACDEV, Stanford, UCL, UCLA and the World Bank for comments.

[a] Stanford Economics, SCID, CEP and NBER; [b] Berkeley Economics; [c] Stanford Economics and SCID; [d] The World Bank, IZA and BREAD; [e] Stanford GSB

I. INTRODUCTION

Economists have long puzzled over why there are such astounding differences in productivity across both firms and countries. For example, US plants in homogeneous industries like cement, block-ice, white pan bread and oak flooring display 100% productivity spreads between the 10^{th} and 90^{th} percentile (Foster, Haltiwanger and Syverson, 2008).

A natural explanation for these productivity differences lies in variations in management practices. Indeed, the idea that "managerial technology" determines the productivity of inputs goes back at least to Walker (1887) and is central to the Lucas (1978) model of firm size.[1] Yet while management has long been emphasized by the media, business schools and policymakers, economists have typically been skeptical about its importance.

One reason for skepticism is the belief that competition will drive badly managed firms out of the market. As a result any residual variations in management practices will reflect firms' optimal responses to differing market conditions. For example, firms in developing countries may not be adopting quality control systems because wages are so low that repairing defects is cheap. Hence, their management practices are not "bad", but the optimal response to low wages.

A second reason for this skepticism is the complexity of management, making it hard to measure and quantify.[2] Yet recent work has down-played the "soft skill" attributes of good managers or leaders such as charisma, ingenuity and the ability to inspire – which can be difficult to measure, let alone change – in order to focus on specific management practices which can be measured, taught in business schools and recommended by consultants. Examples of such practices include key principles of Toyota's "lean manufacturing," the continual analysis and refinement of quality control procedures, inventory management, and advanced human resource practices. A growing literature in both economics and management measures many such management practices and finds large variations across establishments and a strong association between these practices and higher productivity and profitability.[3]

This paper seeks to provide the first experimental evidence of the importance of management practices in large firms. We use a randomized consulting design and collect detailed time-series data on management practices and plant performance. The experiment takes large multi-plant Indian textile firms and randomly allocates their plants to management treatment and control groups. Treatment plants received five months of extensive management consulting from a large

[1] Francis Walker's 1887 paper entitled "*On the sources of business profits*" discussed the extent to which variations in management across firms were responsible for their differences in profitability. Walker was an important character in the early years of the economics discipline as the founding president of the American Economics Association, the second president of MIT, and the Director of the 1880 Economic Census.

[2] Lucas (1978, p. 511) notes that his model "does not say anything about the tasks performed by managers, other than whatever managers do, some do it better than others".

[3] See for example, Osterman (1994), Huselid (1995), MacDuffie (1995), Ichniowski, Prennushi and Shaw (1998), Cappelli and Neumark (2001) and Bloom and Van Reenen (2007). A prominent early example is Pack (1987), which, like the present study, deals with textile firms in developing countries. In related work, Bertrand and Schoar (2003) use a manager-firm matched panel and find that manager fixed effects matter for a range of corporate decisions. Lazear and Oyer (2009) and Bloom and Van Reenen (2010) provide extensive surveys.

international consulting firm, which diagnosed areas for improvement in a set of management practices in the first month, followed by four months of intensive support in implementation of these recommendations. The control plants received only the one month of diagnostic consulting.

The treatment intervention led to significant improvements in quality, inventory and production efficiency (defined as the fraction of time the weaving machines were running) The result was an increase in productivity of about 10% and a substantial increase in profitability of about $250,000. Firms also spread these management improvements from their treatment plants to other plants they owned, providing revealed preference evidence on their beneficial impact.

The improvements were substantial because our sample of plants had very poor management practices prior to the intervention. Most of them had not adopted basic procedures for efficiency, inventory or quality control that have been commonly used for several decades in European, US and Japanese firms. Since these practices do not typically require any capital expenditure and were successfully introduced during the intervention period (albeit with the help of the consulting firm), this raises the question of why these practices had not been adopted previously.

Our evidence suggests that one important factor was informational constraints – the Indian firms were not aware of many existence or impact of modern management practices that are common in Western and Japanese firms. Management practices evolve over time, with innovations like the Taylor's Scientific Management, Sloan's M-form corporation and Toyota's lean production spreading slowly across firms and countries. For example, the US automotive industry took at least two decades to understand and adopt Japanese lean manufacturing.

A related question is why product market competition did not drive these badly managed firms out of business. One reason is the reallocation of market share to well managed firms is restricted by span of control constraints on firm growth. In every firm in our sample only members of the owning family are in senior managerial positions. The reason is that the family owners are worried about non-family members stealing from the firm. For example, if they let non-family managers procure yarn they might buy this at inflated rates from friends and receive kick-backs. Since the rule of law is weak in India, legal sanctions are not much of a deterrent against this type of theft.

The number of adult males available to fill senior positions thus becomes a binding constrain on growth. In fact, the best predictor of the size of the firms in our sample was the number of male family members. All the biggest firms, which all had multiple plants, were run by multiple brothers, while the best managed firm had only one plant because the founder had no brothers or sons. Hence, well managed firms do not always grow large and drive unproductive firms out of the market.[4] Meanwhile, entry is limited by the financing costs, since our textile firms have an average of $13m of assets.[5]

---

[4] This helps to explain the lack of reallocation across firms in China and India (Hsieh and Klenow, 2009).

[5] Another related question is given the large profits from improving management practices why don't consulting firms generate more business? One obvious constraint is that potential client firms are approached all the time by companies offering cost saving products – from cheap telephone lines to better weaving machines – so simply contacting firms to tell them about the huge profits from consulting need not be effective. Of course the consultants could offer their services in return for profit sharing with the firms. But profit sharing is hard to enforce ex post as firms can hide their profit numbers from the consultants, as they do frequently from the tax authorities. Moreover,

We also find two other major results for the impact of better management practices. First, the firms' owners decentralized greater decision making power over hiring, investment and pay to their plant managers. This happened in part because the improved collection and dissemination of information enabled owners to monitor their plant managers better, reducing the risk of managerial theft. As a result owners felt more comfortable in delegating decision making power to middle managers.

Second, the extensive data collection and processing requirements of modern management led to a rapid increase in computer use. For example, installing production quality control systems requires firms to record individual quality defects and then to analyze these by shift, loom, weaver and design. So modern management practices appear to be a skill-biased technical change (SBTC) as increased computerization raises the demand for educated employees. A large literature has highlighted SBTC as a key factor increasing income inequality since the 1970s, with this experiment providing some evidence on the role of modern management as one mechanism.[6]

The major drawback of our experiment is the small cross-sectional sample size. We have data on only 28 plants across 17 firms. So to address concerns over statistical inference in small samples we extend some permutations tests from bio-statistics to develop non-asymptotic significance tests. We also exploit our large time series of around 100 weeks of data per plant by using some recent large T (rather than large N) estimators. We believe these approaches are useful both for addressing sample concerns in our paper and also potentially for other field experiments on large organizations, regions or villages where the data has a small cross-section but large time series.

This paper relates to several strands of literature. First, there is the extensive productivity literature which reports large spreads in total-factor productivity (TFP) across plants and firms in dozens of developed countries. From the outset this literature has attributed much of this spread to differences in management practices (Mundlak, 1961), but problems in measurement and identification have made this hard to confirm (Syversson, 2010). This dispersion in productivity appears even larger in developing countries (Banerjee and Duflo, 2005, Hsieh and Klenow, 2009). Despite this, there are still very few experiments on productivity in firms (McKenzie, 2009) and none involving the sort of large multi-plant firms studied here.

Second, our paper builds on the literature on the management practices of firms. This has a long debate between the "best-practice" view that some management practices are universally good and all firms would benefit from adopting these (Taylor, 1911) and the "contingency view" that every firm is already adopting optimal practices but these are different from firm to firm (e.g. Woodward, 1958). Much of the empirical literature trying to distinguish between these views has traditionally been case-study or survey based, making it hard to distinguish between the different explanations and resulting in little consensus in the management literature.[7] This paper provides experimental evidence that a core set of management best practices do exist.

---

the firms worry this will lead the consultants to bias their work to short-term results at the cost of longer term performance. As a result in India (as in the rest of the world) consulting is rarely offered on a profit-sharing basis.

[6] See, for example, the surveys in Acemoglu 2002 and Autor, Katz and Kearney 2008.

[7] See, for example, the surveys in Delery and Doty (1996) and Bloom and Van Reenen (2010).

Third, it links to the large theoretical literature on the organization of firms. Papers generally emphasize optimal decentralization either as a way to minimize information processing costs or as a way to trade off incentives and information within a principal-agent model.[8] But the empirical evidence on delegation is limited, focusing on natural experiments like the adoption of on-board computers in trucking (e.g. Baker and Hubbard, 2004) or de-layering in large publicly traded US firms (Rajan and Wulf, 2006, Guadalupe and Wulf, 2010). In this paper we have the first experimental evidence on decentralization.

Fourth, it links the rapidly growing literature on Information Technology (IT) and productivity. A growing body of work has emphasized the relationship between technology and productivity, emphasizing both the direct productivity impact of IT and also its complementarity with modern management and organizational practices (i.e. Bresnahan et al. 2002 and Bartel, Ichniowski and Shaw, 2007). But again the evidence on has focused on panel IT and organizational survey data, with no prior experimental data. Our experimental evidence suggests one route for the impact of computers on productivity is via facilitating better management practices, and this occurs simultaneously with the decentralization of production decisions.

Finally, recently a number of other field experiments in developing countries (for example Karlan and Valdivia, 2010, Bruhn et al. 2010 and Drexler et al. 2010) have begun to estimate the impact of business training in microenterprises. This work focuses on training the owners in tasks such as separating business and personal finances, basic accounting, marketing and pricing. This research generally finds significant effects of these business skills on performance in small firms, supporting our results on management practices in larger firms.

II. MANAGEMENT IN THE INDIAN TEXTILE INDUSTRY

II.A. Why work with firms in the Indian textile industry?

Despite rapid growth over the past decade, India's one billion people still have a per-capita GDP (in PPP terms) of only one-seventeenth of the United States. Labor productivity is only 15 percent of that in the U.S. (McKinsey Global Institute, 2001). While average levels of productivity are low, most notable is the large variation in productivity, with a few highly productive firms and a lot of low-productivity firms (Hsieh and Klenow, 2009).

In common with other developing countries for which data is available, Indian firms are also typically poorly managed. Evidence from this is seen in Figure 1, which plots results from the Bloom and Van Reenen (2010) double-blind telephone surveys of manufacturing firms in the US and India. The Bloom and Van Reenen (BVR) methodology scores establishments from 1 (worst practices) to 5 (best practices) on specific management practices related to monitoring, targets

[8] See, for example, Bolton and Dewatripont (1994) and Garicano (2000) for examples of information processing models and Aghion and Tirole (1997), Baker, Gibbons and Murphy (1999), Rajan and Zingales (2001), Hart and Moore (2005), Acemoglu et al. (2007) and Alonso et al. (2008) for examples of principal-agent models. Recent reviews of this literature are contained in Garicano and Van Zandt (2010), Mookherjee (2010) and Gibbons and Roberts (2010).

and incentives. This yields a basic measure of the use of modern management practices that is strongly correlated with a wide range of firm performance measures like productivity, profitability and growth. The top panel of Figure 1 plots the these management practice scores for a sample of 751 randomly chosen US manufacturing firms with 100 to 5000 employees and the second panel for similarly sized Indian ones. The results reveal a thick tail of badly run Indian firms, leading to a much lower average management score (2.69 for India versus 3.33 for US firms). Indian firms tend not to collect and analyze data systematically in their factories, they tend to use less effective target-setting and monitoring, and they typically employ ineffective promotion and reward systems. The scores for Brazil and China in the third panel, with an average of 2.67, are very similar, suggesting that Indian firms are representative of large manufacturing firms in the emerging economies.

In order to implement a common set of management practices across firms and measure a common set of outcomes, we focus on one industry. We chose textile production since it is the largest manufacturing industry in India, accounting for 22% of manufacturing employment. The fourth panel shows the management scores for the 232 textile firms in our Indian sample, which look very similar to Indian manufacturing in general.

Within textiles, our experiment was carried out on 28 plants operated by 17 firms in the woven cotton fabric industry. These plants weave cotton yarn into cotton fabric for suits, shirting and home furnishing. They purchase yarn from upstream spinning firms and send their fabric to downstream dyeing and processing firms. As shown in the bottom panel of Figure 1, the 17 firms involved had an average BVR management score of 2.60, very similar to the rest of Indian manufacturing. Hence, our particular sample of 17 Indian firms also appears broadly similar in terms of management practices to manufacturing firms in the developing countries.[9]

II.B. The selection of firms for the field experiment

The firms were randomly chosen from the population of all public and privately owned textile firms in Maharashtra state, based on lists provided to us by the Ministry of Corporate Affairs (MCA). We restricted attention to firms with between 100 to 1000 employees, yielding a sample of 529.[10] We chose 100 employees as the lower threshold because by this size firms require systematic management practices to operate efficiently. We chose 1000 employees as the upper bound to avoid working with conglomerates and multinationals. Within this group we focused on firms in the cotton weaving industry (US SIC code 2211) as the largest single 4-digit SIC group within textiles. Geographically we focused on firms in the towns of Tarapur and Umbergaon because these provide the largest concentrations of textile firms in the area, and concentrating on two nearby towns reduced travel time for the consultants. This yielded a sample of 66 potential subject firms.

[9] Interestingly, prior work on the Indian textile industry suggested its management practices were also inferior to those in Europe in the early 1900s (Clark, 1987).

[10] The MCA list comes from the Registrar of Business, with whom all public and private firms are legally required to register annually. Of course many firms do not register in India, but this is generally a problem with smaller firms, not with 100+ employee manufacturing firms which are too large and permanent to avoid Government detection.

All of these 66 firms were then contacted by telephone by Accenture, our partnering international consulting firm. Accenture offered free consulting, funded by Stanford University and the World Bank, as part of a management research project. We paid for the consulting to ensure we controlled the intervention, so we could provide a homogeneous management treatment to all firms. We were concerned that if the firms made any co-payments they might have tried to direct the consulting, for example asking for help on marketing or finance.

Of this group of firms, 34 expressed an interest in the project and were given a follow-up visit and sent a personally signed letter from Stanford. Of the 34 firms, 17 agreed to commit senior management time to the consulting program.[11] We compared these program firms with the 49 non-program firms and found no significant differences in observables.[12]

The experimental firms have typically been in operation for 20 years and all are family-owned. They produce fabric for the domestic market, with many firms also exporting, primarily to the Middle East. Table 1 reports some summary statistics for the textile manufacturing parts of these firms (many of the firms have other businesses in textile processing, retail and real estate). On average these firms had about 270 employees, current assets of \$13 million and sales of \$7.5m a year. Compared to US manufacturing firms these firms would be in the top 2% by employment and the top 5% by sales,[13] and compared to India manufacturing in the top 1% by both employment and sales (Hsieh and Klenow, 2010). Hence, by this criterion, as well as by most formal definitions, these are large manufacturing firms.[14]

These firms are complex organizations, with a median of 2 plants per firm and 4 reporting levels from the shop-floor to the managing director. In all the firms, the managing director is the single-largest shareholder, reflecting the lack of separation of ownership and management in many Indian firms. All directors are family members. One of the firms is publicly quoted on the Mumbai Stock Exchange, although more than 50% of the equity is held by the managing director and his father.

In exhibits (1) to (7) we include a set of photographs of the plants. These are included to provide some background information to readers on their size, production process and initial state of management. As is clear these are large establishments (Exhibit 1), with several multi-story buildings per site, and typically two production sites per firm, plus a head office in Mumbai. They operate a continuous production process that runs constantly (Exhibit 2). Their factories' floors were (initially) often rather disorganized (Exhibits 3 and 4), and their yarn and spare-parts inventory stores lacking any formalized storage systems (Exhibits 5 and 6).

---

[11] The two main reasons for refusing free consulting given on the telephone and during the visits was that the firms did not believe they needed management assistance or that it required too much time from their senior management (1 day a week). But it is also possible the real reason is these firms were suspicious of this offer, given many firms in India have tax and regulatory irregularities.

[12] For example, the program firms had slightly less assets (\$12.8m) compared to the non-program firms (\$13.9m), but this difference was not statistically significant (p-value 0.841). We also compared the two groups of firms on management practices using the BVR scores, and found they were almost identical (difference of 0.031, p-value 0.859).

[13] Dunn & Bradstreet (August 2009) lists 778,000 manufacturing firms in the US with only 17,300 of these (2.2%) with 270 or more employees and only 28,900 (3.7%) with \$7.5m or more sales.

[14] Most international agencies define large firms as those with more than 250+ employees.

III. THE MANAGEMENT INTERVENTION

III.A. Why use management consulting as an intervention

The field experiment aimed to improve management practices in the treatment plants. To achieve this we hired a management consultancy firm to work with our treatment plants as the operationally easiest way to rapidly change plant level management. We selected the consulting firm using an open tender. The winner was Accenture, a large international management, IT and outsourcing consultancy. It is headquartered in the U.S., but has about 40,000 employees in India. The full-time team of (up to) 6 consultants working on the project at any time all came from the Mumbai office. These consultants were educated at top Indian business and engineering schools, and most of them had prior experience working with US and European multinationals.

Selecting a high profile international consulting firm substantially increased the cost of the project. But it meant that our experimental firms were more prepared to trust the consultants, which was important for getting a representative sample group. It also offered the largest potential to improve the management practices of the firms in our study.

