Working paper

International Growth Centre

Does Management Matter?

Evidence from India

Nicholas Bloom
Benn Eifert
Aprajit Mahajan
David McKenzie
John Roberts

December 2010

DIRECTED BY

FUNDED BY

Abstract

A long-standing question in social science is to what extent 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 these management practices had three main effects. First, it raised average productivity by 11 per cent through improved quality and efficiency and reduced inventory. Second, it increased decentralization of decision making, as better information flow enabled owners to delegate more decisions to middle 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 technology that diffuses slowly between firms, with many Indian firms initially unaware of its existence or impact. Since competition was limited by constraints on firm entry and growth, badly managed firms were not rapidly driven from the market.

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

Keywords: management, organization, IT, productivity and India.

Acknowledgements

Financial support was provided by the Alfred Sloan Foundation; the Freeman Spogli Institute, the International Initiative and the Graduate School of Business at Stanford; the International Growth Centre; IRISS; the Kauffman Foundation; the Murthy Family; the Knowledge for Change Trust Fund; the National Science Foundation; the Toulouse Network for Information Technology; and the World Bank. 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, Kusha Goyal, Shruti Rangarajan, Jitendra Satpute, Shreyan Sarkar, and Ashutosh Tyagi, and the research support of Troy Smith. We thank our formal discussants Susantu Basu, Ray Fisman, Naushad Forbes, Vojislov Maksimovic, Ramada Nada, Paul Romer, and Steve Tadelis, as well as seminar audiences at the AEA, Barcelona GSE, Berkeley, BREAD, Boston University, Chicago, Columbia, Cornell, the EBRD, Harvard Business School, IESE, Katholieke Universiteit Leuven, Kellogg, the LSE, Maryland, the NBER, NYU, PACDEV, Stanford, TNIT, Toronto, UBC, UCL, UCLA, UCSC, Wharton, and the World Bank for comments.

I. Introduction

Economists have long puzzled over why there are such astonishing 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 per cent productivity spreads between the 10th and 90th 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' affects the productivity of inputs goes back at least to Walker (1887) and is central to the Lucas (1978) model of firm size. 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 adopt 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.[1] Recent work, however, has focused on specific management practices which can be measured, taught in business schools and recommended by consultants. Examples of these practices include key principles of Toyota's 'lean manufacturing', such as quality control procedures, inventory management, and human resource management. A growing literature measures many such practices and finds large variations across establishments and a strong association between these practices and higher productivity and profitability.[2]

This paper provides the first experimental evidence on the importance of management practices in large firms. The experiment takes large, multi-plant Indian textile firms and randomly allocates their plants to treatment and control groups. Treatment plants received five months of extensive management consulting from a large international consulting firm. This consulting diagnosed opportunities for improvement in a canonical set of management practices during the first month, followed by four months of intensive support for the 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 output. The result was an increase in productivity of 11 per cent and an increase in annual profitability of about $230,000. Firms also spread these management improvements from their treatment plants to other plants they owned, providing revealed preference evidence on their beneficial impact.

Given these results, the natural question is why firms had not previously adopted these practices. Our evidence suggests that informational constraints were an important factor. Firms were often not aware of the existence of many modern management practices, like inventory norms and standard operating procedures, or did not appreciate how these could improve performance. For example, many firms claimed their quality was as good as other local firms and so did not need to introduce a quality control process.

We also find two other major impacts of better management practices. First, owners delegated greater decision making power over hiring, investment and pay to their plant managers. This happened in large part because the improved collection and dissemination of information that was part of the change process enabled owners to monitor their plant managers better. As a result, owners felt more comfortable delegating.

1 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'.

2 See for example, Osterman (1994), Huselid and Becker (1996), MacDuffie (1995), Ichniowski, Shaw and Prennushi (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.

Second, the extensive data collection and processing requirements of modern management led to a rapid increase in computer use. For example, installing quality control systems requires firms to record individual quality defects and then analyze these by shift, loom, and design. So modern management appears 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. Our experiment provides some evidence on the role of modern management in driving SBTC.[3]

The major challenge of our experiment is the small cross-sectional sample size. We have data on only 28 plants across 17 firms. To address concerns over statistical inference in small samples we implement permutations tests that have exact finite sample size. We also exploit our large time series of around 100 weeks of data per plant by using estimators that rely on large T (rather than large N) asymptotics. We believe these approaches are useful for addressing sample concerns in our paper, and also potentially for other field experiments where the data has a small cross-section but long time series.

This paper relates to several strands of literature. First, there is the long literature showing large productivity differences across plants in dozens of countries. From the outset this literature has attributed much of these spreads to differences in management practices (Mundlak, 1961), but problems in measurement and identification have made this hard to confirm (Syverson, 2010). This productivity dispersion appears even larger in developing countries (Banerjee and Duflo, 2005, Hsieh and Klenow, 2009). Despite this, there are still few experiments on productivity in firms (McKenzie, 2010a) and none involving large multi-plant firms.

Second, our paper builds on the literature on the management practices of firms. There has been a long debate between the 'best-practice' view that some management practices are universally good so that all firms would benefit from adopting these (Taylor, 1911) and the 'contingency view' that every firm is already adopting optimal practices but these differ firm by firm (eg, 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 different explanations and resulting in little consensus in the management literature.[4] This paper provides experimental evidence that a core set of best practices do exist, at least in one industry.

Third, the paper links to the large theoretical literature on the organization of firms. These papers generally emphasize optimal decentralization as driven either by minimizing learning and information processing costs or by optimizing incentives.[5] But the empirical evidence on decentralization is limited, focusing primarily on de-layering in large publicly traded US firms (Rajan and Wulf, 2006).

Fourth, the paper contributes to the literature on Information Technology (IT) and productivity. A growing body of work has examined the relationship between technology and productivity, emphasizing both the direct productivity impact of IT and also its complementarity with modern management and organizational practices (eg, Bresnahan et al. 2002 and Bartel et al. 2007). But again the evidence has focused on survey data rather than experimental data. Our experimental evidence suggests one route for computers to affect productivity is by facilitating better management practices, and this occurs simultaneously with the decentralization of 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 basic business training and advice in micro- and small enterprises. This research has found significant effects of some forms of training on performance in smaller firms, supporting our results on in larger firms.

[3] See, for example, the survey in Autor, Katz and Kearney (2008).
[4] See, for example, the surveys in Delery and Doty (1996) and Bloom and Van Reenen (2010).
[5] See the recent reviews in Garicano and Van Zandt (2010), Mookherjee (2010) and Gibbons and Roberts (2010).

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 labor productivity that is only 15 per cent of US productivity (McKinsey Global Institute, 2001). While average productivity is 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) surveys of manufacturing firms in the US and India. The Bloom and Van Reenen (BVR) methodology scores firms from 1 (worst practices) to 5 (best practices) on specific management practices related to monitoring, targets and incentives. Aggregating 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 these management practice scores for a sample of 751 randomly chosen US manufacturing firms with 100 to 5,000 employees and the second panel for similarly sized Indian ones. The results reveal a thick tail of badly run Indian firms, leading to a 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 not to set and monitor clear targets for performance, and they do not explicitly link pay or promotion with performance. The scores for Brazil and China in the third panel, with an average of 2.67, are similar, suggesting that Indian firms are broadly representative of large firms in 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 per cent of manufacturing employment. The fourth panel shows the management scores for the 232 textile firms in the BVR 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, shirts 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 developing countries.

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

The sample firms were randomly chosen from the population of all publicly and privately owned textile firms in Maharashtra, based on lists provided by the Ministry of Corporate Affairs.[6] We restricted attention to firms with between 100 to 1,000 employees to focus on larger firms but avoided multinationals. Geographically we focused on firms in the towns of Tarapur and Umbergaon (the largest two textile towns in the area) since this reduced the travel time for the consultants. This yielded a sample of 66 potential subject firms.

All of these 66 firms were then contacted by telephone by our partnering international consulting firm. They offered free consulting, funded by Stanford University and the World Bank, as part of a management research project. We paid for the consulting services to ensure that we controlled the intervention and 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.

[6] 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 notice.

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.[7] We compared these program firms with the 49 non-program firms and found no significant differences in observables.[8]

The experimental firms have typically been in operation for 20 years and all are family-owned. They all produce fabric for the domestic market, and some also export. 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 per cent by employment and the top 5 per cent by sales,[9] and compared to India manufacturing in the top 1 per cent by both employment and sales (Hsieh and Klenow, 2010). Hence, these are large manufacturing firms.[10]

These firms are complex organizations, with a median of two plants per firm (plus a head office in Mumbai) and four reporting levels from the shop-floor to the managing director. In all the firms, the managing director is the largest shareholder, and all directors are family members. One firm is publicly quoted on the Mumbai Stock Exchange, although more than 50 per cent of the equity is held by the managing director and his father.

In Exhibits (1) to (7) in the Appendix 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. Each plant site involves several multi-story buildings (Exhibit 1). The plants operate a continuous production process that runs constantly (Exhibit 2). The factories' floors were rather disorganized (Exhibits 3 and 4), and their yarn and spare-parts inventory stores lacked any formalized storage systems (Exhibits 5 and 6).

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 the plants as the easiest way to rapidly change plant-level management. We selected the consulting firm using an open tender. The winner was a large international management consultancy which is headquartered in the US 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 their Mumbai office. These consultants were educated at leading 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.[11] However, 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 was US$1.3 million, approximately $75,000 per treatment plant and $20,000 per control plant. Note this is very different from what the firms themselves would pay for this consulting, which would be probably

---

7 The main reasons we were given for refusing free consulting were 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 these firms were suspicious of the offer, given many firms in India have tax and regulatory irregularities.

8 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 groups on management practices using the BVR scores, and found they were almost identical (difference of 0.031, p-value 0.859).

9 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.

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

11 At the bottom of the consulting quality distribution in India consultants are cheaper, but their quality is poor. At the top end, rates are similar to those in the US because international consulting companies target multinationals and employ consultants that are often US or European educated and have access to international labor markets.

about $250,000. The reason for our much cheaper costs per plant is that, because it was a research project, the consultancy charged us pro-bono rates (50 per cent of commercial rates), provided free partner time and enjoyed economies of scale working across multiple plants.

While the intervention offered high-quality management consulting, 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 field experiments, 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 industry experience, the consultants identified 38 key practices on which to focus. These 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:

• Factory Operations: Regular maintenance of machines and recording the reasons for breakdowns to learn from failures. Keeping the factory floor tidy to reduce accidents and ease the movement of materials.

• Quality control: Recording quality defects by type, analyzing these records daily, and formalizing procedures to address defects to prevent them recurring.

• Inventory: Recording yarn stocks on a daily basis, with optimal inventory levels defined and stock monitored against these. Yarn sorted, labeled and stored in the warehouse by type and color, and this information logged onto a computer.

• Human-resource management: Performance-based incentive system for workers and managers. Job descriptions defined for all workers and managers.

• Sales and order management: Tracking production on an order-wise basis to prioritize customer orders by delivery deadline. Using design-wise efficiency analysis so pricing can be based on design (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 changes in management practices as a result of the consulting intervention.[12] 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 extracting historical data from existing records. For example, to facilitate quality monitoring on a daily basis, a single metric, termed the Quality Defects Index (QDI), was constructed as a severity-weighted average of the major types of defects. At the end of the diagnostic phase the consulting firm provided each plant with a detailed analysis of its current management practices and performance. This phase involved about 15 days of consulting time per plant.

[12] 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.

The second step was a four month implementation phase given only to the treatment plants. In this phase, the consulting firm followed up on the diagnostic report to help introduce as many of the 38 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 readily be carried out by employees. For example, one of the practices was daily meetings for management to review production and quality data. The consultant attended these meetings for the first few weeks to help the managers run them, provided feedback on how to run future meetings, and adjusted their design. This phase also involved about 15 days a month of consulting time per plant.

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 all treatment and control plants. In return for the firms' continuing to provide this data, the consultants provided some light consulting advice to both the treatment and control plants. This phase involved about 1.5 days a month of consulting time per plant.

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

Cross-sectional sample size: We worked with 17 firms. We considered hiring cheaper local consultants and providing more limited consulting to a sample of several hundred plants in more locations. 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 per cent as noted in section II.B) and retention (100 per cent, 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 that this led to a small cross-sectional sample size. We discuss the estimation issues this generates in section III.D below.

Treatment and control plants: The 17 firms in our sample had 28 plants. Due to manpower constraints we could collect detailed performance data from only 20 plants, so we designated 20 plants as 'experimental' plants and randomly picked six 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.[13] The remaining eight plants were then the 'non-experimental plants': three in control firms and five in treatment firms. These non-experimental plants did not themselves receive consulting services, but data on their management practices and organizational and IT outcomes were collected in bi-monthly visits.

Timing: The consulting intervention was executed in three waves because of the capacity constraint of the six-person consulting team. The first wave started in September 2008 with four treatment plants. In April 2009 a second wave of ten treatment plants was initiated, and in July 2009 the diagnostic phase for the six control plants was carried out. Firm records usually allowed us to collect data going back to a common starting point of April 2008.

[13] Treatment and control plants were never in the same firms. The 6 control plants were randomly selected first, and then the 14 treatment firms randomly selected from the remaining 11 firms which did not have a control plant.

We started with a small first wave because we expected the intervention process to get easier over time due to accumulated 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 implementation waves.[14] 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 issues. A first potential concern is whether the sample size is too small to identify significant impacts. A second is what type of statistical inference is appropriate given the sample size. 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 took to address them.

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. 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 time dummies control for most external shocks. Fourth, we collected weekly data, which provides high-frequency observations over the course of the treatment and the use of these repeated measures can dramatically reduce the sample size needed to detect a given treatment effect (McKenzie, 2010b). Finally, the intervention was intensive, leading to large treatment effects – for example, the point estimate for the reduction in quality defects was over 50 per cent.

Statistical inference: A second concern is over using statistical tests which rely on asymptotic arguments in the N dimension to justify the normal approximation. We use three alternatives to address this concern. First, we use firm-clustered bootstrap standard errors (Cameron et al, 2008). Second, we implement permutation procedures (for both the Intent to Treat (ITT) and Instrumental Variables estimators) that have exact finite sample size and so do not rely upon asymptotic approximations. Third, we exploit our large T sample to implement procedures that rely upon asymptotic approximations along the time dimension (with a fixed N).

Permutation Tests: Permutation tests use the fact that order statistics are sufficient and complete statistics to derive critical values for test procedures. We first implement this for the null hypothesis of no treatment effect against the two sided alternative for the ITT parameter. This calculates the ITT coefficient for every possible combination of 11 treatment firms out of our 17 total firms (we run this at the firm level to allow for firm-level correlations in errors). Once this is calculated for the 12,376 possible treatment assignments (17 choose 11), the 2.5 per cent and 97.5 per cent 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 per cent level. Permutation tests for the IV estimator are more complex, involving implementing a procedure based on Greevy et al. (2004) and Andrews and Marmer (2008) (see Appendix B).

T-asymptotic clustered standard errors: An alternative approach is to use asymptotic estimators that exploit the large time dimension for each firm. To do this we use the recent results by Ibramigov and Mueller (2009) to implement a t-statistic based estimator that is

[14] Each wave had a one-day kick-off meeting involving presentations from senior partners from the consulting firm. This helped impress the firms with the expertise of the consulting firm and highlighted the potential for performance improvements. Since this meeting involved a project outline, and we did not tell firms about the existence of treatment and control groups, we could not mix the groups in the meetings.

robust to substantial heterogeneity across firms as well as to considerable autocorrelation across observations within a firm. This approach requires estimating the parameter of interest separately for each treatment firm and then treating the resultant set of 11 estimates as a draw from a t distribution with ten degrees of freedom (see Appendix B). Such a procedure is valid in the sense of having correct size (for fixed N) so long as the time dimension is large enough that the estimate 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 how representative it is 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 its largest industry (textiles). Comparing our sample to the population of large (100 to 5,000 employee) firms in India, both overall and in textiles, suggests that our small sample is at least broadly representative in terms of management practices (see Figure 1). In section V.D we also report results on a plant-by-plant basis to further demonstrate the results are not driven by any particular plant outlier. While we have a small sample, the results are relatively stable across the individual sample plants.