The project ran from August 2008 until August 2010, and the total cost of this was US$1.3 million, or approximately $75k per treatment plant and $20k per control plant.[15] Note this is very different from what the firms themselves would pay for comparable consulting, which would be probably $500k or more. The reason for our much cheaper costs per plant is that, because it was a research project, Accenture charged us pro-bono rates (50% of commercial rates) and provided partners' time for free, As well, it realized economies of scale in working across multiple similar plants.

While the intervention offered high-quality management consulting services, the purpose of our study was to use the improvements in management generated by this intervention to understand if (and how) modern management practices affect firm performance. Like many recent development projects, this intervention was provided as a mechanism of convenience – to change management practices – and not to evaluate the management consultants themselves.

III.B. The management consulting intervention

The intervention aimed to introduce a set of standard management practices. Based on their prior textile experience, the consultants identified 38 key practices on which to focus. These 38 practices encompass a range of basic manufacturing principles that are standard in almost all US, European and Japanese firms, and can be grouped into five areas:

---

[15] These costs may seem high in the Indian context. Certainly, at the bottom of the consulting quality distribution in India consultants are extremely cheap, but their quality is very poor. At the top end, rates are more comparable to those in the US and Europe (Accenture's India rates are about a third of its US rates). This is because the international consulting companies target large domestic firms and multinationals and because the consultants these firms employ are often US or European educated and have access to international labor markets. In fact 2 of our consultants had previously worked in the US and returned to India for family reasons.

- Factory Operations and Planning: Plants were encouraged to undertake regular maintenance of machines (rather than repairing machines only when they broke down) and record the reasons for machine downtime so they could learn from past failures to reduce future downtime. They were encouraged to keep the factory floor tidy and organized in order to reduce accidents and facilitate the movement of materials. Spare parts were to be organized and inventory levels controlled, to ensure they were available to reduce machine downtime.

- Quality control: Plants were encouraged to record quality defects by type at every stage of the production process, analyze these records daily, and formalize procedures to address defects to prevent them occurring repeatedly.

- Inventory: Plants were encouraged to record yarn stocks, on a daily basis, with optimal inventory levels being defined and stock monitored against these. Yarn was to be sorted, labeled and stored in the warehouse by type and color, and this information logged onto a computer so yarn could be located when required for production.

- Human-resource management: Plants were encouraged to introduce a performance-based incentive system for workers and managers. Worker incentives were to be linked to output, quality and attendance (to reduce absenteeism). Job descriptions were to be defined for all workers and managers to improve clarity on roles and responsibilities.

- Sales and order management: Plants were encouraged to track production on an order-wise basis to prioritize customer orders by delivery deadline. Design-wise efficiency analysis was suggested so that pricing could be based on design-wise (rather than average) production costs.

These 38 management practices (listed in Appendix Table A1) form a set of precisely defined binary indicators that we can use to measure improvements in management practices as a result of the consulting intervention.[16] We recorded these indicators on an on-going basis throughout the study. 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. Thus, while 93 percent of the treatment plants recorded quality defects before the intervention, only 29 percent monitored them on a daily basis or by the particular sort of defect, and none of them had any standardized analysis and action plan based on this defect data.

The consulting treatment had three stages. The first stage, 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. Construction of this database involved setting up processes for measuring a range of plant-level metrics – such as output, efficiency, quality, inventory and energy use – on an ongoing basis, plus developing and recording historical data from existing plant records. For example, to facilitate quality

[16] We prefer these indicators to the BVR management score for our work here, since they are all binary indicators of specific practices, which are directly linked to the intervention. In contrast, the BVR indicator measures practices at a more general level on a 5-point ordinal scale. Nonetheless, the sum of our 38 pre-intervention management practice scores is correlated with the BVR score at 0.404 (p-value of 0.077) across the 17 firms.

monitoring on a daily basis, a single metric, termed the Quality Defects Index (QDI), was defined. The QDI is a severity-weighted average of the major types of defects. .

At the end of the diagnostic phase the consulting firm provided each treatment and control plant with a detailed analysis of its current management practices and performance. The treatment plants were given this diagnostic phase as the first step in improving their management practices. The control plants were given it only because we needed to construct historical performance data for them and help set up systems to generate ongoing data.

The second phase was a four month *implementation* phase which was given only to the treatment plants. In this the consulting firm followed up on the diagnostic report to help introduce as many of the 38 key 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 they could be readily carried out by employees. For example, one of the practices implemented was daily meetings for management to review production and quality data. The consultant attended these meetings for the first few weeks of the implementation phase to help the managers run them, provided feedback on how to run future meetings, and fine-tuned their design to the specific plant's needs.

The third phase was a *measurement* phase which lasted until August 2010. This phase involved only three consultants and a part-time manager, who collected performance and management data from the plants. In return for continuing to provide this data the consultants continued to provide some light consulting advice to both the treatment and control plants.

So, in summary, the control plants were provided with just the diagnostic phase and then the measurement phase (totaling 225 consultant hours on average), while the treatment plants were provided with the diagnostic, then implementation and then measurement phases (totaling 733 consultant hours on average).

III.C. The experimental design

We wanted to work with large firms because their complexity means management practices are likely to be important. However, providing consulting to large firms is expensive, which necessitated a number of trade-offs which are detailed below.

Sample size:
We worked with the 28 plants within our 17 experimental firms. We considered hiring cheaper local consultants and providing more limited consulting to a sample of several hundred plants. But two factors pushed against this. First, many large firms in India are reluctant to let outsiders into their plants because of their lack of compliance with tax, labor and safety regulations. To minimize selection bias we offered a high quality intensive consulting intervention that firms would value enough to take the risk of allowing outsiders into their plants. This helped maximize initial take-up (26% as noted in section II.B) and retention (100%, as no firms dropped out). Second, the consensus from discussions with Indian business people was that achieving a measurable impact in large firms would require an extended engagement with high-quality

consultants. Obviously the trade-off was this led to a small sample size,. We discuss the estimation issues this generates in section III.D below.

On-site and off-site plants: Due to manpower constraints we could collect detailed performance data from only 20 plants. Building data collection systems and compiling historic databases required the consultants spending several hours each week on-site at each plant. As a result the performance regressions are run only on these 20 "experimental" plants. However, data on management, organizational and IT outcomes were gathered for all 28 plants, as this required only bi-monthly visits to the other eight. These eight plants are referred to as "non-experimental".

Treatment and control plants: Within the group of 20 experimental plants we randomly picked 6 control plants and 14 treatment plants. As Table 1 shows, the treatment and control firms were not statistically different across any of the characteristics we could observe.[17]

Timing: The consulting intervention had to be initiated in three batches because of the capacity constraint of the six person consulting team. So the first wave started in September 2008 with 4 treatment plants. In April 2009 a second wave of 10 treatment plants was initiated, and in July 2009 the third wave of 6 control plants was initiated. While this design raises some potential timing concerns, in all regressions we include a full set of weekly time dummies, and provide weekly performance figures (figures 3, 4 and 5).

We started with a small first wave because we expected the intervention process would more difficult initially than it would be later, due to experience. The second wave included all the remaining treatment firms because: (i) the consulting interventions take time to affect performance and we wanted the longest time-window to observe the treatment firms; and (ii) we could not mix the treatment and control firms across waves.[18] The third wave contained the control firms. We picked more treatment than control plants because the staggered initiation of the interventions meant the different treatment groups provided some cross identification for each other, and because we believed the treatment plants would be more useful for understanding why firms had not adopted management practices before.

III.D. Small sample size

The focus on large firms meant we had to work with a small sample of firms. This raises three broad concerns. A first potential concern is whether the sample size is too small to identify significant impacts of management. Second is whether testing the significance of any impact may be impossible because standard testing procedures use asymptotic distributions. Third, the sample may be too small to be representative of large firms in developing countries. We discuss each concern in turn and the steps we have taken to try and address some of these shortcomings.

[17] Treatment and control plants were never in the same firms. The 6 control plants were randomly selected first, and the 14 treatment firms then randomly selected from the remaining firms which did not have a control plant.

[18] Each wave had a one-day kick-off meeting with all the firms, involving presentations from a range of senior partners from the consulting firm. This helped impress the firms with the expertise of the consulting firm and highlighted the huge potential for improvements in management. This meeting involved a project outline, which was slightly different for the treatment and control firms because of the different interventions. Since we did not tell firms about the existence of treatment and control groups we could not mix the treatment and control groups.

Significance of results:

Even though we have only 20 experimental plants across 17 firms we obtain statistically significant results. There are five reasons for this. First, these are large plants with about 80 looms and about 130 employees each, so that idiosyncratic shocks – like machine breakdowns or worker illness – tend to average out.[19] Second, the data were collected directly from the machine logs, so have very little (if any) measurement error. Third, the firms are homogenous in terms of size, product, region and technology, so that most external shocks can be controlled for with time dummies. Fourth, we collected weekly data, which provides high-frequency observations over the course of the treatment. Finally, the intervention was intensive, leading to large treatment effects – for example, the point estimate for the reduction in quality defects was over 50%.

Statistical inference:

A second concern is over using standard statistical tests, which assume an asymptotically normal distribution of errors across firms (the N dimension). So while our main results use firm-clustered bootstrap standard errors (see Bertrand, Duflo and Mullainathan, 2004, and Cameron et al, 2008), we also implement size-robust permutation procedures. We also exploit the fact we have a large T sample and take asymptotics across time.

Permutation Tests: Permutation tests use the fact that order statistics are sufficient and complete statistics to derive critical values for test statistics.[20] We first implement this for the null hypothesis of no treatment effect against the two sided alternative for the Intent to Treat (ITT) parameter. This calculates the ITT coefficient for every possible combination of 11 treatment firms out of our 17 total firms (we run this as the firm rather than plant level to allow for firm-level correlations in errors). Once this is calculated for all possible 12,376 treatment assignments (17 choose 11) the 2.5% and 97.5% confidence intervals are calculated as the 2.5th and 97.5th percentiles of the treatment impact. A treatment effect outside these bounds can be said to be significant at the 5% level.

Permutation tests for the instrumental variables estimator are more complex. For these we implement an IV version of the permutation procedure based on the proposals in Greevy et al (2004), Imbens and Rosenbaum (2005) and Andrews and Marmer (2008). The basic idea is to consider an outcome Y, endogenous regressor D, a randomized treatment assignment Z, and the model $Y = \beta D + \epsilon$. Suppose that we are interested in testing the null hypothesis that $\beta = b$. Under the null hypothesis $Y - bD$ is independent of the random treatment assignment variable Z.[21] This suggests using measures of the dependence between Z and $Y - bD$ as the basis of a testing procedure. The problem is now reduced to testing for the independence between $Y - bD$ and Z, which again we can use a permutation distribution (see Appendix B for more details).

T-asymptotic clustered standard errors: An alternative approach is to use asymptotic estimators that exploit our large time dimension for each firm. To do this we use the recent results by

[19] To illustrate this note that, for example, we actually have more employees in our 20 plants than in the 391 micro-enterprises in De Mel et al. 2008.

[20] See Lehmann and Romano (2005) for a textbook exposition of permutation tests.

[21] If the null hypothesis is false then $Y - bD = Y - \beta D + (\beta - b)D = \epsilon + (\beta - b)D$ which will be correlated with Z since the endogenous regressor D is correlated with the instrument.

Ibramigov and Mueller (2009) (henceforth IM) to implement a t-statistic based estimator that is robust to substantial heterogeneity across firms as well as to considerable autocorrelation across observations within a firm. The IM approach requires estimating the parameter of interest separately for each firm and then treating the resultant set of 17 estimates as a draw from a t distribution with 16 degrees of freedom. IM show that such a procedure is valid in the sense of having correct size (for fixed N and large T) so long as the time dimension is large enough that the estimator for each firm can be treated as a draw from a normal distribution. In our application we have on average over 100 observations for each firm, so this requirement is likely to be met.

Representativeness of the sample:
A third concern with our small sample is whether it is representativeness of large firms in developing countries. In part this concern represents a general issue for field experiments, which are often run on individuals, villages or firms in particular regions or industries. In our situation we focus on one region and one industry, albeit India's commercial hub (Mumbai) and it's largest industry (textiles). Comparing our sample to the population of large (100 to 5000 employee) firms in India, both overall and in textiles, suggests that while our sample is small it is at least broadly representative in terms of management practices (see Figure 1). In section V we also report results on a plant by plant basis to further demonstrate these are not driven by any particular outlier amongst the treatment or control plants. As such while we have a small sample the results are relatively stable across the individual plants within the sample.

IV. THE IMPACT ON MANAGEMENT PRACTICES

In Figure 2 we plot the average management practice adoption of the 38 practices for the 14 treatment plants, the 6 control plants and the 8 non experimental plants. This data is shown at 2 month intervals before and after the diagnostic phase. Data from the diagnostic phase onwards was compiled from direct observation at the factory. Data from before the diagnostic phase was collected from detailed interviews of the plant management team based on any changes to management practices during the prior year. Figure 2 shows five key results:

First, the plants in all groups started off with low baseline adoption rates of the set of 38 management practices.[22] Among the 28 individual plants the initial adoption rates varied from a low of 7.9% to a high of 55.3%, so that even the best managed plant in the group had in place just over half of the 38 key textile manufacturing management practices. This is consistent with the results on poor general management practices in Indian firms shown in Figure 1. For example, many of the plants did not have any formalized system for recording or improving production quality, which meant that the same quality defect could arise repeatedly. Most of the plants also had not organized their yarn inventories, so that yarn stores were mixed by color and type, without labeling or computerized entry. The production floor was often blocked by waste, tools and machinery, impeding the flow of workers and materials around the factory.

[22] The difference between the treatment, control and other plant groups is not statistically significant, with a p-value on the difference of 0.248 (see Table A1).

Second, the intervention did succeed in changing management practices. The treatment plants increased their use of the 38 management practices over the period by 37.8 percentage points on average (an increase from 25.6% to 63.4% of practices implemented).

Third, the treatment plants' adoption of management practices occurred gradually. In large part this reflects the time it takes for the consulting firm to gain the confidence of the firm's directors. Initially many directors were skeptical about the suggested management changes, and they often started by piloting the easiest changes around quality and inventory. Once these started to generate substantial improvements, these changes were rolled out and the firms started to consider introducing the more complex improvements around operations and HR.

Fourth, the control plants, which were given only the 1 month diagnostic, increased their adoption of these management practices, but by only 12% on average. This is substantially less than the increase in adoption in the treatment wave firms, indicating that the four months of the implementation the treatment plants received was important in changing management practices[23]. The control firms typically did not adopt the more complex practices like daily quality meetings, formalizing the yarn monitoring process or defining roles and responsibilities.

Fifth, the non experimental plants also saw a substantial increase in the adoption of management practices. In these 8 plants the management adoption rates increased by 11%. Most of this increase was driven by the 5 non experimental plants which were in the same firm as the treatment plants, which increased the adoption of practices by 17.5%, compared with the 3 non experimental plants which were in the same firm as the control plants which increasing their adoption by just 1%. This spillover of management practices was driven by the directors copying the new management practices from their experimental plants to their other plants.

IV.A. Management practice spillovers across plants within firms

To test formally whether the intervention has differentially changed management practices between the treatment and control plants we run the following plant-level panel regression:

$$MANAGEMENT_{i,t} = \alpha_i + \beta_t + \lambda_1 OWN\_TREAT_{i,t} + \lambda_2 SPILLOVER\_TREAT_{i,t} + \varepsilon_{i,t} \qquad (1)$$

where $MANAGEMENT_{it}$ is the fraction of the 38 practices adopted by plant I at date t, α_i are plant fixed effects, β_t are calendar month fixed effects, $OWN\_TREAT_{i,t}$ = log(1+cumulative months since the implementation phase began), and $SPILLOVER\_TREAT_{i,t}$ = log(1+cumulative months since implementation began in all other plants in the same firm). We use this logarithmic functional form because of concave adoption path of management practices shown in Figure 2. The parameter λ_1 estimates the semi-elasticity of the plants management practices with respect to the months of their own on-site consulting, while λ_2 estimates the semi-elasticity of spillovers from on-site consulting plants in other plants within the firm. The standard errors are bootstrap clustered by firm.

[23] This is very much consistent with the Nelson and Winter's (1982) discussions of tacit knowledge and the development and transferability of routines.

The results are shown in Table 2. We report in column (1) that management practices significantly respond to own plant treatment, rising by about 0.109 for every unit change in log(1+months treatment). We also see a response of 0.041 to log(1+months treatment) in other plants within the same firm. This coefficient is just over one third of the magnitude of the direct impact, suggesting substantial spillovers of management practices across plants within the same firm. In column (2) we add the three month lagged spillover term to investigate the timing of any potential spillover and find the lagged term dominates. This is consistent with a delay in transferring management practices across plants.[24] This arises because the firms' directors typically evaluate the impact of the new management practices in their treatment plants before transferring these over to their other plants. In column (3) we use just the three month lag and find a coefficient of 0.047, at almost half the direct effect. In columns (4) and (5) we see that using an even longer six-month lags leads to similar coefficient of 0.050. So whatever the exact specification, this data provides evidence of gradual spillovers of better management practices across plants within firms. Importantly for our study, these results also show that the experiment differentially changed management practices between treatment and control plants, providing variation which we can use to examine the impacts of management on performance.

V. THE IMPACT OF MANAGEMENT ON PERFORMANCE

The previous literature has shown a strong correlations between management practices and firm performance in the cross-section, with a few papers (e.g. Ichniowski et al. 1998) also showing this in the panel.[25] Our unique panel data on management practices and plant level performance, coupled with the experiment, enables us to confirm to what extent these results are causal.