III.E. The potential conflict of interest in having the consulting firm measuring performance

A final design challenge was the potential for a conflict of interest in having our consulting firm measuring the performance of the experimental firms. To address this about every other month one of the research team visited the firms in India, meeting with the firms' directors and presenting in detail the quality, inventory and output data the consultants had sent us. This was not only a useful way to initiate discussions on the impact of the experiment, but also important for confirming the data we were receiving reflected reality. Moreover, when visiting the factories we could visually confirm whether the interventions had led to the reorganization of the factory floor, reduced inventory and improved quality control.

IV. The impact on management practices

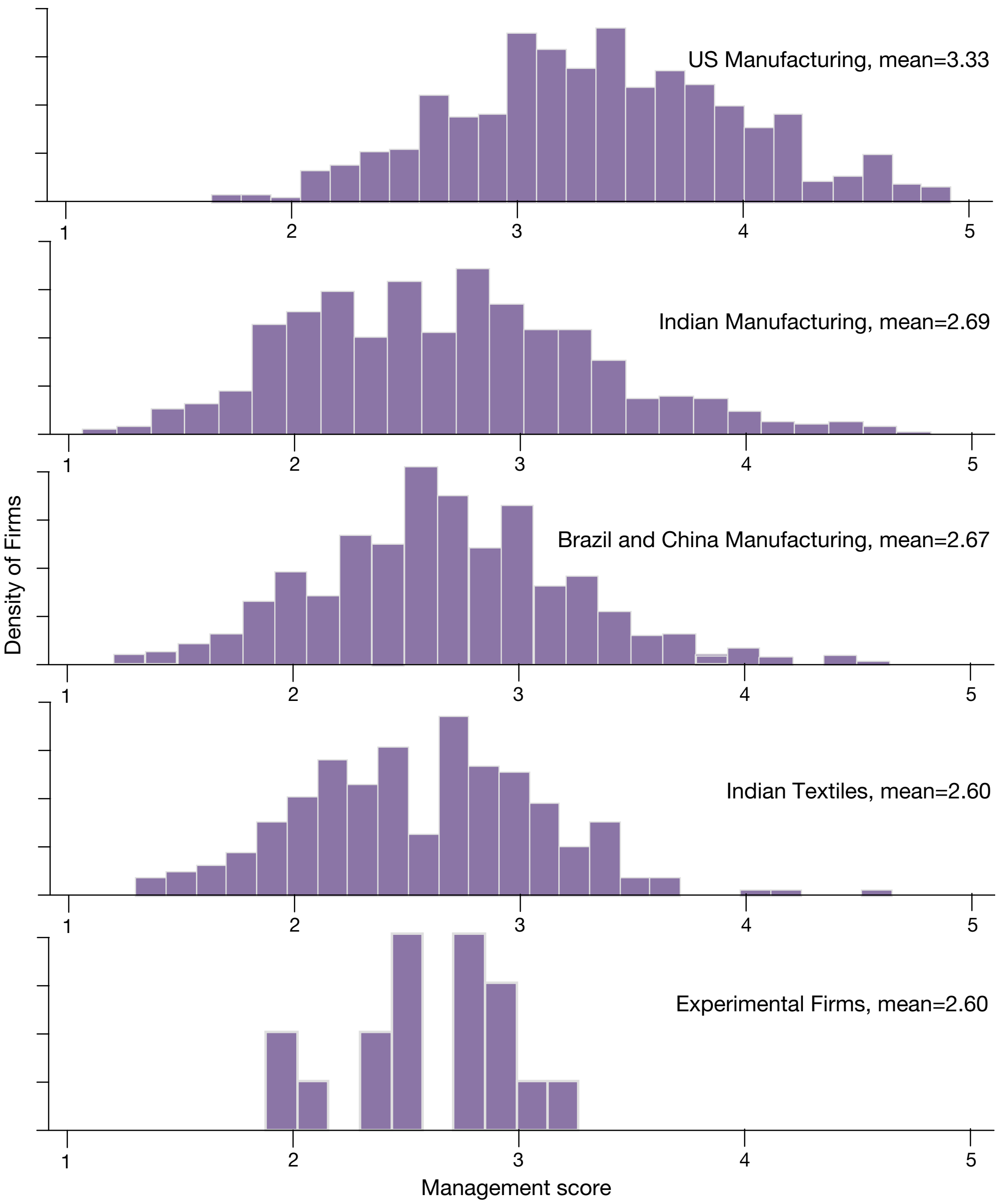
In Figure 2 we plot the average management practice adoption of the 38 practices for the 14 treatment plants, the six control plants, and the eight non-experimental plants. This data is shown at two month intervals before and after the diagnostic phase. Data from the diagnostic phase onwards was compiled 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, all plants started off with low baseline adoption rates of the set of 38 management practices.[15] Among the 28 individual plants the initial adoption rates varied from a low of 7.9 per cent to a high of 55.3 per cent, so that even the best managed plant in the group had just over half of the key textile-manufacturing practices in place. 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.

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

[15] The pre-treatment 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).

Figure 1: Management practice scores across countries



Notes: Management practice histograms using Bloom and Van Reenen (2007) methodology. Double-blind surveys used to evaluate firms' monitoring, targets and operations. Scores from 1 (worst practice) to 5 (best practice). Samples are 695 US firms, 620 Indian firms, 1083 Brazilian and Chinese firms, 232 Indian textile firms and 17 experimental firms

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

Fourth, the control plants, which were given only the 1 month diagnostic, increased their adoption of these management practices, but by only 12 per cent on average. This is substantially less than the increase in adoption in the treatment firms, indicating that the four months of the implementation phase were important in changing management practices. 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 in the treatment firms also saw a substantial increase in the adoption of management practices. In these five plants the adoption rates increased by 17.5 per cent. This increase occurred because the owners of the treatment firms copied the new practices from their experimental plants over to their other plants.

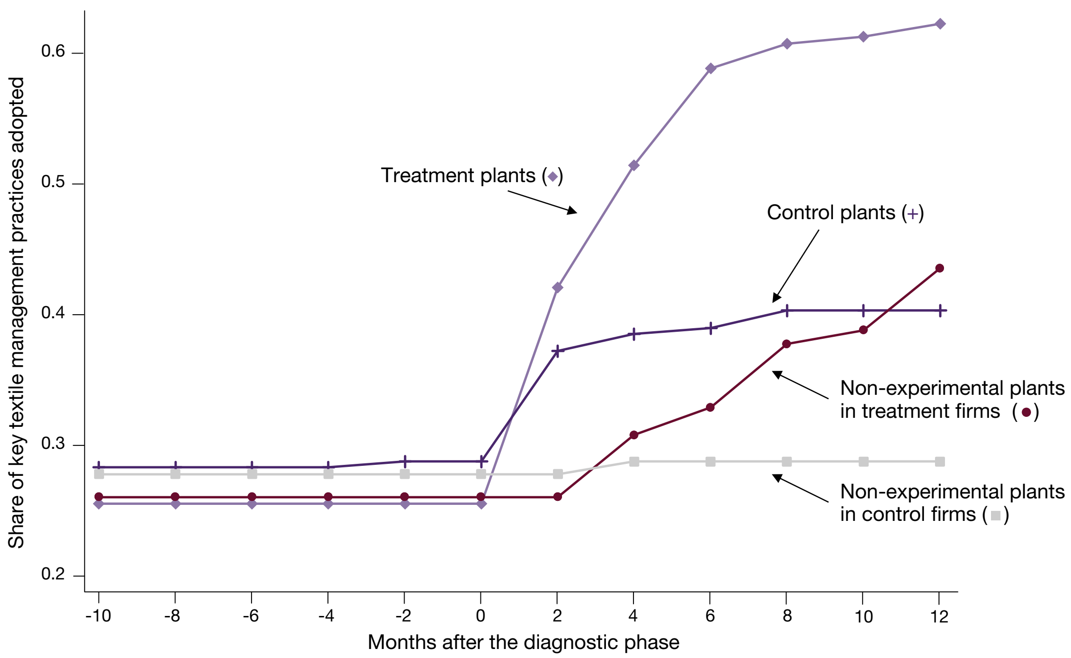
V. The impact of management on performance

Previous work has shown a strong correlation between management practices and firm performance in the cross-section, with a few papers (eg, Ichniowski et al. 1998) also showing this in the panel.[16] Our unique panel data on management practices and plant level performance, coupled with the experiment, enables us to examine the extent to which these relations are causal. We begin with a panel fixed-effects specification:

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

where OUTCOME will be one of the key performance metrics of quality, inventory and output. The concern is that management practices are not exogenous to the outcomes that are being assessed, even in changes. For example, a firm may start monitoring quality only when it starts 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 product quality as part of a major upgrade of workers and equipment, in which case we would misattribute quality improvements from better capital and labor to better management.

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 (diamond symbol), six control plants (plus symbol), the five non-experimental plants in the treatment firms which the consultants did not provide any direct consulting assistance to (round symbol) and the three non-experimental plants in the control firms (square symbol). Scores range from 0 (if none of the group of plants have adopted any of the 38 management practices) to one (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.

[16] Note that most papers using repeated surveys have found no significant panel linkage between management practices and performance (Cappelli and Neumark (2001) and Black and Lynch (2004)).

To overcome this endogeneity problem, we instrument the management practice score with log (1+weeks since the implementation phase began)[17]. We use this logarithmic form because of the concave adoption path of management practices shown in Figure 2, with the results robust to alternative functional form specifications such as linear or quadratic. The exclusion restriction is that the intervention affected the outcome of interest only 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 management practices in their recommendations to firms, and firms did not buy new equipment or hire new labor as a result of the intervention during the period of our study. 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 IV will consistently estimate the marginal effect of improvements in management practices, telling us how much management matters for the average plant participating in the study. However, if the effects of better management are heterogeneous, then the IV estimator will consistently estimate 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 better management if instead the plants that change management only when free consulting is provided are 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 the management practice score ranging from 26.3 to 60 percentage points. 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 to be a parameter of policy interest, since if governments are to employ policies to try to 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 which improved management practices via the following equation:

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

where TREATi,t is a 1/0 variable for whether plants have started the implementation phase or not. The parameter *c* then gives the ITT, which is the average impact of the intervention in the treated plants compared to the control plants.

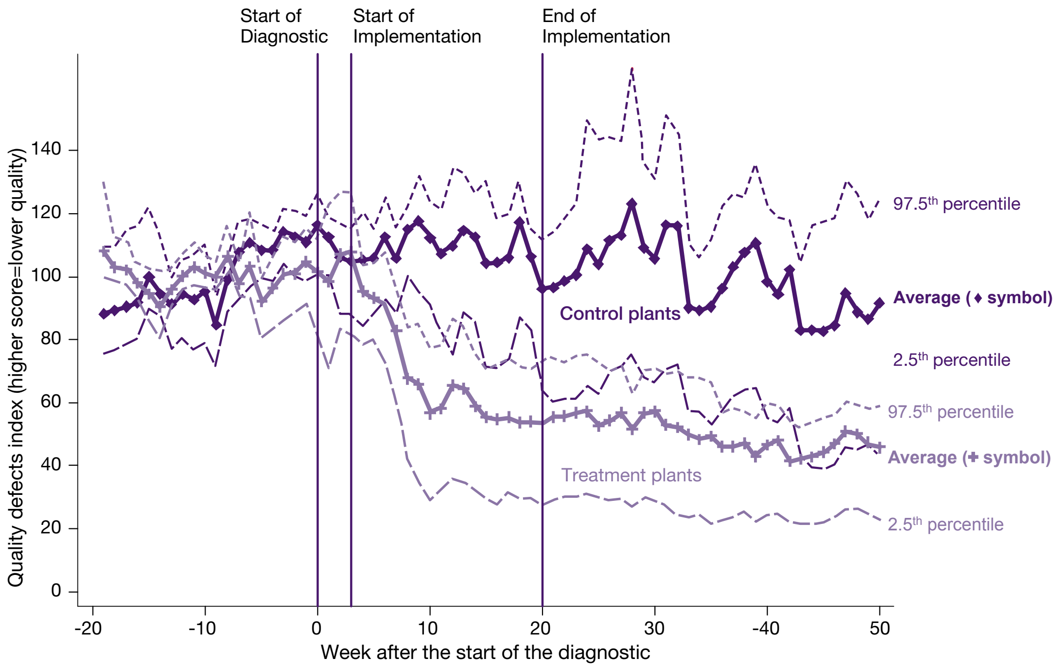
V.A Quality

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.[18] 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.

[17] Note that this is defined as zero for control plants and for treatment plants pre-implementation.

[18] 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.

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



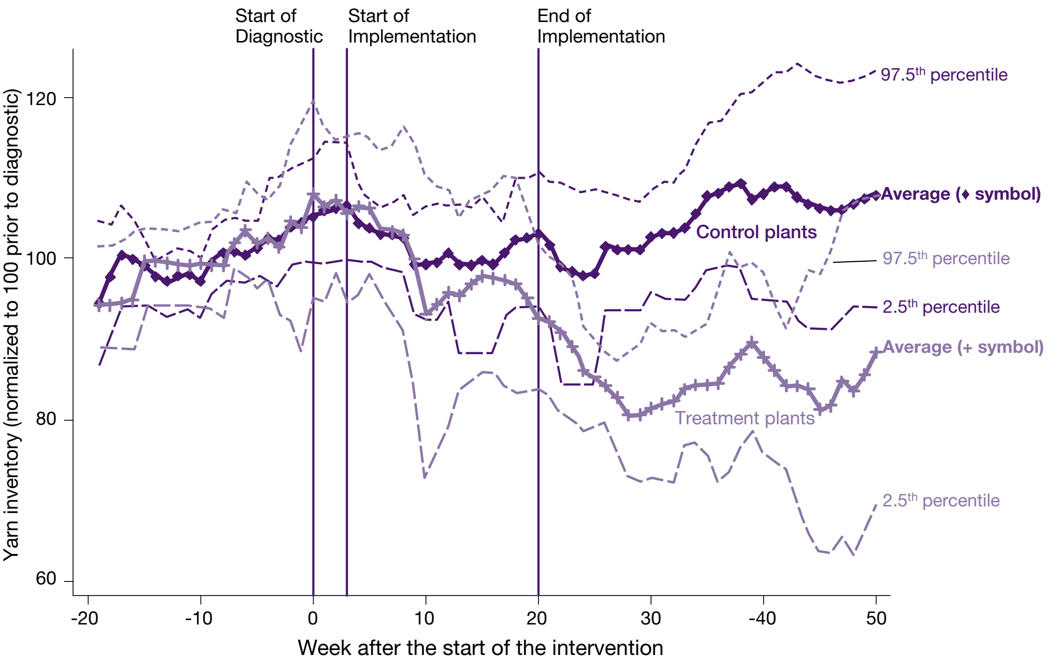
Notes: Displays the average weekly quality defects index, which is a weighted index of quality defects, so a higher score means lower quality. This is plotted for the 14 treatment plants (+ symbols) and the 6 control plants (♦ symbols). Values normalized so both series have an average of 100 prior to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

It is clear the treatment plants started to reduce their QDI scores (ie, improve quality) significantly and rapidly from about week five onwards, which was the beginning of the implementation phase following the initial one 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.

Table 2 in columns (1) to (4) examines whether management practices improve quality using regression analysis. In column (1) we present the fixed-effects OLS results which regresses the weekly log(QDI) score on plant level management practices, plant fixed effects, and a set of weekly time dummies. The standard errors are bootstrap clustered at the firm level to allow for any correlation across different experimental plants within the same firm. The -0.561 coefficient implies that increasing the adoption of management practices by 1 percentage point would be associated with about a 0.6 per cent reduction in defects, although this is not statistically significant.