We begin with a panel fixed-effects specification:

$$OUTCOME_{i,t} = \alpha_i + \beta_t + \theta MANAGEMENT_{i,t} + \nu_{i,t} \qquad (2)$$

where outcome will be on of the three key performance metrics of quality, inventory and output. The concern is of course that management practices are not exogenous to the outcomes that are being assessed, even in changes. For example, a firm may only start monitoring quality when it is starting to experience a larger than usual number of defects, which would bias the fixed-effect estimate towards finding a negative effect of better management on quality. Or firms may start monitoring quality as part of a major upgrade in worker quality and equipment, in which case we would misattribute quality improvements from better capital and labor to better management.

To overcome this endogeneity problem, we instrument the management practice score with log(1+weeks of treatment). The exclusion restriction is that the intervention affected only the outcome of interest through its impact on management practices, and not through any other channel. A justification for this assumption is that the consulting firm focused entirely on the 38

[24] In comparison the 3 month lagged own plant consulting term was actually negative and insignificant when added, so that the cross-plant spillovers is delayed while the within plant treatment effect seems to happen instantly.

[25] Note that other papers using repeated surveys have found no significant panel linkage between management practices and performance (Cappelli and Neumark (2001) and Black and Lynch (2004)), probably because of measurement error issues with repeated surveys.

management practices in their recommendations to firms, and firms did not buy new equipment or hire new labor as a result of the intervention (at least in the short run).[26] The IV estimator will then allow us to answer the headline question of this paper – does management matter?

If the impact of management practices on plant-level outcomes is the same for all plants, then the IV estimator will provide a consistent estimate of the marginal effect of improvements in management practices, telling us how much management matters for the average firm participating in the study. However, if the effects of better management are heterogeneous, then the IV estimator will provide a local average treatment effect (LATE). The LATE will then give the average treatment effect for plants which do change their management practices when offered free consulting. If plants which stand to gain more from improving management are the ones who change their management practices most as a result of the consulting, then the LATE will exceed the average marginal return to management. It will understate the average return to management if instead the plants that change management only when free consulting is provided fare those with the least to gain.

There was heterogeneity in the extent to which treatment plants changed their practices, with the before-after change in average total management practice score ranging from 26.3% to 60%. The feedback from the consulting firm was that to some extent it was firms with the most unengaged, uncooperative managers who changed practices least, suggesting that the LATE may underestimate the average impact of better management if these firms have the largest potential gains from better management. Nonetheless, we believe the LATE estimate to be a parameter of policy interest, since if governments are to employ policies to try and improve management, information on the returns to better management from those who actually change management practices when help is offered is informative.

We can also directly estimate the impact of the consulting services intervention on management practices via the following equation:

$$OUTCOME_{i,t} = a_i + b_t + cTREAT_{i,t} + e_{i,t} \qquad (3)$$

Where $TREAT_{i,t}$ is a 1/0 variable for whether plants have started the implementation phase or not. The parameter *c* then gives the intention to treat effect (ITT), which is the average impact of the intervention in the treated plants compared to the control plants.

In all cases we include plant and time fixed effects, and we bootstrap cluster the standard errors at the firm level. We have daily data on many outcomes, but aggregate them to the weekly level to reduce higher-frequency measurement errors.

V.A Quality

[26] The exceptions to this were that the firms hired on average $34 (1,700 rupees) of extra manual labor to help organize the stock rooms and clear the factory floor, spent $418 (10,900 rupees) on plastic display boards for the factory floor, standard-operating procedure notices and racking for the store rooms, spent an additional $800 on salary and prizes (like a radio and a watch) for managerial and non managerial staff and about $1000 each on new computers (discussed in section VI below). These and any other incidental expenditures are too small to have a material impact on our profitability and productivity calculations.

Our measure of quality is the Quality Defects Index (QDI), a weighted average score of quality defects, which is available for all but one of the plants. Higher scores imply more defects. Figure 3 provides a plot of the QDI score for the treatment and control plants relative to the start of the treatment period. This is September 2008 for Wave 1 treatment, April 2009 for Wave 2 treatment and control plants.[27] This is normalized to 100 for both groups of plants using pre-treatment data. To generate point-wise confidence intervals we block bootstrapped over firms.

It is very clear the treatment plants started to reduce their QDI scores (i.e. improve quality) significantly and rapidly from about week 5 onwards, which was the beginning of the implementation phase following the initial 1 month diagnostic phase. The control firms also showed a mild downward trend in their QDI scores from about week 30 onwards, consistent with their slower take-up of these practices in the absence of a formal implementation phase. These differences in trends between the treatment and control plants are also significant, as indicated by the non-overlapping 95% confidence intervals towards the end of the period.

Table 3 in columns (1) to (3) examines whether management practices improve quality using a regression approach. In column (1) we present the fixed-effects OLS results which regresses the monthly log(Quality Defects Index) score on plant level management practices, plant fixed effects, and a set of monthly time dummies. The standard errors are bootstrap clustered at the firm level to allow for any potential correlation across different experimental plants within the same firm. The coefficient of -0.561 implies that increasing the adoption of management practices by 10 percentage points would be associated with a reduction of 5.61% in the quality defects index.

In Table 5, column (2), we instrument management practices using the experimental intervention to identify the causal impact of better management on quality. After doing this we see a significant point estimate of -2.028, suggesting that increasing the management practice adoption rate by 10% would be associated with a reduction in quality defects of 20.3%. The large rise in the point estimate from the OLS to the IV estimator suggests firms are endogenously adopting better management practices when their quality starts to deteriorate. There was some anecdotal evidence for the latter, in that the consulting firm reported some plants with improving quality were less keen to implement the new management practices because they felt these were unnecessary. This suggests that the fixed-effects estimates for management and performance in prior work like Ichniowski, Prennushi and Shaw (1997) may be substantially underestimating the true impact of management on performance.

The reason for this large effect is that measuring defects allows firms to address quality problems rapidly. For example, a faulty loom that creates weaving errors would be picked up in the daily QDI score and dealt with in the next day's quality meeting. Without this, the problem would often persist for several weeks since the checking and mending team had no system (or incentive) for resolving defects. In the longer term the QDI also allows managers to identify the largest sources of quality defects by type, design, yarn, loom and weaver, and start to address these systematically. For example, designs with complex stitching that generate large numbers of quality defects can be dropped from the sales catalogue. This ability to improve quality

[27] Since the control plants have no treatment period we set their timing to zero to coincide with the 10 Wave 2 treatment plants. This maximizes the overlap of the data.

dramatically through systematic data collection and evaluation is a key element of the successful lean manufacturing system of production (see, for example, Womack, Jones and Roos, 1992).

Finally, in column (3) we look at the intention to treat (ITT), which is the average reduction in the quality defects index in the period after the intervention in the treatment plants versus the control plants. We see this is associated with a 32% (exp(-.386)-1) reduction in the QDI index.

At the foot of table 5 we also present our Ibragimov-Mueller (IM) and permutation alternative significance tests. First, looking at the IM tests that exploit asymptotics in T rather than N , we find that the IV and ITT results are both significant at the 5% level (zero is outside the 95% confidence intervals). For the standard permutation tests the ITT is again significant at the 5% level (the p-value is 0.0168), while for the IV-permutation tests the estimate is significant at the 10% level although not at the 5% level. So overall the small sample statistical tests also find the IV and ITT effects of management on quality significant at the 10% level and usually also the 5% level.

V.B Inventory

Figure 4 shows the plot of inventory levels over time for the treatment and control groups. It is clear that after the intervention the inventory levels in the treatment group fall relative to the control group, with this being point-wise significant by about 30 weeks after the intervention.

The reason for this effect is that these firms were carrying about 4 months of inventory on average before the intervention, including a large amount of dead-stock. Often, because of poor records and storage practices, firms did not even know they had these stocks. By cataloguing the yarn and sending the shade-cards to the design team to include in new products[28], selling dead yarn stock, introducing restocking norms for future purchases, and monitoring inventory on a daily basis, the firms dramatically reduced their inventories. But this took time as the reduction in inventories primarily arose from lowering stocking norms and consuming old yarn into new products.

Table 5 columns (4) to (6) shows the regression results for raw material (yarn) inventory. In all columns the dependent variable is the log of raw materials, so the coefficients can be interpreted as the percentage reduction in yarn inventory. The results are presented for the 18 plants for which we have yarn inventory data (two plants do not maintain yarn stocks on site). In column (4) we present the fixed-effects results which regresses the monthly yarn on the plant level management practices, plant fixed effects, and a set of monthly time dummies. The coefficient of -0.639 says that increasing management practices adoption rates by 10 percentage points would be associated with a yarn inventory reduction of about 6.39%. In Table 5, column (5s), we see the impact of management instrumented with the intervention displays a point estimate of -0.929, again somewhat higher than the FE estimates in column (1). Again, the IV estimator is higher than the OLS estimator, suggesting that the adoption of better management practices may be endogenous, driven in part by poor inventory performance. In column (6) we see the intervention is associated with an average reduction in yarn inventory of (exp(-.179)-1=) 16.4%.

---

[28] Shade cards comprise a few inches of sample yarn, plus information on its color, thickness and material. These are sent to the design teams in Mumbai who use these to design new products using the surplus yarn.

These numbers are substantial, but in fact US automotive firms achieved much greater reductions in inventory levels (as well as quality improvements) by adopting Japanese lean manufacturing technology beginning in the 1980s. Many firms reduced inventory levels from several months to a few hours by moving to just-in-time production (Womack, Jones and Roos, 1991).

Finally, as with the quality defects estimates the IM confidence intervals for the IV estimator find the coefficient significance at the 5% level. However, the IV permutation tests can not exclude zero so are not finding significance at standard levels. Looking at the ITT coefficient we see that under IM the results are significant at the 10% level although again not significant using the standard permutation tests.

V.C Output

In Figure 5 we plot output over time for the treatment and control plants. Output is measured in physical terms, as production picks. The results here are less striking, although output of the treatment plants has clearly risen on average relative to the control firms, and this difference is point-wise statistically significant in some weeks towards the end of the period.

In columns (7) to (9) in table 5 we look at this in a regression setting with plant and time dummies. In column (7) we see that for the OLS specification increasing the adoption of management practices by 10 percentage points would be associated with a 1.27% increase in efficiency. In column (8), we see the impact of management instrumented with the intervention displays a higher significant point estimate of 0.346, suggesting a 10% increase in management adoption would lead to a 3.46% increase in output. As with quality and inventory the IV estimator is again notably higher than the OLS estimator, again indicating an endogenous adoption of better management when output falls.

Finally, in column (9) we look at the intention to treat (ITT) and see a point estimate of 0.056, implying a 5.4% increase in output (exp(0.056)-1), although this only significant at the 11% level. Looking at the small-sample standard errors we find both the IM and permutation tests are, however, significant at the 5% level for both the IV and ITT estimates.

There are several reasons for these increases in output. Undertaking routine maintenance of the looms, especially following the manufacturers' instructions, reduces breakdowns that stop production. Collecting and monitoring the breakdown data also helps highlight looms, shifts, designs and yarn-types that are associated with more breakdowns, facilitating pro-actively addressing these. And visual displays around the factory floor together with the incentive schemes based on these performance metrics motivate workers to improve operating efficiency. Finally, keeping the factory floor clean and tidy reduces the number of accidents, reducing incidents like tools falling into machines or fires damaging equipment. Again the experience from lean manufacturing is that the collective impact of these procedures can lead to extremely large improvements in operating efficiency, raising output levels.

V.D Are the improvements in performance due to Hawthorne effects?

Hawthorne effects are named after the experiments carried out by industrial engineers in the Hawthorne Works in the 1920s and 1930s which attempted to raise productivity. The results apparently showed that simply running experiments led to an improvement in performance, with the most cited result being that both reducing and increasing ambient light levels led to higher productivity. While these putative Hawthorne effects in the original experiments have long been disputed (e.g. Levitt and List, 2009), there raise a serious concern that some form of the Hawthorne effect is causing our observed increase in plant performance.

However, we think this is unlikely, for a series of reasons. First, our control plants also had the consultants on site over a similar period of time as the treatment firms. Both sets of plants got the initial diagnostic period and the follow-up measurement period, with the only difference being the treatment plants also got an intensive intermediate 4 month implementation stage while the control plants had briefer, but frequent, visits from the consultants collecting data. The control plants were not told they were in the control group. Hence, it cannot be simply the presence of the consultants or the measurement of performance that generated the improvement in performance. Second, the improvements in performance took time to arise and they arose in quality, inventory and efficiency, where the majority of the management changes took place. Third, these improvements persisted for many months after the implementation period, so are not some temporary phenomena due to increased attention. Finally, the firms themselves also believed these improvements arose from better management practices, which was the motivation for them spreading these practices out to their other plants not involved in the experiments.

VI. IMPACT OF MANAGEMENT ON ORGANIZATIONAL STRUCTURE AND COMPUTERIZATION

V.A The impact of management practices on firm organization

Although our interventions were never intended to change the treatment firms' organizational designs, theory gave us some reason to believe that organizational changes might follow from increasing the amount and quality of information available to decision makers. In recent years a large theoretical literature on the economics of organization has developed dealing with the locus of decision-making within firms.[29] However, this literature does not lead to clear-cut predictions about the effects of increased availability of information to top executives and lower-level managers. On the one hand, models like Garicano (2000) of hierarchy as a specialization in knowledge acquisition would suggest that more decisions ought to be taken at lower levels if the amount of information available to all levels is increased. Similarly, a standard agency perspective might also suggest that more decisions would be delegated if new or more accurate performance measures become available, especially if (as in our sample) the directors are under significant time constraints. However, to the extent that the plant managers were initially better informed than their bosses by virtue of being closer to the operations, the availability of the better measures might have reduced their information advantage, favoring the directors' making more decisions. While the theoretical literature is expansive, the empirical literature is extremely limited, so we collected extensive data on the locus of decision making in our firms.

[29] For recent surveys see Garicano and Van Zandt (2011), Mookherjee (2011), and Bolton and Dewatripont (2011).

To measure decentralization we collected data on eight variables: the locus of decision-making for weaver hiring, manager hiring, spares purchases, maintenance planning, weaver bonuses, investment, and departmental co-ordination, and the number of days per week the owner spends at the factory. Because firms' organizational designs change slowly over time, we collected this data at lower frequencies – pre-intervention, in March 2010 and in August 2010. For every decision except investment and days at the factory we scored decentralization on a 1 to 5 scale, where 1 was defined as no authority of the plant manager over the decision and 5 as full authority (see Appendix Table A2 for the full survey and Table A4 for descriptive statistics). So, for example, we measured decentralization for the plant manager over weaver hiring from a scale of 1 defined as "No authority – not even for replacement hires" to 5 defined as "Complete authority – it is his decision entirely", with intermediate scores like 3 defined as "Requires sign-off from the Director based on the business case. Typically agreed about 80% or 90% of the time". These questions and scoring were based on the survey methodology in Bloom, Sadun and Van Reenen (2009), which measured decentralization across countries and found developing countries like India, China and Brazil typically have very centralized decision-making within firms. The measure of the decentralization for investment was in terms of "The largest expenditure (in rupees) a plant manager (or other managers) could typically make without a Directors signature", which had an average of 12,608 rupees (about \$250). Finally, the number of days the owners spend each week at the factory is a revealed preference measure of decentralization. The owners are usually located either at their head-offices in Mumbai (which they prefer as it reduces their commute) or at the factory (if this needs more direct management).

To combine all these eight decentralization measures into one index we took the principal factor component of the measures, which we called the decentralization index. Changes in this index were strongly and significantly correlated with changes in management across firms. Firms which had substantial improvements in management practices during the experiment also tended to delegate more production decisions to their plant managers.

Table 4 looks at this in a regression format by estimating the following specification

$$\text{DECENTRALIZATION}_{i,t} = a_i + b_t + c\text{MANAGEMENT}_{i,t} + e_{i,t} \qquad (3)$$

where DECENTRALIZATION is our index of plant decentralization, and a_i and b_t are plant fixed effects and time dummies. In column (1) we run the OLS estimation and find a significant and positive coefficient, indicating that firms which improved their management practices during the experiment have also delegated more decisions to their plant managers. Given that the decentralization index has a standard deviation of 1 the magnitude of this coefficient is large – increasing the adoption of management practices by 37.8% (the mean change for the treatment group) is associated with a 0.55 standard-deviation change increase in decentralization. In column (2) we run the IV estimation, using the log(1+weeks treatment) as the instrument, and again find a positive and significant impact. Finally, in column (3) we report a positive intention to treat impact.

Our evidence from direct discussion with the owners is that better management leads to decentralization because it increases their trust in their plant managers. In large part this is because the improved monitoring of the factory operations allows them to delegate more decision making without fear of being exploited (the monitoring channel in our principal-agent group of organizational theories). For example, with daily inventory, quality and output data it is harder for the factory manager to appropriate inventory or output without detection. A second channel seems to be that better management practices enables the plant managers to more effectively run the factory without assistance from the owner. For example, after introducing daily factory meetings the owners told-us they felt less obliged to visit the factories daily to oversee production (the informational channel in our first group of theories).

It is worth noting that even these decentralizing Indian factories are still extremely centralized compared to factories in Europe and the US. For example, using the Bloom, Sadun and Van Reenen (2009) data we know that plant managers in developed countries are typically able to hire full-time employees with pretty minimal control from their headquarters (compared to very limited authority in our Indian factories) and can invest about $52,000 without central clearance (compared to about $250 in India). So, these improvements in management practices have increased delegation but still leave Indian factories very centralized compared to plants in developed countries.