In Table 2 column (2) we report the first stage from using the experimental intervention to identify the causal impact of better management on quality. The coefficient on log cumulative treatment is extremely significant, reflecting the fact that the intervention substantially increased the adoption of management practices. In column (3) we report the second stage, finding a significant point estimate of -2.028, suggesting that increasing the practice adoption rate by 1 percentage point would lead to a reduction in quality defects of about 2 per cent. The large rise in the point estimate from the OLS to the IV estimator suggests firms may be endogenously adopting better management practices when their quality starts to deteriorate. There was anecdotal evidence for the latter, in that the consulting firm reported plants with worsening quality were often the most keen to implement the new management practices because of their concern over quality problems. This has some conceptual similarities with the broader empirical literature showing that tough times – measured by higher competition – raises productivity (eg, Syverson 2004a), presumably in part because firms respond by improving management.

Figure 4: Yarn inventory for the treatment and control plants



Notes: Displays the average weekly quality defects index, which is a weighted index of quality defects, so a Notes: Displays the weekly average yarn inventory plotted for 12 treatment plants (+ symbols) and the six 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. Two treatment plants maintain no on-site yarn inventory.

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 mechanism (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 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 (4) we look at the ITT, which is the average reduction in the defects index after the intervention in the treatment plants versus the control plants. We see a 32 per cent (= exp(-.386)-1) fall in the QDI index, meaning the intervention cut quality defects by about a third.

At the foot of table 2 we also present our Ibramigov-Mueller (IM) and permutation 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 per cent level (zero is outside the 95 per cent confidence intervals). For the standard permutation tests the ITT is again significant at the 5 per cent level (the p-value is 0.0168), as are the IV-permutation tests.

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 four 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,[19] selling dead yarn stock, introducing restocking norms for future purchases, and monitoring inventory on a daily basis, the firms reduced their inventories. But this took time as the reduction in inventories primarily arose from lowering stocking norms and using old yarn for new products.

Table 2 columns (5) to (7) shows the regression results for log of raw material (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 (5) we present the fixed-effects result which regresses the weekly yarn on the plant level management practices, plant fixed-effects, and a set of weekly time dummies. The coefficient of -0.639 says that increasing management practices adoption rates by 1 percentage point would be associated with a yarn inventory reduction of about 0.6 per cent. In Table 2, column (6), we see the impact of management instrumented with the intervention displays a point estimate of -0.929, somewhat higher than the FE estimates in column (1).[20] Again, the IV estimator is higher than the OLS estimator, suggesting that the adoption of better management practices may be endogenous (or at least downward biased by measurement error). In column (7) we see the intervention causes an average reduction in yarn inventory of (exp(-.179)-1=) 16.4%.

These numbers are substantial but not unprecedented. Japanese automotive firms achieved much greater reductions in inventory levels (as well as quality improvements) from the adoption of lean manufacturing technology. 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 interval for the IV estimator finds the coefficient significant at the 5 per cent level. However, the IV permutation tests cannot exclude zero. Looking at the ITT coefficient, we see that under IM the results are significant at the 10 per cent 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[21]. 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 (8) to (10) in table 2 we look at this in a regression setting with plant and time dummies. In column (8) the OLS coefficient of 0.127 implies increasing the adoption of management practices by 1 percentage point would be associated with about a 0.1 per cent increase in output. In column (9), we see the impact of management instrumented with the intervention displays a higher significant point estimate of 0.346. 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 (10) we look at the ITT and see a point estimate of 0.056, implying a 5.4 per cent increase in output (exp(0.056)-1), although this only significant at the 11 per cent level.[22] Looking at the small-sample standard errors we find the IM and permutation tests are all significant at the 5 per cent or 10 per cent level.

There are several reasons for these increases in output. Undertaking routine maintenance of the looms reduces breakdowns. Collecting and monitoring the breakdown data also helps highlight looms, shifts, designs and yarn-types that are associated with more breakdowns. Visual displays around the factory floor together with the incentive schemes motivate workers to improve operating efficiency. Finally, keeping the factory floor clean and tidy reduces the number

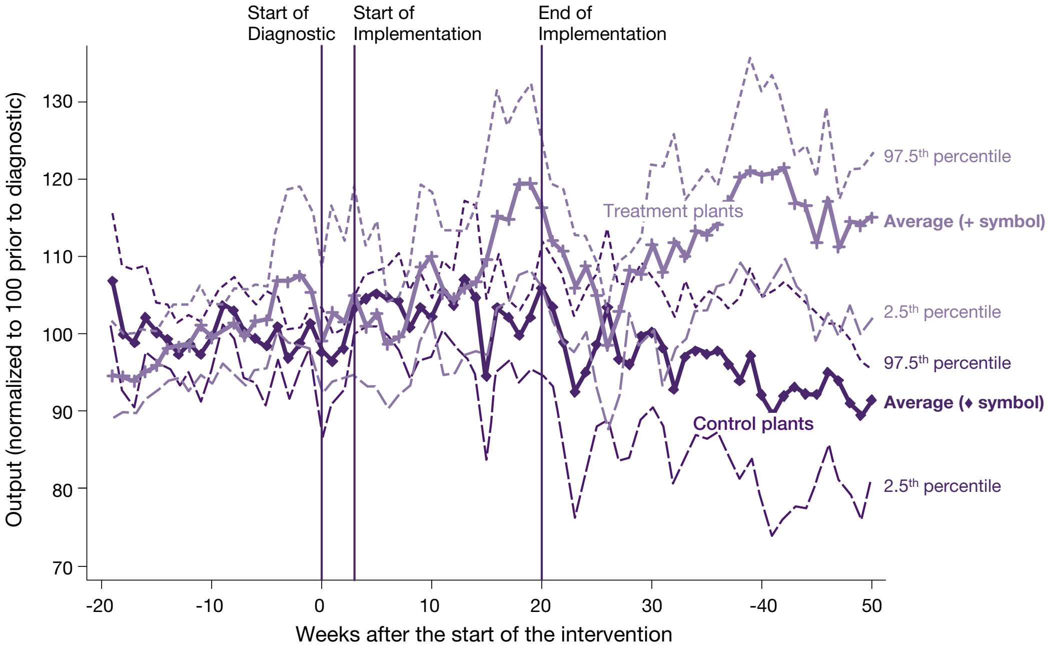
[19] Shade cards comprise a few inches of sample yarn, plus information on its color, thickness and material. These are sent to the design teams in Mumbai who use these to design new products using the surplus yarn.

[20] We do not report the IV first-stage as this is very similar to the first stage for quality shown in column (2).

[21] A production pick is a single crossing of the shuttle, representing the weaving of one thread of weft yarn.

[22] The IV is significant (and not the ITT) because the first stage of the IV uses log(cumulative treatment) rather than the binary 1/0 treatment variable, with the former more correlated with the gradual improvement in performance. Running the reduced-form for log(output) returns a coefficient (s.e.) of 0.028 (0.009) on log(cumulative treatment).

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.

of untoward incidents like tools falling into machines or factory fires. 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 Results by plant

We can also examine the difference in quality, inventory and output after treatment on a plant by plant basis. Figure 6 plots the histograms of the before-after changes in our performance measures for the treatment and control plants. No outliers are driving these differences, with all treatment plants improving their quality (top-left plot), nine of the treatment plants improving their inventory (top-right plot) and all treatment plants improving their output (bottom left plot). In comparison the control plants appear to be fairly randomly distributed around the zero impact point. We can also test the statistical difference of these changes between the two groups, and find the p-value on the difference in differences is 0.035 for quality, is 0.096 for inventory and 0.010 for output.[23]

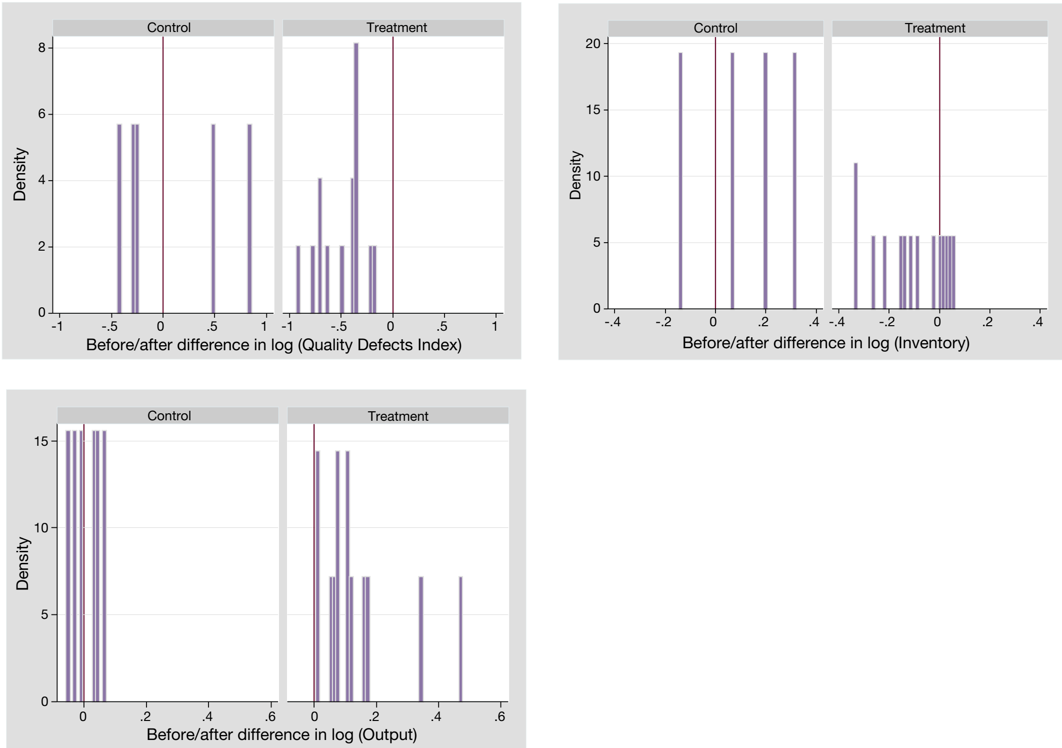
V.E Are the improvements in performance due to Hawthorne effects?

Hawthorne effects are named after a series experiments carried out at the Hawthorne Works in the 1920s and 1930s. The results apparently showed that just running experiments and collecting data can improve performance, raising concerns that our results could be spurious.

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 consulting during the 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

[23] Formally, we test this by regressing the 20 plant level differences on a 1/0 dummy variable for being a treatment firm, and report the p-value on that dummy, clustering at the parent firm level.

Figure 6 Plant level changes in performance



Notes: Displays the histogram of plant by plant changes in log (Quality Defects Index), log (Inventory) and log (Output) between the post and pre treatment periods.

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 extensively copying these practices out to their other non-experimental plants (see Figure 2).

VI. The impact of management practices on organization and computerization

VI.A The impact of management practices on firm organization

Although our interventions were never intended to directly change the treatment firms' organizational design, theory gave us some reason to believe that organizational changes might follow as a result of better management practices due to changes in the 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. However, this literature does not lead to clear-cut predictions about the effects of increased availability of information to managers. On the one hand, models of hierarchy as specialization in knowledge acquisition (like Garicano, 2000) 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. But while the theoretical literature is large, the empirical literature is very limited.

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 spent 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 survey and Table A4 for descriptive statistics). These questions and scoring were based on the survey methodology in Bloom, Sadun and Van Reenen (2009b), 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 Director's 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 dramatically reduces their commute) or at the factory (if it needs direct management from them).

To combine all eight decentralization measures into one index we took the first principal component, which we called the decentralization index. We found changes in this index were strongly and significantly correlated with changes in management across firms, as better management led to more decentralization. Table 3 looks at this in a regression format:

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

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 per cent (the mean change for the treatment group) is associated with a 0.55 standard-deviation change increase in decentralization. So typically this would mean the owner reduces his factory visits from daily to three times a week, while also letting the plant manager make hiring decisions for weavers, award small weaver bonuses, and plan the weekly maintenance schedule. In column (2) we run the IV estimation, using the log(1+weeks since the implementation phase began) as the instrument, and again find a positive and significant impact. Finally, in column (3) we report a positive ITT.

The consultants provided no advice on delegation and decentralization. It occurred in large part because the better monitoring of the factory operations allowed owners to delegate more decisions without fear of being exploited (the monitoring channel in the principal-agent group of organizational theories). For example, with daily inventory, quality and output data it is harder for the factory manager to steal inventory or output without detection by the owner.

VI.B The impact of management practices on computerization

A major topic over the last decade has been the relationship between IT and productivity. A growing literature finds that the productivity impact of IT is substantially larger than its cost share (eg, Bresnahan, Brynjolfsson and Hitt, 2002). The literature argues 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 because management and organizational practices are typically an unmeasured residual.[24] But none of this literature has any experimental evidence.

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 Enterprise

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

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 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 Tables A3 for the 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 table A4 it is readily apparent that as firms adopted more modern management practices they significantly increased the computerization of their operations. Table 3 looks at this in a regression format:

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

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 use a week, a rise of over 100 per cent 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. The exclusion restriction here is that the consulting intervention did not directly change computerization, apart from its effect through the management practices. The consultants were not told to discuss computerization apart from its use in implementing the management practices, and in our own discussions with the owners we did not come across cases where they mentioned the consultants discussing computerization for other reasons. 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.

These finding also relate to another major IT literature that 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 for example 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 skilled-biased, in that computer users in India are relatively skilled due to the need for literacy and numeracy. As a result modern management practices are a skill-biased technology, driving both the use of computers and the demand for skilled workers.

VII. Why do badly managed firms exist?

Given the evidence in section (IV) the obvious question is whether these management changes also increased profitability and productivity, and, if so, why they were 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. We could not obtain accounting data on these firms' profits and losses. Public accounts data are available only with a lag of 2-3 years at the firm level (rather than plant, which is what we would want), and in our interviews with firm owners they told us they under-report profits to avoid tax and also move profits to years when they want a loan (to have proof of income). When asked for their internal accounts the firms were evasive and would not provide them, beyond occasional comments that profits were in the range of $0.5m to $1m per year.[25] So we estimated the changes from the quality, inventory and efficiency improvements. Our methodology is simple: for example, if an 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 firm's total wage bill. These estimates are medium-run because, for example, it takes a few months for the firms to reduce their mending manpower.

[25] It is not even clear if firms actually keep correct records of their profits given the risk these could find their way to the tax authorities. For example, any employee that discovered these could use these to blackmail the firm.

These estimates for increases in profits are potentially biased. There is a downward bias because we take firms' initial 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 (eg, Milgrom and Roberts, 1990). However, the intervention time-horizon was too short to change many of the complementary human-resource practices. The estimates are upward biased if the firms backslide on the management changes once the consultants leave.

To estimate the net increase in profit for these improvements in management practices we also need to calculate the costs of implementing these changes (ignoring for now any costs of consulting). These costs were small, averaging less than $3,000 per firm.[26] So given the $250,000 this consulting would have cost these firms, this implies about a 90 per cent one-year rate of return.

Productivity: We estimate a total increase in productivity of 11.1 per cent, detailed in Table A5. Our methodology is again very simple, assuming a constant-returns-to-scale Cobb-Douglas production function $Y=AL^{\alpha}K^{1-\alpha}$ where Y is value-added (output – materials and energy costs), L is hours of work and K is the net capital stock. Using this we can back out changes in productivity after estimating changes in output and inputs. So, for example, reducing the yarn inventory by 16.4 per cent lowers capital by 1.3 per cent (yarn is 8% of the capital stock), increasing productivity by 0.6 per cent (capital has a factor share of 0.42). Our estimated productivity impact will also be subject to a number of the biases discussed above for profitability.

VII.B. Why are firms badly managed?