VI.B The impact of management practices on computerization

One of the major topics over the last decade has been the relationship between IT and productivity. Until the 1990s, convincing evidence on the aggregate impact of computers on productivity was so hard to find that Robert Solow famously quipped in 1987 that "you see computers everywhere but in the productivity statistics". In more recent periods, however, the paradox has reversed, with a growing literature now finding that the productivity impact of IT is substantially larger than its cost share (e.g. Bresnahan, Brynjolfsson and Hitt, 2002, and Brynjolfsson and Hitt, 2003). The literature has argued this is because IT is complementary with modern management and organizational practices, so that as firms invest in IT they also improve their management practices. This leads to a positive bias on IT in productivity estimates as management and organizational practices are typically an unmeasured residual.[30] But none of this literature has any direct experimental evidence, instead relying on identification from observed changes in IT and management and organizational survey data.

A second related IT literature has argued that skill biased technical change (SBTC) has been the major factor driving the increase in income inequality observed in the US and most other countries since the 1970s (see surveys in Acemoglu 2002 and Autor, Katz and Kearney 2008). But SBTC is usually inferred as the residual in inequality regressions, with rather limited direct evidence on specific skill-biased technologies. Our experimental changes in management practices are clearly skilled-biased, in that computer users are highly skilled due to the need for literacy, numeracy and computer familiarity. As a result modern management practices are a skill-biased technology, whose adoption is driving both the increased use of computers and the demand for relatively skilled workers.

So to investigate the potential complementarity between IT and management practices we collected computerization data on nine aspects of the plants, covering the use of Electronic Resource Planning (ERP) systems, the number of computers, the age of the computers, the number of computer users, the total hours of computer use, the connection of the plant to the

[30] See, for example, Bartel, Ichniowski and Shaw (2007) and Bloom, Sadun and van Reenen (2009).

internet, the use of e-mail by the plant manager and the director, the existence of a firm website and the depth of computerization of production decisions (see Appendix Table A3 for the full survey and Table A4 for descriptive statistics). As with the organizational changes we collected this data once from before the intervention, in March 2010 and in August 2010. Even in the d-stata table A4 it is readily apparent that as firms adopted more modern management practices they significantly increased the computerization of their operations.

Table 4 looks at this in a regression form by estimating the following specification

$$\text{COMPUTERIZATION}_{i,t} = a_i + b_t + c\text{MANAGEMENT}_{i,t} + e_{i,t} \quad (4)$$

where COMPUTERIZATION is measured in terms of the number of computer users (in columns (4) to (6)) or in terms of the overall computerization index (in columns (7) to (9)).

In column (4) we see that the full adoption of all management practices is associated with an increase of 16.76 hours of computer user a week, an average rise of over 100% given the pre-sample mean was 13.66 hours per week. In columns (5) and (6) we report the IV and ITT estimates, which show a similar result. In columns (7) to (9) we report similar OLS, IV and ITT results for the computerization index, which is a broader measure of computer use, and again see highly significant increases from the management intervention.

For context we should note, however, that even after the intervention these firms had extremely low rates of computer use compared to firms in developed countries. For example, the median numbers of computers per employee in our Indian plant was 0.023 compared to 1.18 in similar sized US and European factories (see Bloom, Sadun and Van Reenen (2009)). And in US and European factories all plant managers have a personal e-mail, compared to just 25% in India. Our results suggest one reason for this low overall level of computer use in developing countries may be the lack of modern management practices which these technologies support.

VII. WHY DO BADLY MANAGED FIRMS EXIST?

Given the evidence in section (IV) on the substantial impact of better management practices on plants' quality, inventory and output, the obvious question is whether these management changes increased profitability, and if so why where these not introduced before.

VII.A. The estimated impact of management practices on profits and productivity

Profits:
Overall we estimate a total increase in profits of around \$228,000, with our calculations outlined in Table A5. Firms did not provide us with any profit and loss accounts, so we estimated this from the quality, inventory and efficiency improvements.[31] Our methodology is very simple: for

[31] We could obtain the public profit and loss accounts, but it was unclear how accurate these were, and they were not at the plant level. We did not ask firms for their private profit and loss accounts (if they even kept them) as they would have been likely to refuse, given their fears over the information leaking out to the Indian tax authorities.

example, if a given improvement in practices is estimated to reduce inventory stock by X tons of yarn, we map this into profits using conservative estimates of the cost of carrying X tons of yarn. Or if it reduces the numbers of hours required to mend defects we estimated this reduction in hours on the firms total wage bill. These estimates are medium-run because, for example, it will take a few months for the firms to reduce their mending manpower.

These estimates for increases in profits are potentially biased. They are a downward biased because we take firms' choice of capital, labor and product range as given. But in the long-run the firms can re-optimize, for example, with more machines per weaver if quality improves (as dealing with breakdowns is time consuming). Furthermore, many of the management practices are complementary, so they are much more effective when introduced jointly (e.g. Milgrom and Roberts, 1990). However, the intervention time-horizon was too short to change many of the complementary human-resource practices, so the full rewards would not be realized. The intervention was also narrow in focus in that other management practices around activities like finance, strategy, marketing and procurement were not been addressed. They are clearly upward biased if once the consultants leave the factory the firms backslide on the management changes.

To estimate the net increase in profit for these improvements in management practices we also need to calculate the costs of these changes (ignoring for now any costs of consulting). These costs were extremely small, averaging less than \$3000 per firm.[32] So in the absence of any costs of consulting to introduce these new management practices – which would have been substantial if firms had paid themselves – it would clearly be highly profitable to do so.

Productivity:

We estimate a total increase in productivity of 11.1%, detailed in Table A5. Our methodology is very simple, assuming a constant-returns-to-scale Cobb-Douglas production function:

$$Y=AL^{\alpha}K^{1-\alpha} \tag{1}$$

where Y is value-added (output – materials and energy costs), L is hours of work and K is the net capital stock. Using equation (1) we can back out changes in productivity after estimating changes in output and inputs. So, for example, reducing the yarn inventory by 16.4% lowers capital by 1.3% (yarn is 8% of the total capital stock), increasing productivity by 0.6% (capital has a factor share of 0.42).

Our estimated productivity impact will be subject to a number of biases similar to those discussed above for profitability. However, the relatively substantial 11% impact from this narrow short-run management experiment suggests that bad management does play a potentially important role in explaining the aggregate productivity gap between India and the US.

VII.B. Why are firms badly managed?

[32] The \$35 of extra labor to help organize the stock rooms and clear the factory floor, about \$200 on plastic display boards, about \$200 for extra racking for stores rooms, about \$1000 on rewards, and about \$1000 for extra computer equipment (this is bought second hand).

Given the evidence in section (VI.A) above on the large increase in profitability from the introduction of these modern management practices, the obvious question is: why had firms not already adopted them? To investigate this we asked our consultants to document every other month the reason for any non-adoption of the 38 practices in each plant. To do this consistently we developed a flow-chart (see Exhibit 7) which runs through a series of questions to understand the root cause for the non-adoption of each individual practice. They collected this data from extensive discussions with owners, managers and workers, plus their own observations from working daily in the plants.

As an example of how this flow chart works, imagine a plant that does not record quality defects. The consultant would first ask if there was some external constraint, like labor regulations, preventing this, which we found never to be the case.[33] They would then ask if the plant was aware of this practice, which in the example of quality recording systems typically was the case, as it is a well known practice. The consultants would then check if the plant could adopt the practice with the current staff and equipment, which again for quality recording systems was always true. Then they would ask if the owner believed it would be profitable to record quality defects, which was often the constraint on adopting this practice. The owner frequently argued their quality was so good they did not need to record quality defects. This view was mistaken, however, because while these plants' quality might have been good compared to other low-quality Indian textile plants, but by international standards their quality was very poor. So, as shown in Figure 3, when they did adopt basic quality control practices they substantially improved their production quality. So, in this case the reason for non-adoption would be "incorrect information" as we believed the owner had incorrect information on the cost-benefit calculation for quality control processes.

The overall results for non-adoption of management practices are tabulated in Table 5, for the treatment plants, control plants and the non-experimental plants. This is tabulated at two-month intervals starting the month before the intervention phase. The rows report the different reasons for non-adoption as a percentage of all practices. Looking across the table several results are apparent.

First, a major initial barrier to the adoption of these modern management practices was a lack of information about their existence. About 15% of practices were not adopted because the firms were simply not aware of them. These practices tended to be the more advanced practices of regular quality, efficiency and inventory review meetings, posting standard-operating procedures and visual aids around the factory, the use of historical efficiency data for design pricing, and scientific inventory methods. Many of these are derived from the Japanese inspired lean manufacturing revolution, and are common across Europe, Japan and the US but apparently have yet to permeate Indian manufacturing.

Second, another major initial barrier was incorrect information, in that firms may have heard of these practices but thought they did not apply profitably to them. For example, many of the firms were aware of preventive maintenance but few of them thought it was worth doing this. They

[33] This does not mean labor regulations do not matter for some practices – for example firing underperforming employees – but they did not directly impinge adopt the immediate adoption of the 38 practices.

preferred to keep their machines in operation until they broke down, and then repair them. This accounted for slightly over 45% of the initial non-adoption of practices.

Third, as the intervention progressed the lack of information constraint was rapidly addressed. It was easy for the consultants to inform the firms about modern management practices. However, the incorrect information constraints were harder to address. This was because the owners had their prior beliefs about the efficacy of a practice and it took time to change these. This was often done using pilot changes on a few machines in the plant or with evidence from other plants in the experiment. For example, the consultants typically started by persuading the managers to undertake preventive maintenance on a set of trial machines, and once it was proven successful it was rolled out to the rest of the factory. And as the consultants demonstrated the positive impact of some of these initial practice changes, the owners increasingly trusted them and would adopt more of the recommendations, like introducing performance incentives for managers.[34]

Fourth, once the informational constraints were addressed, other constraints arose. For example, even if the owners became convinced of the need to adopt a practice, they would often take several months to adopt it. This was particularly pertinent in the non-experimental plants, where the consultants were not on-site to drive the changes. A major reason is the owners were severely time constrained, working an average of 68 hours per week already. There was also evidence on procrastination in that some owners would defer on taking quick decisions. This matches up with the evidence on procrastination in other development contexts, for example African farmers investing in fertilizer (Duflo, Kremer and Robinson, 2009) or growing pineapples (Conley and Udry, 2010).

Finally, somewhat surprisingly, we did not find evidence for the direct impact of a set of other factors that have been highlighted in the literature on capital investment. One such factor is capital constraints, which are a significant obstacle to the expansion of micro-enterprises (e.g. De Mel et al, 2008). Our evidence suggested that the relatively large firms that were involved in our experiment were not cash-constrained. We collected data on all the investments for our 17 firms over the period April 2008 until April 2010 and found the firms invested a mean (median) of $880,000 ($140,000). For example, several of the firms were setting up new factories or adding machines, apparently often financed by bank loans. Certainly, this scale of investment suggests that investment on the scale of $2000 (the first-year costs of these management changes, ignoring the consultants' fees) to improve the factories' management practices is unlikely to be directly impeded by financial constraints.

Of course financial constraints could impede hiring in international consultants. The market cost of our free consulting would be at least $500,000, and as an intangible investment it would be difficult to collateralize.[35] Hence, while financial constraints do not appear to directly block the

[34] These sticky priors highlight one reason why management practices appear to take several years to change in the US and Europe. The evidence on this is anecdotal, but for example, the private equity industry has a 3 year minimum estimate for the time needed for a management turnaround. Similarly, consulting firms typically take at least 18 months to execute large change management programs at their clients.

[35] Our international consulting firm estimated that to offer a standard consulting team to these firms at market rates would cost at least $500,000. This is much more expensive than our costs per firm because: (I) we achieved substantial scale economies from working with a large number of firms simultaneously; and (II) we had 50% rates on the consultants and no partner charges.

implantation of better management practices, they may hinder firms' ability to improve their current management practices using external consultants. On the other hand, our estimates of the incremental profitability from adopting modern management practices suggest cost recovery in as little as eighteen months.

VII.C. How do badly managed firms survive?

We have shown that management matters, with improvements in management practices improving plant-level outcomes. One response from economists might then be to argue that poor management can at most be a short-run problem, since in the long run better managed firms should take over the market. Yet many of our firms have been in business for 20 years and more.

One reason why better run firms do not dominate the market is constraints on growth through managerial span of control. In every firm in our sample, only members of the owning family are company directors – that is, in managerial positions with major decision-making power over finances, purchases, operations or employment. Non-family members are given junior managerial positions that have power only over low-level, day-to-day activities. The reason is the family members do not trust the non-family members not to steal from the firm. For example, they are concerned if they let their plant managers run procurement they might buy yarn at inflated rates from friends and receive kick-backs.

A key reason for this inability to decentralize is the poor rule of law in India. Even if directors found managers stealing, their ability to successfully prosecute them and recover the assets is minimal because of the inefficiency of Indian courts. In contrast, in the US if a manager was found stealing from a firm it is likely they could be successfully prosecuted and much of the assets recovered. A compounding reason for the inability to decentralize in Indian firms is bad management, as this means the owners cannot keep good track of materials and finance, so may not even able to identify theft within their firms.[36]

As a result of this inability to delegate, firms can expand beyond the size that can be managed by a single director only if male family members are available to share directorial duties. Thus, an important correlate of firm size in our firms was the number of male family members of the owners. For example, the number of brothers and sons of the leading director has a correlation of 0.689 with the total employment of the firm, compared to a correlation of 0.223 with their average management score across plants. In fact the best managed firm in our sample had only one (large) production plant, in large part because the owner had no brothers or sons to help run a larger organization. This matches the ideas of the Lucas (1978) span of control model, that there are diminishing returns to how much additional productivity better management technology can generate from a single manager. In the Lucas model, the limits to firm growth restrict the ability of highly productive firms to drive lower productivity ones from the market. In our Indian

[36] Another compounding factor is these firms had poor human resources management practices. None of the firms had a formalized development or training plan for their managers, and managers could not be promoted because only family members could become directors. As a result managers lacked career motivation within the firm. In contrast, in Indian software and finance firms place a huge emphasis on development and training to motivate employees, which is essential for delegation without a strong legal system (see also Banerjee and Duflo (2000)).

firms, this span of control restriction is definitely binding, so unproductive firms are able to survive because more productive firms cannot expand.

Entry of new firms into the industry also appears limited by several factors, several related to the difficulty of separating ownership from control. The supply of new firms is constrained by the number of families with finance and male family members available to build and run textiles plants. As well, the rapid growth of other industries in India – like software and real-estate – means the attractiveness of new investment in textile manufacturing is limited. (Even our firms were often taking cash from their textile businesses to invest in other businesses, like real-estate and retail.) Finally, even if an entrant had funding, given the informational problems identified earlier, there is no obvious guarantee its management practices would be much better than the incumbent firms.

Beyond all this, a 50% tariff on fabric imports from China insulates against foreign competition. Hence, the equilibrium appears to be that, with Indian wage rates are extremely low, firms can survive while operating with poor management practices. Because spans of control are constrained productive incumbent firms are limited from expanding and so do not drive out the badly run firms. And because entry is limited new firms do not enter rapidly. As a result the situation approximates a Melitz (2003) style model where firms have very high decreasing returns to scale, entry rates are low, and initial productivity draws are low (because good management practices are not widespread). The resultant equilibrium has a low average level of productivity, a low wage level, a low average firm-size, and a large dispersion of firm-level productivities.

VII.D. Why do firms not use more management consulting?

Finally, why do these firms not hire consultants themselves, given the large gains from better management? The primary reason is these firms are not aware they are badly managed, as illustrated in Table 5.

Of course consulting firms could still approach firms for business, pointing out that their practices were bad and offering to fix them. But Indian firms, much like US firms, are bombarded with solicitations from businesses offering to save them money on everything from telephone bills to raw materials, and so are unlikely to be particularly receptive. Of course consulting firms could go further and offer to provide their advice in return for an *ex post* profit-sharing deal. But monitoring this would be hard. First, how are the returns attributable to the consulting to be determined? Further, in the Indian context, firms very often conceal profits from the tax authorities and would be tempted to do the same with partnering consultants. Finally, the client in such an arrangement might worry that the consultant would twist its efforts to increase the short-term returns it shares at the expense of long-term results. On top of all these problems, many Indian firms are breaching tax, labor and health-and-safety laws (see Exhibits 3 to 7) and so are reluctant to let unknown outsiders into their firms. Our project benefited from the endorsement of Stanford and the World Bank, but a local firm offering free consulting would probably find it much harder to gain the trust of potential clients.

VIII. CONCLUSIONS

Management does matter. We have implemented a randomized experiment which gave managerial consulting services to textile plants in India. This experiment led to improvements in basic management practices, with plants adopting lean manufacturing techniques which have been standard for decades in the developed world. These improvements in management practice led to plants' improving the quality of their product, reducing inventory levels, and improving efficiency. The result was a significant improvement in profitability and productivity. Firms also delegated decisions more because the improved informational flow from the modern management practices enabled the owners to reduce their oversight of the plant. At the same time computer use increased substantially, driven by the need to collect, process and disseminate data as required by modern management practices.

What are the implications of this for public policy? First, our results suggest that firms were not implementing the better practices on their own because of lack of information and knowledge, and that to really improve management quality, firms needed persuasion that these practices would improve productivity and how they could implement them. This suggests a need for better benchmarking programs – to convince firms of the need to improve management practices – and knowledge and training programs in India and other developing countries more generally. This would include high quality educational programs to teach better management practices, and a more vibrant local consulting industry with the ability to signal quality through reputation building. While both these are private sector activities, they depend on the government for a regulatory environment which makes entry easy and which allows quality to be the main determinant of success.

A second method for knowledge transference comes from the presence of multinationals. Indeed, many of the consultants working for the international consulting firm hired by our project had worked for multinationals in India, learning from their state-of-the-art manufacturing management processes. Yet a variety of legal, institutional, and infrastructure barriers have limited the extent of multinational expansion within India, limiting the spread of knowledge on better manufacturing among the Indian managerial labor force. Finally, our results also suggest that a weak legal environment has limited the scope for well-managed firms to grow. So that improving the legal environment should encourage productivity-enhancing reallocation, helping to drive out badly managed firms.

BIBLIOGRAPHY

[Aprajit – could we slightly trim the econometrics references are slightly too many, thanks]

Acemoglu, Daron (2002) "Technical Change, Inequality and the Labor Market", *Journal of Economic Literature,* 40(1): 7-72.

Acemoglu Daron, Philippe Aghion, Claire Lelarge, John Van Reenen, and Fabrizio Zilibotti (2007) "Technology, Information and the Decentralization of the Firm", *Quarterly Journal of Economics*, 122(4), 1759–1799.

Aghion, Philippe, and Jean Tirole (1997) "Formal and Real Authority in Organizations", *Journal of Political Economy*, 105(1), 1-29.

Alonso, Ricardo, Wouter Dessein and Niko Matouschek (2008) "When Does Coordination Require Centralization", *American Economic Review*, 98(1), 145-179.

Andrews, D.W.K and Marmer, V. (2008) "Exactly Distribution-free Inference in Instrumental Variables Regression with Possibly Weak Instruments, " *Journal of Econometrics*. 142(1): 183-200

Autor, David, Katz, Lawrence, and Kearney, Melissa (2008), "Trends in US wage inequality: revising the revisionists", *Review of Economics and Statistics,* 90 (2): 300-323.

Baker, George, Robert Gibbons, and Kevin Murphy (1999) "Informal Authority in Organizations", *Journal of Law, Economics, and Organization*, 15(1), 56-73.

Baker, George, and Thomas Hubbard (2004) "Contractibility and Asset Ownership: On Board Computers and Governance in US Trucking", *Quarterly Journal of Economics*, 119(4), 1443-1479.

Banerjee, Abhijit and Ester Duflo (2000) "Reputation effects and the limits of contracting: a study of the Indian software industry", *Quarterly Journal of Economics* vol 115(3), pp. 989-1017.

Banerjee, Abhijit and Ester Duflo (2005) "Growth Through the Lens of Development Economics", in Philippe Aghion and Stephen Durlauf (eds), *Handbook of Economic Growth, Volume 1* of Handbook of Economic Growth, Chapter 7, pp. 473-552. Amsterdam: Elsevier.

Bartel, Ann, Casey Ichniowski and Kathryn Shaw, 2007. 'How Does Information Technology Really Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills', *Quarterly Journal of Economics*, 122(4), 1721-1758.

Bertrand, Marianne and Antoinette Schoar (2003) "Managing with Style: The Effects of Managers on Corporate Policy", *Quarterly Journal of Economics* 118(4): 1169-1208.

Bertrand, Marianne and Duflo, Esther and Mullainathan, S. (2004) "How Much Should We Trust Difference-in-Difference Estimators," *Quarterly Journal of Economics*. 119(1): 249-275

Bester, C.A and Conley, T.G and Hansen, C.B. (2008) "Inference with Dependent Data Using Cluster Covariance Estimators" *Working Paper*

Black, Sandra, and Lisa Lynch. 2001. "How to Compete: The Impact of Workplace Practices and Information Technology on Productivity." *Review of Economics and Statistics*, 88(3): 434-45.

Black, Sandra and Lisa Lynch. 2004. 'What's Driving the New Economy? The Benefits of Workplace Innovation', *Economic Journal*, 114(493), 97-116.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2009) "Americans do IT Better: American Multinationals and the Productivity Miracle", forthcoming American Economic Review.

Bloom, Nicholas, Sadun, Raffaella, and John Van Reenen (2009) "The organization of firms across countries", NBER Working Paper No. 15129.

Bloom, Nicholas, and John Van Reenen (2007) "Measuring and Explaining Management Practices across Firms and Countries", *Quarterly Journal of Economics,* 122(4), 1341-1408.

Bloom, Nicholas, and John Van Reenen (2010) "Human Resource Management and Productivity", draft chapter for the *Handbook of Labor Economics.*

Bloom, Nicholas, and John Van Reenen (2010) "Why do management practices differ across firms and countries", *Journal of Economic Perspectives*, 24(1), 203-224.

Bresnahan, Timothy, Erik Brynjolfsson, and Lorin Hitt (2002) "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-level Evidence", *Quarterly Journal of Economics*, 117(1), 339-376.

Bruhn, Miriam, Karlan, Dean and Schoar, Antoinette (2010), “The impact of offering consulting services to small and medium enterprises: evidence from a randomized trial in Mexico”, mimeo.

Cameron, Colin, Jonah Gelbach and Douglas Miller (2008) “Bootstrap-Based Improvements for Inference with Clustered Errors”, *Review of Economics and Statistics* 90(3): 414-27.

Cappelli, P and Neumark, D. 2001. “Do ‘High Performance” Work Practices Improve Establishment-Level Outcomes?” *Industrial and Labor Relations Review,* 737-775.

Clark, Greg (1987), “Why isn’t the whole world developed”, *Journal of economic history,* 47, pp. 141-173.

Conley, Tim and Udry, Christopher, (2010), “Learning about a new technology: pineapple in Ghana”, *American Economic Review* 100(1): 35-69.

Davis, Steve, Haltiwanger, John and Schuh, Scott, (1996), *Job Creation and Destruction*, MIT Press.

De Mel, Suresh, David McKenzie and Christopher Woodruff (2008) “Returns to Capital in Microenterprises: Evidence from a Field Experiment”, *Quarterly Journal of Economics* 113(4): 1329-1372.

Delery, John and Doty, Harold (1996), “Modes of theorizing in strategic human resource management: test of universalistic, contingency and configurational performance predictions”, *Academy of Management Review,* vol. 39(4), pp 802-835.

Drexler, Alejandro, Fischer, Greg and Schoar, Antoinette (2010), “Financial literacy training and rule of thumbs: evidence from a field experiment”, mimeo.

Duflo, Esther, Kremer, Michael and Robinson, Jonathan (2009), “Nudging farmers to use fertilizer: theory and experimental evidence from Kenya”, Harvard mimeo.

Foster, Lucia, John Haltiwanger and Chad Syverson (2008) “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, 98(1), 394-425

Garicano, Luis (2000) “Hierarchies and the Organization of Knowledge in Production”, *Journal of Political Economy*, 108(5), 874-904.

Garicano, Luis and Van Zandt, Timothy (2010), “Hierarchy”, draft chapter for publication in Robert Gibbons and John Roberts, editors, *Handbook of Organizational Economics*, Princeton University Press.

Gibbons, Robert and John Roberts (2010) *The Handbook of the Organizational Economics*, Princeton: Princeton University Press.

Greevy, R. and Silber, J.H. and Cnaan A. and Rosenbaum, P.R. (2004) Randomization Inference with Imperfect Compliance in the ACE-inhibitor after anthracycline randomized trial. *Journal of the American Statistical Association.* 99(465): 7-15

Guadalupe, Maria and Julie Wulf (2007) “The Flattening Firms and Product Market Competition: The Effects of Trade Costs and Liberalization”, Columbia University mimeo.

Hsieh, Chiang-Tai and Pete Klenow (2009), “Misallocation and Manufacturing TFP in China and India”, *Quarterly Journal of Economics* 124(4): 1403-1448.

Hsieh, Chiang-Tai and Pete Klenow (2010), “Development accounting”, American Economic Journal: Macroeconomics 2(1):207-223.

Huselid, Mark and Brian Becker, 1996. ‘Methodological Issues in Cross-sectional and Panel Estimates of the Human Resource-firm Performance link’, Industrial Relations, 35, 400-422.

Ibragimov, R. and Muller, U.K. (2009) “t-statistic Based Correlation and Heterogeneity Robust Inference,” *Journal of Business and Economic Statistics.*

Imbens, G.W and Rosenbaum, P.R. (2005) “Robust, Accurate Confidence Intervals with a Weak Instrument: Quarter of Birth and Education,” *Journal of the Royal Statistical Society: Series A.* vol 168:1 pp.109-126

Ichniowski, Casey, Kathryn Shaw and Giovanna Prenushi. (1997), “The Effects of Human Resource Management: A Study of Steel Finishing Lines”, *American Economic Review*, 87(3), 291-313.

Jorgenson, Dale, Mun Ho and Kevin Stiroh, 2008. ‘A Retrospective Look at the US Productivity Growth Resurgence’, *Journal of Economic Perspectives*, 22(1), 3-24.

Karlan, Dean and Martin Valdivia (2009) "Teaching Entrepreneurship: Impact Of Business Training On Microfinance Clients and Institutions", *Review of Economics and Statistics*, forthcoming.

Lazear, Edward, and Paul Oyer. 2009. "Personnel Economics." in Robert Gibbons and John Roberts, eds. forthcoming in *Handbook of Organizational Economics.*

Lehmann, E.L. and Romano, J.P. (2005) *Testing Statistical Hypotheses.* Springer-Verlag.

Levitt, Steven and List, John, (2009), "Was there really a Hawthorne Effect at the Hawthorne Works? An Analysis of the Original Illumination Experiments", NBER WP15016.

Lucas, Robert E. (1978): "On the size distribution of business firms." *Bell Journal of Economics*, 9:508-523.

MacDuffie, John Paul. (1995), "Human Resource Bundles and Management Performance: Organizational Logic and Flexible Production Systesm in the World Auto Industry," *Industrial and Labor Relations Review*, **48**(2): 197-221.

McKenzie, David (2009) "Impact Assessments in Finance and Private Sector Development: What have we learned and what should we learn?", *World Bank Research Observer*, forthcoming.

McKinsey Global Institute (2001) *India: The Growth Imperative*. http://www.mckinsey.com/mgi/publications/India.asp

Milgrom, Paul and Roberts, John (1990), "The economics of modern manufacturing: technology, strategy and organization", *American Economic Review*, 80 (3), pp. 511-528.

Mookherjee, Dilip (2010) draft chapter for publication in Robert Gibbons and John Roberts, editors, *Handbook of Organizational Economics*, Princeton *University Press.*

Nelson, Richard and Winter, Sydney (1982) An Evolutioinary Theory of Economics Change, Harvard University Press.

Osterman, Paul, 1994. 'How Common Is Workplace Transformation and Who Adopts It?', Industrial and Labor Relations Review, 47(2), 173-188 .

Pack, Howard (1987) *Productivity, Technology, and Industrial Development: A Case Study in Textiles,* World Bank Publications.

Rajan, Raghuram, and Luigi Zingales (2001) "The Firm as a Dedicated Hierarchy: A Theory of the Origin and Growth of Firms", *Quarterly Journal of Economics*, 116(3), 805-851.

Rajan, Raghuram, and Julie Wulf (2006) "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies", *Review of Economics and Statistics*, 88(4), 759-773.

Syverson, Chad (2004a), "Market Structure and Productivity: A concrete example", *Journal of Political Economy*, 112(6), 1181-1222.

Syverson, Chad (2004b), "Product substitutability and Productivity Dispersion", *Review of Economics and Statistics*, 86 (2), 534-50.

Syverson, Chad (2010), ''What determines productivity at the micro level?", draft manuscript for the Journal of Economic Literature.

Taylor, Fredrick (1911), Principles of Scientific management, Harper and brothers, New York and London.

Walker, Francis (1887), "The source of business profits", *Quarterly Journal of Economics*, 1(3), pp. 265-288.

Womack, James, Jones, Daniel and Roos, Daniel (1991), "The machine that changed the world", Harper Collins publishers, New York, USA.

Wooldridge, J. (2004) "Cluster Sample Methods in Applied Econometrics*," American Economic Review.*

Woodward J. 1958 *Management and Technology*, Cambridge: Cambridge University Press.

APPENDIX A: DATA

Our estimates for profits and productivity impacts are laid out in Table A5, with the methodology outlined below. We calculate the numbers for the median firm.

A. Estimations of profitability and productivity impacts.

We first generate the estimated impacts on quality, inventory and efficiency. To do this we take the Intention to Treat (ITT) numbers from Table 3, which shows a reduction of quality defects of 32% (exp(-0.386)-1), a reduction in inventory of 16.4% (exp(-0.179)-1) and an increase in output of 5.4% (exp(0.056)-1).

Mending wage bill:

Estimated by recording the total mending hours, which is 71,700 per year on average, times the mending wage bill which is 36 rupees (about $0.72) per hour. Since mending is undertaken on a piece-wise basis – so defects are repaired individually – a reduction the severity weighted defects should lead to a proportionate reduction in required mending hours.

Fabric revenue loss from non grade-A fabric:

Waste fabric estimated at 5% in the baseline, arising from cutting our defect areas and destroying and/or selling at a discount fabric with unfixable defects. Assume increase in quality leads to a proportionate reduction in waste fabric, and calculate for the median firm with sales of $6m per year.

Inventory carrying costs:

Total carrying costs of 22% calculated as interest charges of 15% (average prime lending rate of 12% over 2008-2010 plus 3% as firm-size lending premium – see for example http://www.sme.icicibank.com/Business_WCF.aspx?pid), 3% storage costs (rent, electricity, manpower and insurance) and 4% costs for physical depreciation and obsolescence (yarn rots over time and fashions change).

Increased profits from higher output

Increasing output is assumed to lead to an equi-proportionate increase in sales because these firms are small in their output markets, but would also increase variable costs of energy and raw-materials since the machines would be running. The average ratio of (energy + raw materials costs)/sales is 63%, so the profit margin on increased efficiency is 37%.

Labor and capital factor shares:

Labor factor share of 0.58 calculated as total labor costs over total value added using the "wearing apparel" industry in the most recent (2004-05) year of the Indian Annual Survey of industry. Capital factor share defined as 1-labor factor share, based on an assumed constant returns to scale production function and perfectly competitive output markets.

APPENDIX B: ECONOMETRICS

We briefly outline in this section the v. The proposed procedure by Ibragimov-Mueller (2009) (IM) is useful for our case where the number of entities (firms) is small but the number of observations per entity is large. Their approach can be outlined as follows: Implement the estimation method (OLS, IV, ITT) on each firm separately and obtain 17 firm-specific estimates. Note that we cannot do this for the control firms when estimating the ITT (since there is no within-firm variation for the right hand side for the control firms). Therefore the results from this procedure are essentially based on before-after comparisons for the treatment firms.

The procedure requires that the coefficient estimates from each entity are asymptotically independent and Gaussian (but with potentially different variances). In our case this would be justified by an asymptotics in T argument (recall we have about 170 observations per plant). In particular, we can be agnostic about the exact structure of correlations between observations within a firm as long as the parameter estimators satisfy a central limit theorem. Subject to this requirement, the extent of correlation across observations within an entity is unrestricted. In addition, different correlation structures across firms are permissible since the procedure allows for different variances for each firm level parameter. This "asymptotic heterogeneity" considerably relaxes the usual assumptions made in standard panel data contexts (such as those underlying the cluster covariance matrices in our main tables). Finally, IM show that the limiting standard Gaussian distribution assumption (for each firm) can be relaxed to accommodate heterogeneous scale mixtures of standard normal distributions as well.

The procedure comprises estimating separate coefficients for each of the 11 treatment firms (recall that we cannot use control firms for this procedure since they have no within firm variation in the instrument) and treating the resulting estimates as a set of independent random variables drawn from normal distributions.

We next summarize the ideas underlying the permutation based tests.We first describe the permutation test for the ITT parameter. We base the test on the Wei-Lachin (1994) statistic as described in Greevy et al (2004). The reason for using this statistic (as opposed to more commonly used ones) is that the permutation test for the IV parameter is a generalization of this procedure and so it natural to consider this procedure in the first step. Consider the vector of outcomes $\{Y_{i,t}\}_{t=1}^{T}$for plant *i* (we examine each outcome separately). Define the binary random assignment variable for firm *i* . Define the random variable

$$q_{i,j,t} = \mathbb{I}(Z_i > Z_j)\Big(\mathbb{I}(Y_{i,t} > Y_{j,t}) - \mathbb{I}(Y_{i,t} < Y_{j,t})\Big)$$

This variable takes on the values 0, 1 and -1. It is equal to zero if plant i is a control or plant j is a treatment plant of any of the outcome variables for either plant is missing. It is equal to +1 if plant i is a treatment plant, plant j is a control and the outcome for *i* is larger than the outcome for *j*. It is equal to -1 if plant i is a treatment plant, plant j is a control and the outcome for *i* is smaller than the outcome for *j*. The Wei-Lachin statistic can be written as

$$T = \sum_{i=1}^{N} Z_i q_i = \sum_{i=1}^{N} Z_i \sum_{t=1}^{T} \sum_{j=1}^{N} q_{i,j,t}$$

Under the null hypothesis of no treatment effect, the treatment outcomes should not be systematically larger than the control outcomes (note that this means that we do not use pre-treatment observations for all plants). Specifically, under the null hypothesis and conditional upon the order statistics, each possible candidate value of T has an an equal probability of occurring. We use this insight to construct a critical value for the test. Consider one of the $\binom{17}{11}$ combinations of the firm treatment assignment variable *Z*. For each such permutation, compute *T*. Form the empirical distribution of T by considering all possible permutations and record the appropriate quantile for the distribution of *T* thus generated (in the one-sided alternative case this would be the $1-\alpha$ quantile). Finally, reject the null hypothesis of no treatment effect if the original statistic *T* exceeds this quantile. Greevy et al (2004), using previous results (Rosenbaum (2002) and Imbens and Rosenbaum (2005)) show that this test has exact size α for any sample size *n*. Therefore, the conclusions of this test do not rely upon any asymptotic theory. Instead, the results lean heavily on the idea of exchangeability – the property that changing the ordering of a sequence of random variables does not affect their joint distribution. For our application, this notion seems reasonable. Note that exchangeability is weaker than the i.i.d. assumption so for instance outcomes across firms can even be correlated (as long as they are equi-correlated).