Given the evidence in section (VII.A) 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 the non-adoption of any of the 38 practices in each plant. To do this consistently we developed a flow-chart (Appendix Exhibit 7) which runs through a series of questions to understand the root cause for the non-adoption of each practice. The consultants collected this data from discussions with owners, managers, and workers, plus their own observations.

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.[27] They would then ask if the plant was aware of this practice, which in the example of recording quality typically was the case. 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 that 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, it was very poor by international standards. 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 the owner appeared to have incorrect information on the cost-benefit calculation.

The overall results for non-adoption of management practices are tabulated in Table 4, for the treatment plants, control plants and the non-experimental plants. This is tabulated at two-month intervals starting the month before the intervention. The rows report the different reasons for non-adoption as a percentage of all practices. From the table several results are apparent. First,

[26] About $35 of extra labor to help organize the stock rooms and factory floor, $200 on plastic display boards, $200 for extra yarn racking, $1,000 on rewards, and $1,000 for extra computer equipment (this is bought second hand).

[27] 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.

a major initial barrier to the adoption of these practices was a lack of information about their existence. About 15 per cent 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. Many of these are derived from the Japanese-inspired lean manufacturing revolution and are now standard across Europe, Japan and the US.[28]

Second, another major initial barrier was incorrect information, in that firms had heard of the 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. They preferred to keep their machines in operation until they broke down, and then repair them. This accounted for slightly over 45 per cent of the initial non-adoption of practices.

Third, as the intervention progressed the lack of information constraint was rapidly overcome. However, the incorrect information constraints were harder to address. This was because the owners had 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 these initial practice changes, the owners increasingly trusted them and would adopt more of the recommendations, like performance incentives for managers.[29]

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. A major reason is that the owners were severely time constrained, working an average of 68 hours per week already. There was also evidence of procrastination in that some owners would defer on taking quick decisions. This matches up with the evidence on procrastination in other contexts, for example African farmers investing in fertilizer (Duflo, Kremer and Robinson, 2009).

Finally, somewhat surprisingly, we did not find evidence for the direct impact of capital constraints, which are a significant obstacle to the expansion of micro-enterprises (eg, De Mel et al., 2008). Our evidence suggested that these large firms were not cash-constrained, at least for tangible investments. We collected data on all the investments for our 17 firms over the period August 2008 until August 2010 and found the firms invested a mean (median) of $880,000 ($140,000). For example, several of the firms were adding machines or opening new factories, apparently often financed by bank loans. Certainly, this scale of investment suggests that investment on the scale of $2,000 (the first-year costs of these management changes, ignoring the consultants' fees) is unlikely to be directly impeded by financial constraints.

Of course financial constraints could impede hiring international consultants. The market cost of our free consulting would be at least $250,000, and as an intangible investment it would be difficult to collateralize. Hence, while financial constraints do not appear to directly block the implantation of better management practices, they may hinder firms' ability to improve their management using external consultants. On the other hand, our estimates of the return on hiring consultants to improve management practices suggest profitability in just over one year.

[28] This ignorance of best practices seems to be common in many developing contexts, for example in pineapple farming in Ghana (Conley and Udry, 2010).

[29] These sticky priors highlight one reason why management practices appear to change slowly. The anecdotal evidence from private equity and consulting is that firms typically need between 18 months to 3 years to execute a turn around.

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 derived from limited managerial span of control. In every firm in our sample only members of the owning family have positions with major decision-making power over finance, purchasing, operations or employment. Non-family members are given only lower-level managerial positions with authority only over basic day-to-day activities. The principal reason is that family members do not trust non-family members. For example, they are concerned if they let their plant managers procure yarn they may do so 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 civil courts. A compounding reason for the inability to decentralize in Indian firms is bad management practices, as this means the owners cannot keep good track of materials and finance, so may not even able to identify mismanagement or theft within their firms.[30]

As a result of this inability to delegate, firms can expand beyond the size that can be managed by a single director only if other family members are available to share directorial duties. Thus, an important predictor of firm size was the number of male family members of the owners. In particular, 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 between employment and the average management score of 0.223. 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 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 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 textile plants. Since other industries in India – like software, construction and real estate – are growing rapidly the attractiveness of new investment in textile manufacturing is relatively limited (even our firms were often taking cash from their textile businesses to invest in other businesses).

Finally, a 50 per cent tariff on fabric imports insulates Indian textile firms against Chinese competition. Hence, the equilibrium appears to be that, with Indian wage rates being extremely low, firms can survive with poor management practices. Because spans of control are constrained, productive firms are limited from expanding, and so do not drive out badly run firms. And because entry is limited new firms do not enter rapidly. 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.

[30] Another compounding factor is none of these 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 and were often poorly equipped to take on extra responsibilities. In contrast, Indian software and finance firms that have grown management beyond the founding families place a huge emphasis on development and training. (see also Banerjee and Duflo (2000)).

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? A primary reason is that these firms are not aware they are badly managed, as illustrated in Table 4. 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 receptive. Of course consulting firms could go further and offer to provide free advice in return for an ex post profit-sharing deal. But monitoring this would be extremely hard, given the firms' desire to conceal profits from the tax authorities. Moreover, the client firm in such an arrangement might worry that the consultant would twist its efforts to increase short-term profits at the expense of long-term profits.

VIII. Conclusions

Management does matter. We implemented a randomized experiment that provided managerial consulting services to textile plants in India. This experiment led to improvements in basic management practices, with plants adopting lean manufacturing techniques that have been standard for decades in the developed world. These improvements in management practices led to improvements in product quality, reductions in inventory and increased efficiency, raising profitability and productivity. Firms also delegated more decisions because the improved informational flow from adopting modern management practices enabled the owners to reduce their oversight of plant operations. At the same time computer use increased, 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? Certainly we do not want to advocate free consulting, given its extremely high cost. But our results do suggest that, first, knowledge transference from multinationals would be very helpful. Indeed, many of the consultants working for the international consulting firm hired by our project had worked for multinationals in India, learning from their manufacturing management processes. Yet a variety of legal, institutional, and infrastructure barriers have limited multinational expansion within India. Abolishing tariffs could also help, as Indian firms would be driven to improve management practices to survive against lower cost imports from countries like China. Second, our results also suggest that a weak legal environment has limited the scope for well-managed firms to grow. Improving the legal environment should encourage productivity-enhancing reallocation, helping to drive out badly managed firms. Finally, our results suggest that firms were not implementing best practices on their own because of lack of information and knowledge. This suggests that training programs for basic operations management, like inventory and quality control, could be helpful.

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	5,560	5,130	2,260	13,000	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 A1 in the experimental plants. **Annual sales ($m)** and **Current assets ($m)** are both in 2009 US $million values, exchanged at 50 rupees = 1 US Dollar. **Daily mtrs, experimental plants** reports the daily meters of fabric woven in the experiment plants. Note that about 3.5 meters is required for a full suit with jacket and trousers, so the mean plant produces enough for about 1600 suits daily. **Age of experimental plant (years)** reports the age of the plant for the experimental plants. **Raw materials inventory** is the stock of yarn per intervention. **Operating efficiency** is the percentage of the time the machines are producing fabric. **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 modern management practices on plant performance