Consider next the randomization inference based test for the IV case. We first consider the cross-section. Define the counterfactual model for outcomes $Y_d = \tau + \beta d + \epsilon$ and let D_j denote potential treatment status when treatment assignment is j. Define observed treatment status as $D = ZD_1 + (1-Z)D_0$. In our case, the treatment status is the fraction of the 38 practies that the firm has implemented. The maintained assumption is that the potential outcomes are independent of the instrument *Z* or equivalently (ϵ, D_1, D_0) is independent of *Z* and the error term has mean 0. We observe a random sample on (D, Z, Y_D) and wish to test the null hypothesis $\mathrm{H} : \beta = \beta_0$ against the two-sided alternative. Note that under the null hypothesis, $\tilde{Y} \equiv Y - \tau - \beta_0 D = \epsilon$ is independent of *Z* and we use this fact to construct a test along the lines of the previous test. Consider the analogue of the first equation

$$q_{i,j} = \mathbb{I}(Z_i > Z_j)\left(\mathbb{I}(\tilde{Y}_i > \tilde{Y}_j) - \mathbb{I}(\tilde{Y}_i < \tilde{Y}_j)\right)$$

Where we have replaced the response Y by the response subtracted by $\tau + \beta_0 D$. Note that τ is consistently estimable under the null, so without loss of generality we can treat it as known. This point is discussed in greater detail below. The key to this rewriting is to notice that under the null hypothesis, is independent of *Z* so that we can use the same argument as before to derive the finite sample distribution of *T* .For our data, we modify this approach to allow for a panel and covariates (time and plant dummies). This parallels the proposal in Andrews and Marmer (2008) and we can define

$$\tilde{Y_{i,t}} = Y_{i,t} - \beta_0 D_{i,t} - X'_{i,t}\hat{\delta}$$

and we form the statistic as

$$\tilde{T} = \sum_{i=1}^{N} Z_i q_i = \sum_{i=1}^{N} Z_i \sum_{t=1}^{T} \sum_{j=1}^{N} \tilde{q}_{i,j,t}$$

Where

$$\tilde{q}_{i,j,t} = \mathbb{I}(Z_i > Z_j)\left(\mathbb{I}(\tilde{Y}_{i,t} > \tilde{Y}_{j,t}) - \mathbb{I}(\tilde{Y}_{i,t} < \tilde{Y}_{j,t})\right)$$

For each candidate value of β, we form $\{\tilde{Y_{i,t}}\}_{i,t}$ and carry out the permutation test (as described in the ITT case above and noting that we do not use pre-treatment outcomes). We collect the set of values for which we could not reject the null hypothesis (against the two-sided alternative at α=.05) to construct an exact confidence set for β. As Imbens and Rosenbaum (2005) point out, there is no reason for a confidence set constructed in this manner to be a single interval. However, in our estimation, the confidence sets were intervals.

Table A1: The textile management practices adoption rates

Area	Specific practice	Pre-intervention *level*		Post-intervention *change*	
		Treatment	**Control**	**Treatment**	**Control**
Factory Operations	Preventive maintenance is carried out for the machines	0.429	0.667	0.286	0
	Preventive maintenance is carried out per manufacturer's recommendations	0.071	0	0.071	0.167
	The shop floor is marked clearly for where each machine should be	0.071	0.333	0.214	0.167
	The shop floor is clear of waste and obstacles	0	0.167	0.214	0.167
	Machine downtime is recorded	0.571	0.667	0.357	0
	Machine downtime reasons are monitored daily	0.429	0.167	0.5	0.5
	Machine downtime analyzed at least fortnightly & action plans implemented to try to reduce this	0	0.167	0.714	0
	Daily meetings take place that discuss efficiency with the production team	0	0.167	0.786	0.5
	Written procedures for warping, drawing, weaving & beam gaiting are displayed	0.071	0.167	0.5	0
	Visual aids display daily efficiency loomwise and weaverwise	0.214	0.167	0.643	0.167
	These visual aids are updated on a daily basis	0.143	0	0.643	0.167
	Spares stored in a systematic basis (labeling and demarked locations)	0.143	0	0.143	0.167
	Spares purchases and consumption are recorded and monitored	0.571	0667	0.071	0.167
	Scientific methods are used to define inventory norms for spares	0	0	0.071	0
Quality Control	Quality defects are recorded	0.929	1	0.071	0
	Quality defects are recorded defect wise	0.286	0.167	0.643	0.833
	Quality defects are monitored on a daily basis	0.286	0.167	0.714	0.333
	There is an analysis and action plan based on defects data	0	0	0.714	0.167
	There is a fabric gradation system	0.571	0.667	0.357	0
	The gradation system is well defined	0.500	0.5	0.429	0
	Daily meetings take place that discuss defects and gradation	0.071	0.167	0.786	0.167
	Standard operating procedures are displayed for quality supervisors & checkers	0	0	0.714	0
Inventory Control	Yarn transactions (receipt, issues, returns) are recorded daily	0.929	1	0.071	0
	The closing stock is monitored at least weekly	0.214	0.167	0.571	0.5
	Scientific methods are used to define inventory norms for yarn	0	0	0.083	0
	There is a process for monitoring the aging of yarn stock	0.231	0	0.538	0
	There is a system for using and disposing of old stock	0	0	0.615	0.6
	There is location wise entry maintained for yarn storage	0.357	0	0.357	0
Loom Planning	Advance loom planning is undertaken	0.429	0.833	0.214	0
	There is a regular meeting between sales and operational management	0.429	0.500	0.143	0
Human Resources	There is a reward system for non-managerial staff based on performance	0.571	0.667	0.071	0
	There is a reward system for managerial staff based on performance	0.214	0.167	0.286	0
	There is a reward system for non-managerial staff based on attendance	0.214	0.333	0.357	0
	Top performers among factory staff are publicly identified each month	0.071	0	0.357	0
	Roles & responsibilities are displayed for managers and supervisors	0	0	0.643	0
Sales and Orders	Customers are segmented for order prioritization	0	0	0	0.167
	Orderwise production planning is undertaken	0.692	1	0.231	0
	Historical efficiency data is analyzed for business decisions regarding designs	0	0	0.071	0
All	Average of all practices	0.256	0.288	0.378	0.120
p-value for the difference between the average of all practices		0.510		0.000	

Notes: Reports the 38 individual management practices measured before, during and after the management intervention. The columns **Pre Intervention *level* of Adoption** report the pre-intervention share of plants adopting this practice for the 14 treatment and 6 control plants. The columns **Post Intervention *increase* in Adoption** report the changes in adoption rates between the pre-intervention period and 4 months after the end of the diagnostic phase (so right after the end of the implementation phase for the treatment plants) for the treatment and control plants. The **p-value for the difference between the average of all practices** reports the significance of the difference in the average level of adoption and the increase in adoption between the treatment and control groups.

Table A2: The decentralization survey:

For all questions except D7 any score can be given, but the scoring guide is only provided for scores of 1, 3 and 5.

Question D1: “*What authority does the plant manager(or other managers) have to hire a WEAVER (e.g. a worker supplied by a contractor)?*”

	Score 1	Score 3	Score 5
Scoring grid:	No authority – even for replacement hires	Requires sign-off from the Director based on the business case. Typically agreed (i.e. about 80% or 90% of the time).	Complete authority – it is my decision entirely

Question D2: “*What authority does the plant manager(or other managers) have to hire a junior Manager (e.g. somebody hired by the firm)?*”

	Score 1	Score 3	Score 5
Scoring grid:	No authority – even for replacement hires	Requires sign-off from the Director based on the business case. Typically agreed (i.e. about 80% or 90% of the time).	Complete authority – it is my decision entirely

Question D3: “*What authority does the plant manager (or other managers) have to purchase spare parts?*”?

Probe until you can accurately score the question. Also take an average score for sales and marketing if they are taken at different levels.

	Score 1	Score 3	Score 5
Scoring grid:	No authority	Requires sign-off from the Director based on the business case. Typically agreed (i.e. about 80% or 90% of the time).	Complete authority – it is my decision entirely

Question D4: “*What authority does the plant manager (or other managers) have to plan maintenance schedules?*”

	Score 1	Score 3	Score 5
Scoring grid:	No authority	Requires sign-off from the Director based on the business case. Typically agreed (i.e. about 80% or 90% of the time).	Complete authority – it is my decision entirely

Question D5: “*What authority does the plant manager (or other managers) have to award small (<10% of salary) bonuses to workers?*”

	Score 1	Score 3	Score 5
Scoring grid:	No authority	Requires sign-off from the Director based on the business case. Typically agreed (i.e. about 80% or 90% of the time).	Complete authority – it is my decision entirely

Question D6: “*What is the largest expenditure (in rupees) a plant manager (or other managers) could typically make without your signature?*”

Question D7: “*What is the extent of follow-up required to be done by the directors?*”

	Score 1	Score 3	Score 5
Scoring grid:	Directors are the primary point of contact for information exchange between managers	Frequent follow ups on about half of the decisions made by managers	Minimal follow-ups on decisions taken between managers. Only dispute resolution.

Question D8: “*How many days a week did the director spend away from the factory last month?*”

Table A3: The computerization survey:

Question C1: "Does the plant have an Electronic resource planning system?"

Question C2: "How many computers does the plant have?"

Question C3: "How many of these computers are less than 2 years old"

Question C4: "How many people in the factory typically use computers for at least 10 minutes day?"

Question C5: "How many cumulative hours per week are computers used in the plant"?

Question C6: "Does the plant have an internet connection"

Question C7: "Does the plant manager use e-mail (for work purposes)?"

Question C8: "Does the plant manager use e-mail (for work purposes)?"

Question C9: "*What is the extent of computer use in operational performance management?*" *(and score from 1 to 5 is possible, but scores given for 1,3, and 5)*

	Score 1	Score 3	Score 5
Scoring grid:	Computers not used in operational performance management	Around 50% of operational performance metrics (efficiency, inventory, quality and output) are tracked & analyzed through computer/ERP generated reports.	All main operational performance metrics (efficiency, inventory, quality and output) are tracked & analyzed through computer/ERP generated reports.

Table A4: Descriptive statistics for the Decentralization and Computerization survey

	Mean pre-level	**Min pre-level**	**Max pre-level**	**SD pre-level**	**Mean change**	**Correlation of change with treatment status**
Decentralization questions						
D1 (weaver hiring)	4.68	3	5	0.72	0	n/a
D2 (manager hiring)	1.93	1	4	1.05	0.36	0.198
D3 (spares purchases)	2.61	1	4	0.79	0.18	0.121
D4 (maintenance planning)	4.50	1	5	1	0.04	0.133
D5 (worker bonus pay)	2.25	1	4	1.14	0.29	0.375
D6 (investment limit, rupees)	10357	1000	35000	10434	714	0.169
D7 (director coordination)	2.78	2	4	0.69	0.36	0.358
D8 (days director not at the factory per week)	2.69	0	4.75	1.30	0.39	0.282
Decentralization index	0	-1.33	1.52	1	0.44	0.355
Computerization questions						
C1 (ERP)	0.74	0	1	0.44	0	n/a
C2 (number computers)	2.68	0	8	1.98	0.36	0.377
C3 (number new computers)	0.43	0	8	1.55	0.29	0.189
C4 (computer users)	3	0	10	2.21	0.11	0.308
C5 (computer hours)	10	0	48	12.20	5.34	0.439
C6 (internet connection)	0.64	0	1	0.49	0.036	0.133
C7 (plant manager e-mail)	0.29	0	1	0.46	0.04	-0.280
C8 (directors e-mail)	0.82	0	1	0.39	0	n/a
C9 (production computerization)	2.71	1	5	0.98	0.89	0.367
Computerization index	0	-1.58	3.15	1	0.458	0.440

Notes: There are about 50 rupees to the dollar. The mean change measures the different between pre the experiment and August 2010. The decentralization index and the computerization index are normalized to have a mean of zero and standard deviation of unity on the pre-experiment sample.

Table A5: Estimated impact of improved quality, inventory and efficiency

Change	Impact	Estimation approach	Estimated impact
Profits (annual in $)			
Improvement in quality	Reduction in repair manpower	Reduction in defects (32%) times median mending manpower wage bill ($41,000).	$13,000
	Reduction in waste fabric	Reduction in defects (32%) times the average yearly waste fabric (5%) times median average sales ($6m).	$96,000
Reduction in inventory	Reduction in inventory carrying costs	Reduction in inventory (16.4%) times carrying cost of inventory (22%) times median inventory ($230,000)	$8,000
Increased efficiency	Increased sales	Increase in output (5.4%) times margin on sales (37%) times median sales ($6m)	$121,000
Total			$238,000
Productivity (%)			
Improvement in quality	Reduction in repair manpower	Reduction in defects (32%) times share of repair manpower in total manpower (18.7%) times labor share (0.58) in output in textiles (from the 2003-04 Indian Annual Survey of Industries.)	3.5%
	Reduction in waste fabric	Reduction in defects (31.9%) times the average yearly waste fabric (5%)	1.6%
Reduction in inventory	Reduction in capital stock	Reduction in inventory (16.4%) times inventory share in capital (8%) times capital factor share in output in textiles (0.42)	0.6%
Increased efficiency	Increased output	Increase in output (5.4%) without any change in labor or capital	5.4%
Total			11.1%

Notes: Estimated impact of the improvements in the management intervention on firms profitability and productivity through quality, inventory and efficiency using the estimates in Table 3. Figure calculated for the median firm. See Appendix A for details of calculations for inventory carrying costs, fabric waste, repair manpower and factor shares.

Table 1: The field experiment sample

	All				Treatment	Control	Diff
	Mean	Median	Min	Max	Mean	Mean	p-value
Sample sizes:							
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
Firm/plant sizes:							
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	12.8	7.9	2.85	44.2	13.3	12.0	0.837
Daily mtrs, experimental plants	5560	5130	2260	13000	5,757	5,091	0.602
Management and plant ages:							
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
Performance measures							
Operating efficiency (%)	70.77	72.8	26.2	90.4	70.2	71.99	0.758
Raw materials inventory (kg)	59,497	61,198	6,721	149,513	59,222	60,002	0.957
Quality (% A-grade fabric)	40.12	34.03	9.88	87.11	39.04	41.76	0.629

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 2 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. Note that none of the differences between the means of the treatment and control plants are significant. **Raw materials inventory** is the stock of yarn per intervention. **Operating efficiency** is the percentage of the time the machines are producing fabric per intervention. **Quality (% A-grade fabric)** is the percentage of fabric each plant defines as A-grade, which is the top quality grade.

Table 2: The impact of the treatment on management practices within and across plants

Dependent Variable	Overall Management	Overall Management	Overall Management	Overall Management	Overall Management
	(1)	(2)	(3)	(4)	(5)
Own plant treatment$_{i,t}$	0.109***	0.108***	0.108***	0.105***	0.105***
Months consulting in own plant	(0.014)	(0.016)	(0.016)	(0.018)	(0.018)
Spillover treatment$_{i,t}$	0.041**	-0.007		0.030	
Months consulting in other plants within the same firm	(0.016)	(0.023)		(0.024)	
3 month lagged spillover treatment$_{i,t-3}$		0.054**	0.047***		
Months consulting in other plants within the same firm		(0.023)	(0.016)		
6 month lagged spillover treatment$_{i,t-6}$				0.027	0.050***
Months consulting in other plants within the same firm				(0.021)	(0.017)
Time FEs	12	11	11	10	10
Plant FEs	28	28	28	28	28
Observations	336	308	308	280	280
R-squared	0.904	0.909	0.909	0.889	0.889

Notes: The dependent variable is the share of the 38 management practices adopted in each plant. This is regressed against the cumulative weeks of intervention in the own plant ("Own plant treatment"), the cumulative weeks of treatment in other plants within the same firm ("Spillover treatment"), and this variable lagged three and six months ("3 month Lag spillover treatment" and "6 month Lag spillover treatment"). The data is quarterly until April 2009 and bi-monthly thereafter, reflecting the frequency of measurement of management practices. A full set of time-dummies and plant dummies is included. Standard errors are bootstrap clustered at the firm level.

Table 3: The impact of modern management practices on plant performance

Dependent Variable	Quality (log QDI)			Inventory (log tons)			Output (log picks)		
Specification	OLS	IV	ITT	OLS	IV	ITT	OLS	IV	ITT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Management_{i,t}$	-0.561	-2.028***		-0.639***	-0.929**		0.127	0.346**	
Adoption of management practices	(0.440)	(0.685)		(0.242)	(0.389)		(0.099)	(0.171)	
$Intervention_{i,t}$			-0.386**			-0.179**			0.056
Intervention stage initiated			(0.162)			(0.089)			(0.034)
Instrument		Log (1+ weeks treatment)			Log (1+ weeks treatment)			Log (1+ weeks treatment)	
Small sample robustness									
Ibragimov-Mueller (95% CI)	(-4.56,3.59)	(-6.66,-0.18)	(-0.79,-0.28)	(-0.88,0.35)	(-0.69,-0.20)	(-0.17,0.00)	(-.13,0.66)	(0.01,1.24)	(0.03,0.19)
(90%CI)	(-3.83,2.86)	(-6.06,-0.79)	(-0.74,-0.33)	(-0.77,0.24)	(-0.65,-0.24)	(-0.16,-.001)	(-0.06,0.59)	(0.19,1.80)	(0.05,0.18)
Permutation Test I (p-value)			.0168			.1262			.0263
IV Permutation Tests (95% CI)		(-6.05,0.35)			(-3.23,0.84)			(0.16,0.47)	
(90% CI)		(-6.00,-0.03)			(-2.29,0.62)			(0.18,0.45)	
Time FEs	113	113	113	113	113	113	114	114	114
Plant FEs	20	20	20	18	18	18	20	20	20
Observations	1732	1732	1732	1977	1977	1977	2312	2312	2312

Notes: All regressions use a full set of plant and calendar week dummies. Standard errors bootstrap clustered at the firm level. **Quality (log QDI)** is a log of the quality defects index (QDI), which is a weighted average score of quality defects, so higher numbers imply worse quality products (more quality defects). **Inventory (log tons)** is the log of the tons of yarn inventory in the plant. **Output (log picks)** is the log of the weaving production picks. **Management** is the adoption of the 38 management practices listed in table 2. **Intervention (implementation)** is a plant level indicator taking a value of 1 after the implementation phase has started at a treatment plant. **Log(1+weeks of treatment)** is the log of one plus the cumulative count of the weeks since the start of the implementation in each plant (treatment plants only), and value zero before. **OLS** reports results with plant estimations. **IV** reports the results where the management variable has been instrumented with log(1+ cumulative intervention weeks). **ITT** reports the intention to treat results from regressing the dependent variable directly on the 1/0 intervention indicator. **Time FEs** report the number of calendar week time fixed effects. **Plant FEs** reports the number of plant-level fixed effects. Two plants do not have any inventory on site, so no inventory data is available. The **Robustness Checks** implement three different procedures (described in greater detail in Appendix B) to address issues of plant heterogeneity, within plant (and firm) correlation, and small sample concerns, where **95% CI** and **90% CI** report 95% and 90% confidence intervals. **Ibragimov-Mueller** estimates parameters firm-by-firm and then treats the estimates as a draw from independent (but not identically distributed) normal distributions. **Permutation Test I** reports the p-values for testing the null hypothesis that the treatment has no effect for the ITT parameter by constructing a permutation distribution of the ITT estimate using 1000 possible permutations (out of 12376) of treatment assignment. **IV-Permutation** tests implements a permutation test for the IV parameter using 1000 possible permutations (out of 12376) of treatment assignment. These tests have exact finite sample size..

Table 4: The impact of modern management practices on organization and computerization

Dep. variable:	Decentralization Index			Hours of computer use			Computerization index		
Specification	OLS (1)	IV (2)	ITT (7)	OLS (4)	IV (5)	ITT (6)	OLS (7)	IV (8)	ITT (9)
Management$_{i,t}$ Adoption of management practices	1.695*** (0.420)	1.837*** (0.535)		16.761*** (3.457)	23.272** (6.708)		0.951*** (0.275)	1.202*** (0.438)	
Intervention$_{i,t}$ Intervention stage initiated			0.360** (0.164)			4.536*** (1.555)			0.230* (0.121)
Instrument		Log (1+weeks treatment)			Log (1+weeks treatment)			Log (1+weeks treatment)	
Time FEs	3	3	3	3	3	3	3	3	3
Plant FEs	28	28	28	28	28	28	28	28	28
Observations	84	84	84	84	84	84	84	84	84

Notes: All regressions use three observations per firm (pre intervention, March 2010 and August 2010), and a full set of plant dummies and time dummies. Standard errors bootstrap clustered at the firm level. **Management** is the adoption of the 38 management practices listed in table A1. **Decentralization index** is the principal component factor of 7 measures of decentralization around weaver hiring, manager hiring, spares purchases, maintenance planning, weaver bonuses, investment, and departmental co-ordination. This has a standard deviation of 1 and a mean of 0. **Hours of computer use** is the hours of computer use. This has a (pre-intervention) mean and standard deviation of 13.66 and 12.20. **Computerization index** is the principal component factor of 10 measures around computerization, which are the use of an ERP system, the number of computers in the plant, the number of computers less than 2 years old, the number of employees using computers for at least 10 minutes per day, and the cumulative number of hours of computer use per week, an internet connection at the plant, if the plant-manager uses e-mail, if the directors use of e-mail, and the intensity of computerization in production. The other computerization columns show the results for the individual components of this index that changed over time (the omitted components did not change). This has a standard deviation of 1 and a mean of 0. **Log(1+weeks of treatment)** is the log of one plus the cumulative count of the weeks since the start of the implementation in each plant (treatment plants only), and value zero before. **OLS** reports results with plant estimations. **IV** reports the results where the management variable has been instrumented with log(1+ cumulative intervention weeks). **ITT** reports the intention to treat results from regressing the dependent variable directly on the 1/0 intervention indicator. **Time FEs** reports the number of time fixed effects. **Plant FEs** reports the number of plant-level fixed effects. **SD of dep. var.** reports the standard deviation of the dependent variable.

Table 5: Reasons for bad management, as a percentage (%) of all practices, before and after treatment

Non-adoption reason	Firm group	1 month before	1 month after	3 months after	5 months after	7 months after	9 months after
Lack of information (plants not aware of the practice)	Treatment	18.5	13.5	2.0	0.6	0	0
	Control	12.9	9.6	8.0	8.0	8.0	8.0
	Non-experimental	9.3	6.8	3.8	3.8	3.8	3.8
Incorrect information (plants incorrect on cost-benefit calculation)	Treatment	44.4	36.6	33.6	31.3	31.1	30.2
	Control	46.7	45.3	44.2	43.1	42.2	42.2
	Non-experimental	41.2	42.0	38.6	35.6	34.6	33.6
Owner lack of time, low ability or procrastination (the owner is the reason for non adoption)	Treatment	10.3	7.5	7.2	7.5	7.7	6.8
	Control	11.6	10.2	9.3	9.8	8.4	8.4
	Non-experimental	23.5	22.0	27.0	31.5	26.3	26.6
Not profitable (the consultants agree non-adoption is correct)	Treatment	0.5	0.5	0.5	0.5	0.5	0.5
	Control	0	0	0	0	0	0
	Non-experimental	0	0	0	0	0	0
Other (variety of other reasons for non-adoption)	Treatment	0.2	0.2	0.2	0.2	0.2	0.2
	Control	0	0	0	0	0.9	0.9
	Non-experimental	0.3	0.3	0.3	0.3	0.3	0.3
Total (sum of all individual reasons)	Treatment	74.4	58.2	45.5	40.1	39.9	38.1
	Control	71.2	65.1	61.6	60.9	60.6	60.6
	Non-experimental	73.4	71.0	70.7	69.8	65.4	64.7

Notes: Show the percentages (%) of practices not adopted by reason for non-adoption, in the treatment plants, control plants and non-experimental plants. Timing is relative to the start of the treatment phase (the end of the diagnostic phase for the control group and the start of the treatment phase for the other plant in their firm for the non-experimental plants). Covers 532 practices in treatment plants (38 practices in 14 plants), 228 practices in the control plants (38 practices in 6 plants) and 30 practices in the non-experimental plants (38 practices in 8 plants). Non adoption was monitored every other month using the tool shown in Figure 4, based on discussions with the firms' directors, managers, workers, plus regular consulting work in the factories. Note that data is only currently available up to 7 months after the end of diagnostic phase in the control firms.

Exhibit 1: Plants are large compounds, often containing several buildings.

Plant entrance with gates and a guard post

Plant surrounded by grounds

Front entrance to the main building

Plant buildings with gates and guard post

Exhibit 2: These factories operate 24 hours a day for 7 days a week producing fabric from yarn, with 4 main stages of production

(1) Winding the yarn thread onto the warp beam

(2) Drawing the warp beam ready for weaving

(3) Weaving the fabric on the weaving loom

(4) Quality checking and repair

Exhibit 3: Many parts of these factories were dirty and unsafe

Garbage outside the factory

Garbage inside a factory

Flammable garbage in a factory

Chemicals without any covering

Exhibit 4: The factory floors were frequently disorganized

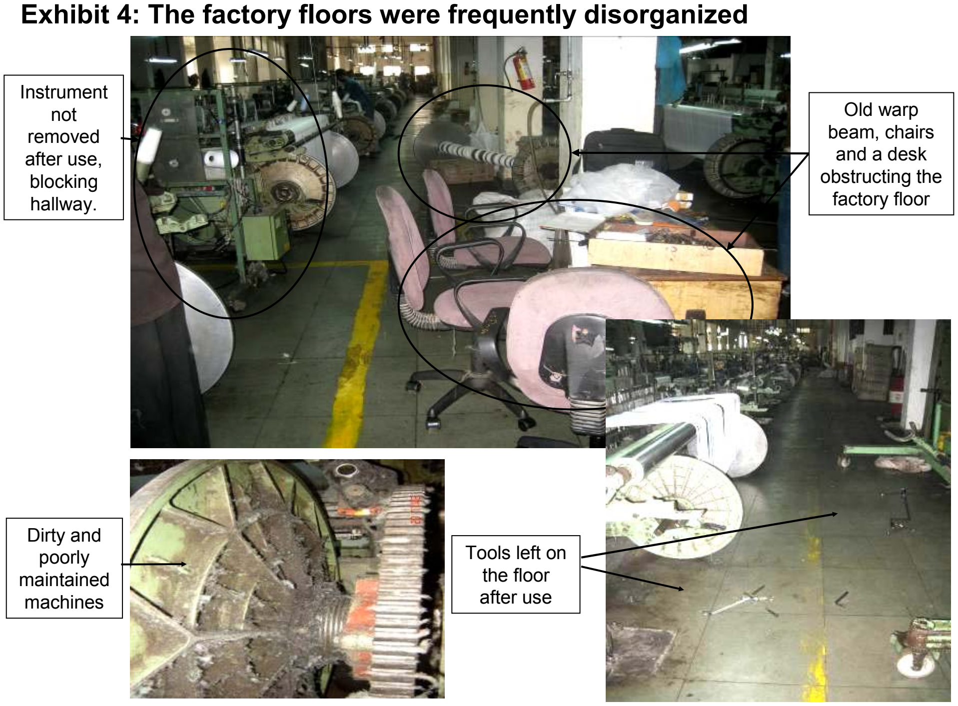


Exhibit 5: Most plants had months of excess yarn, usually spread across multiple locations, often without any rigorous storage system

Exhibit 6: The parts stores were often disorganized and dirty

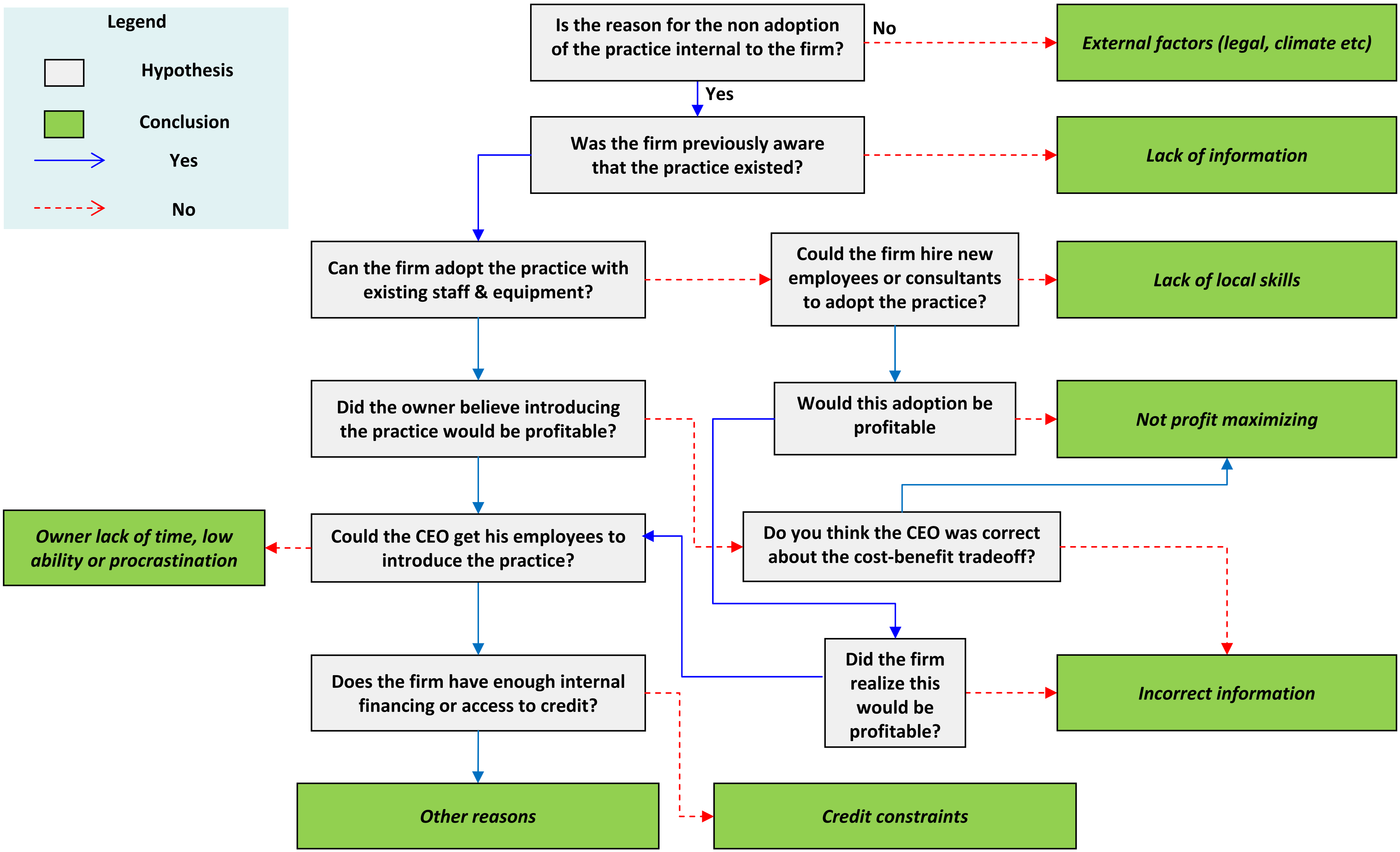
Spares without any labeling or

No protection to prevent damage and rust

Spares without any labeling or order

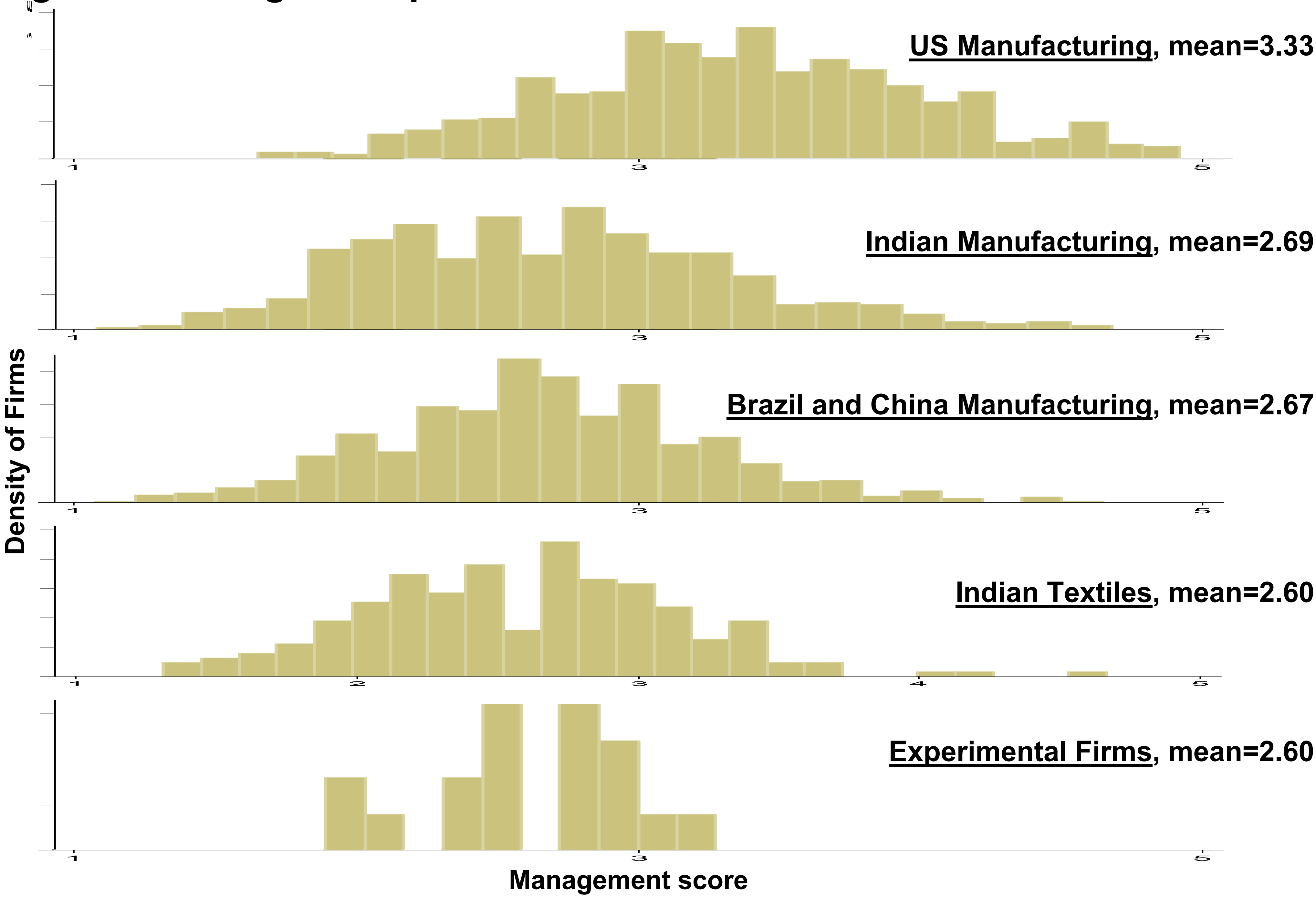
Shelves overfilled and disorganized

Exhibit 7: Non adoption flow chart used by consultants to collect data



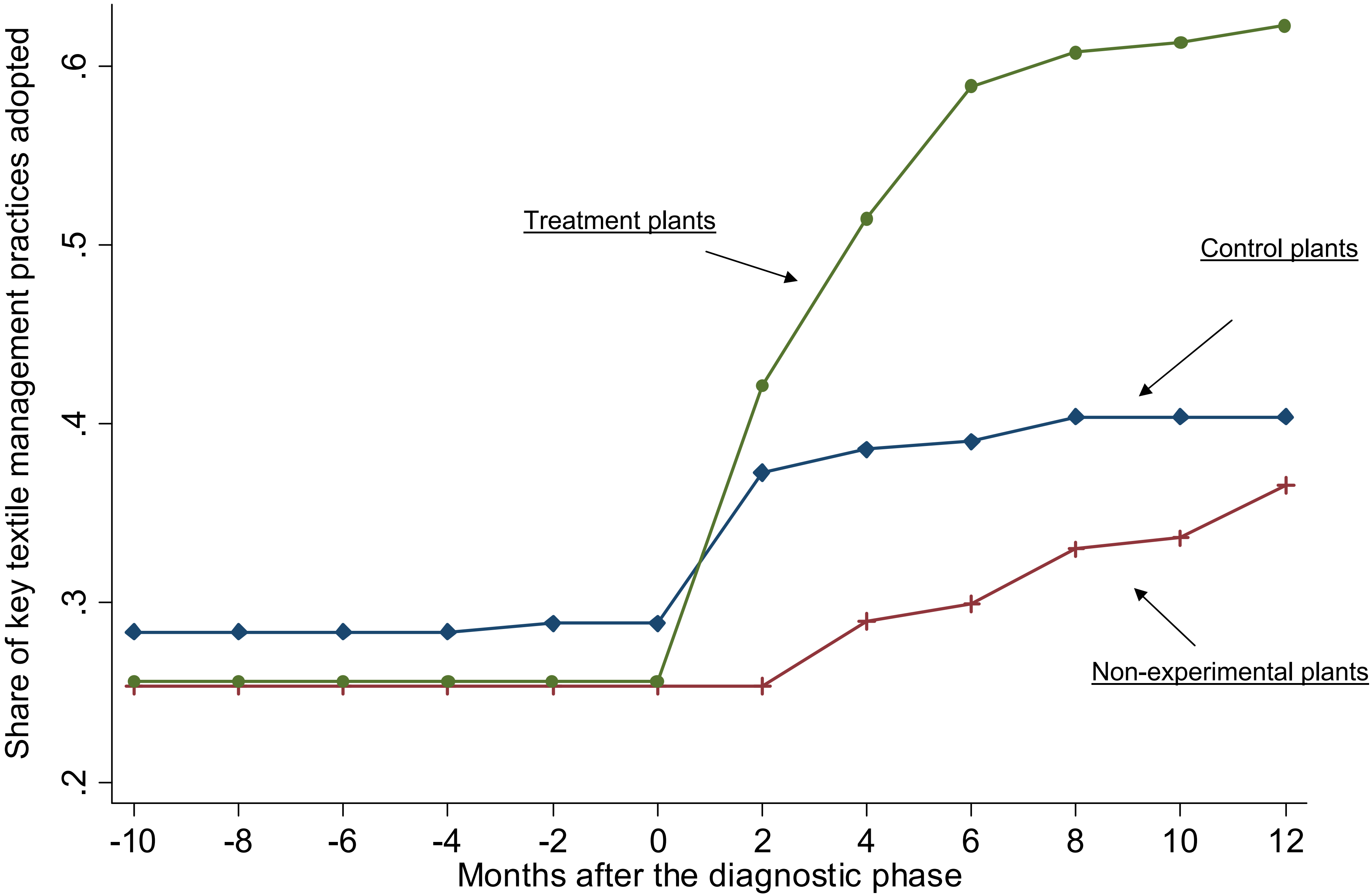
Notes: The consultants used the flow chart to evaluate why each particular practice from the list of 38 in Table 2 had not been adopted in each firm, on a bi-monthly basis. Non adoption was monitored every other month based on discussions with the firms' directors, managers, workers, plus regular consulting work in the factories.

Figure 1: Management practice scores across countries



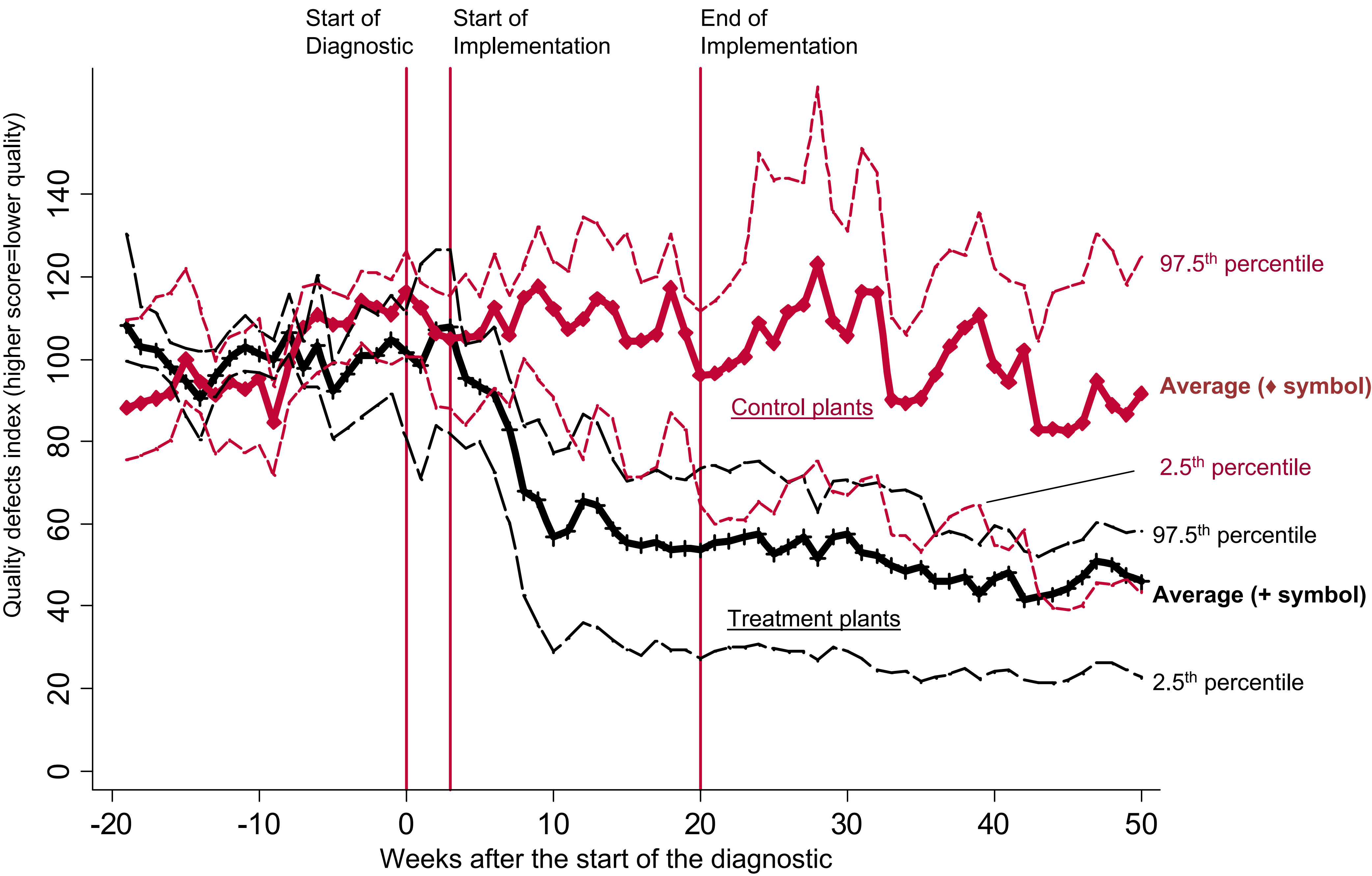
Notes: Management practice histograms using the Bloom and Van Reenen (2010) data. Double-blind surveys used to evaluate firms monitoring, targets and operations. Scores range from 1 (worst practice) to 5 (best practice). Samples are 17 experimental plants, 232 Indian textile firms, 620 Indian manufacturing firms, and 1083 Brazilian and Chinese firms.

Figure 2: The adoption of key textile management practices over time



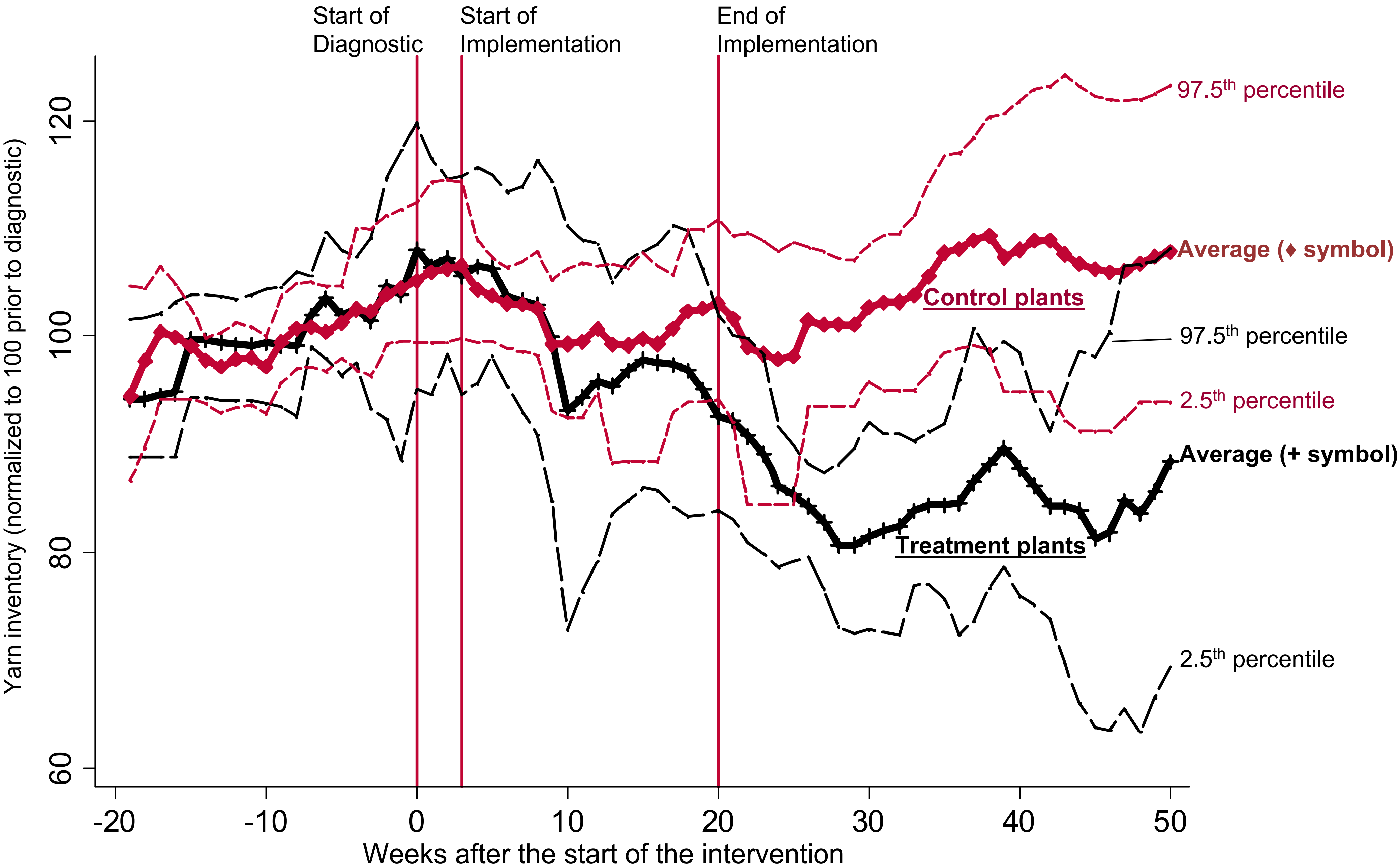
Notes: Average adoption rates of the 38 key textile manufacturing management practices listed in Table 2. Shown separately for the 14 treatment plants (round symbol), 6 control plants (diamond symbol) and the 8 non-experimental plants which the consultants did not provide any direct consulting assistance to (+ symbol). Scores range from 0 (if none of the group of plants have adopted any of the 38 management practices) to 1 (if all of the group of plants have adopted all of the 38 management practices). Initial differences across all the groups are not statistically significant.

Figure 3: Quality defects index for the treatment and control plants



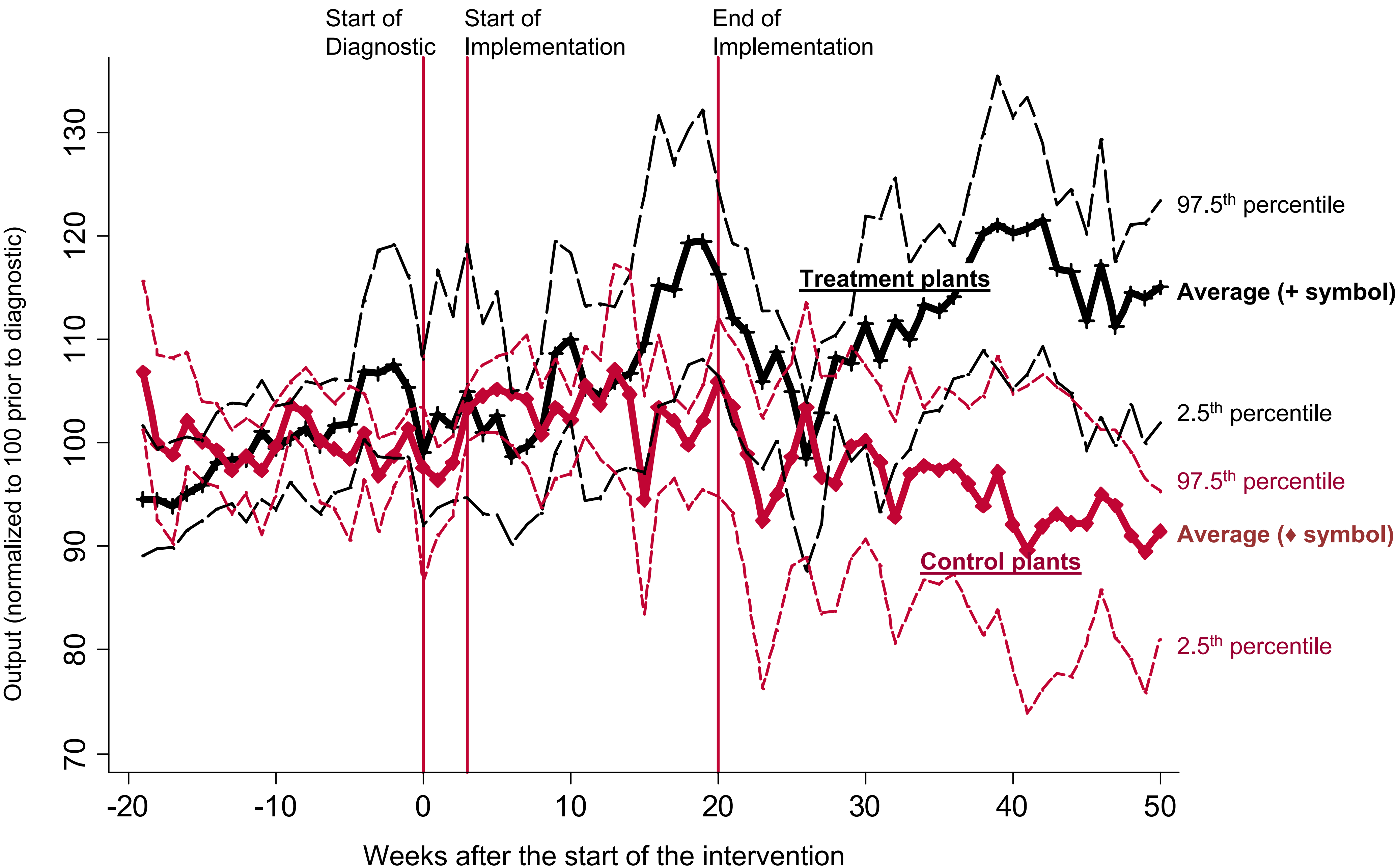
Notes: Displays the average weekly quality defects index, which is a weighted index of quality defects, so a higher score means lower quality. This is plotted for the 14 treatment plants (+ symbols) and the 6 control plants (♦ symbols). Values normalized so both series have an average of 100 prior to the start of the intervention. To obtain confidence intervals we bootstrapped the firms with replacement 250 times.

Figure 4: Yarn inventory for the treatment and control plants



Notes: Displays the weekly average yarn inventory plotted for 12 treatment plants (+ symbols) and the 6 control plants (♦ symbols). Values normalized so both series have an average of 100 prior to the start of the intervention. To obtain confidence intervals we bootstrapped the firms with replacement 250 times. Slow moving fluctuations due to seasonality. 2 treatment plants maintain no on-site yarn inventory so are not included in the figures.

Figure 5: Output for the treatment and control plants



Notes: Displays the weekly average output for the 14 treatment plants (+ symbols) and the 6 control plants (♦ symbols). Values normalized so both series have an average of 100 prior to the start of the intervention. To obtain confidence intervals we bootstrapped the firms with replacement 250 times.