Dependent Variable	Quality defects	Management	Quality defects	Quality defects	Inventory	Inventory	Inventory	Output	Output	Output
Specification	OLS	IV	IV	ITT	OLS	IV	ITT	OLS	IV	ITT
		1st stage	2nd stage			2nd stage			2nd stage	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Managementi,t	-0.561		-2.028**		-0.639***	-0.929**		0.127	0.346**	
Adoption of management practices	(0.440)		(1.013)		(0.242)	(0.386)		(0.099)	(0.147)	
Cumulative treatmenti,t		0.088***								
Week since start of intervention		(0.018)								
Interventioni,t				-0.386**			-0.179**			0.056
Intervention stage initiated				(0.162)			(0.089)			(0.034)
Instrument			**Cumulative treatment**			**Cumulative treatment**			**Cumulative treatment**	
Small sample robustness										
Ibragimov-Mueller (95% CI)	(-4.46,-0.53)		(-5.03,-0.98)	(-0.69,-0.38)	(-0.81,-0.09)	(-0.84,-0.02)	(-0.17,0.02)	(0.22,0.86)	(-.08,2.25)	(0.05,0.26)
(90%CI)	(-4.09,-0.90)		(-4.65,-1.36)	(-0.66,-0.41)	(-0.75,-0.16)	(-0.77,-0.10)	(-0.15,0.00)	(0.28,0.80)	(0.13,2.03)	(0.07,0.24)
Permutation Test I (p-value)				0.02			0.14			0.10
IV Permutation Tests (95% CI)			(-6.05,-.06)			(-2.64,0.66)			(0.07,0.62)	
(90% CI)			(-6.00,-0.23)			(-2.04,0.52)			(0.12,0.54)	
Time FEs	113	113	113	113	113	113	113	114	114	114
Plant FEs	20	20	20	20	18	18	18	20	20	20
Observations	1,732	1,732	1,732	1,732	1,977	1,977	1,977	2312	2312	2,312

Notes: All regressions use a full set of plant and calendar week dummies. Standard errors bootstrap clustered at the firm level. **Quality defects** 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 is the log of the tons of yarn inventory in the plant. **Output** is the log of the weaving production picks. Management is the adoption share of the 38 management practices listed in table A1. **Intervention** is a plant level indicator taking a value of 1 after the implementation phase has started at a treatment plant and zero otherwise. **Cumulative treatment** is the log of one plus the cumulative count of the weeks of since beginning the implementation phase in each plant, and zero otherwise. **OLS** reports results with plant estimations. **IV** reports the results where the management variable has been instrumented with log(1+ weeks since start of implementation phase). First stage results are only shown for quality as the first stage results for inventory and output are very similar. **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. **Small sample robustness** implements 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 per cent and 90 per cent 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 1,000 possible permutations (out of 12376) of treatment assignment. **IV-Permutation** tests implements a permutation test for the IV parameter using 1,000 possible permutations (out of 12376) of treatment assignment. These tests have exact finite sample size.

Table 3: The impact of modern management practices on organization and computerization

Dep. variable:		**Decentralization Index**			**Hours of computer use**			**Computerization index**	
Specification	OLS	IV	ITT	OLS	IV	ITT	OLS	IV	ITT
	(1)	(2)	(7)	(4)	(5)	(6)	(7)	(8)	(9)
$\text{Management}_{i,t}$	1.695***	1.837***		16.761***	23.272**		1.154***	1.497**	
Adoption of management practices	(0.420)	(0.535)		(3.457)	(6.708)		(0.338)	(0.555)	
$\text{Interventioni}_{i,t}$			0.360**			6.168**			0.403**
Intervention stage initiated			(0.164)			(2.163)			(0.148)
Instrument		**Cumulative treatment**			**Cumulative treatment**			**Cumulative treatment**	
Small sample robustness									
Permutation Test I (p-value)			0.06			0.02			0.08
IV Permutation Tests (95% CI)		(-0.74,2.37)			(-18.99,62.89)			(-1.19,3.89)	
(90% CI)		(-.34,2.16)			(-9.88,45.07)			(-0.48,2.98)	
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. **Cumulative treatment** is the log of one plus the cumulative count of the weeks since the start of the implementation phase 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. The **Small sample robustness** implements 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.

Table 4: 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 304 practices in the non-experimental plants (38 practices in 8 plants). Non adoption was monitored every other month using the tool shown in Exhibit 7, based on discussions with the firms' directors, managers, workers, plus regular consulting work in the factories.

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 2, which shows a reduction of quality defects of 32 per cent (exp(-0.386)-1), a reduction in inventory of 16.4 per cent (exp(-0.179)-1) and an increase in output of 5.4 per cent (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 in 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 per cent in the baseline, arising from cutting out defect areas and destroying and/or selling at a discount fabric with unfixable defects. Assume an 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 per cent calculated as interest charges of 15 per cent (average prime lending rate of 12 per cent over 2008-2010 plus 3 per cent as firm-size lending premium – see for example www.sme.icicibank.com/Business_WCF.aspx?pid), 3 per cent storage costs (rent, electricity, manpower and insurance) and 4 per cent 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 per cent, so the profit margin on increased efficiency is 37 per cent.

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 various econometric procedures we implemented to verify the robustness of our results. We first outline the Ibragimov-Mueller procedure and then briefly discuss the two permutation tests and refer the reader to the original papers for a more detailed discussion.

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 summarized as follows: Implement the estimation method (OLS, IV, ITT) on each treatment firm separately and obtain 11 firm-specific estimates. Note that we cannot do this for the control firms 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, after using the control firms to remove time period effects.

The procedure requires that the coefficient estimates from each entity are asymptotically independent and Gaussian (but can have different variances). In our case this would be justified by an asymptotics in T argument (recall we have about 110 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.

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 statistic as described in Greevy et al (2004). The reason for using this statistic is that the permutation test for the IV parameter is a generalization of this procedure and so it is 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

$$qi,j,t = \mathbb{I}(Z_i > Z_j \left(\mathbb{I}(Y_{i,t} > Y_{j,t}) - (Y_{i,t} < Y_{j,t})\right)$$

This variable takes on the values 0, 1 and -1. It is equal to zero if plant is a control or plant j is a treatment plant and 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

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

Under the null hypothesis of no treatment effect, the treatment outcomes should not be systematically larger than the control outcomes. Specifically, under the null hypothesis and conditional upon the order statistics, each possible candidate value of T has 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), 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 $Y_d = \tau + \beta d + e$ 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 practices that the firm has implemented. The maintained assumption is that the potential outcomes are independent of the instrument Z or equivalently (e, 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, $\tilde{Y} \equiv Y - \tau - \beta_0 D = e$ 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

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

Where we have replaced the response by the response subtracted by . Note that is consistently estimable under the null, so without loss of generality we can treat it as known. 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

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

and we form the statistic as

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

Where

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

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. Although the confidence set constructed in this manner need not be a single interval, in all our estimations, the confidence sets were single 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

Area	Specific practice	Pre-intervention *level*		Post-intervention *change*	
		Treatment	Control	Treatment	Control
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 (eg, 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 (ie, 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 (eg, 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 (ie, 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 (ie, 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 (ie, 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 (ie, 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:

<table>
<tr><td colspan="4">Question C1: ‘Does the plant have an Electronic resource planning system?’</td></tr>
<tr><td colspan="4">Question C2: ‘How many computers does the plant have?’</td></tr>
<tr><td colspan="4">Question C3: ‘How many of these computers are less than two years old’</td></tr>
<tr><td colspan="4">Question C4: ‘How many people in the factory typically use computers for at least ten minutes day?’</td></tr>
<tr><td colspan="4">Question C5: ‘How many cumulative hours per week are computers used in the plant’?</td></tr>
<tr><td colspan="4">Question C6: ‘Does the plant have an internet connection’</td></tr>
<tr><td colspan="4">Question C7: ‘Does the plant manager use email (for work purposes)?’</td></tr>
<tr><td colspan="4">Question C8: ‘Does the plant manager use email (for work purposes)?’</td></tr>
<tr><td colspan="4">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)</td></tr>
<tr><td></td><td>Score 1</td><td>Score 3</td><td>Score 5</td></tr>
<tr><td>Scoring grid:</td><td>Computers not used in operational performance management</td><td>Around 50% of operational performance metrics (efficiency, inventory, quality and output) are tracked and analyzed through computer/ERP generated reports.</td><td>All main operational performance metrics (efficiency, inventory, quality and output) are tracked and analyzed through computer/ERP generated reports.</td></tr>
</table>

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)	1,0357	1,000	35,000	10,434	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 median 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 2. Figure calculated for the median firm. See Appendix A for details of calculations for inventory carrying costs, fabric waste, repair manpower and factor shares.

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 seven days a week producing fabric from yarn, with four 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

Instrument not removed after use, blocking hallway.

Old warp beam, chairs and a desk obstructing the factory floor

Dirty and poorly maintained machines

Tools left on the floor after use

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

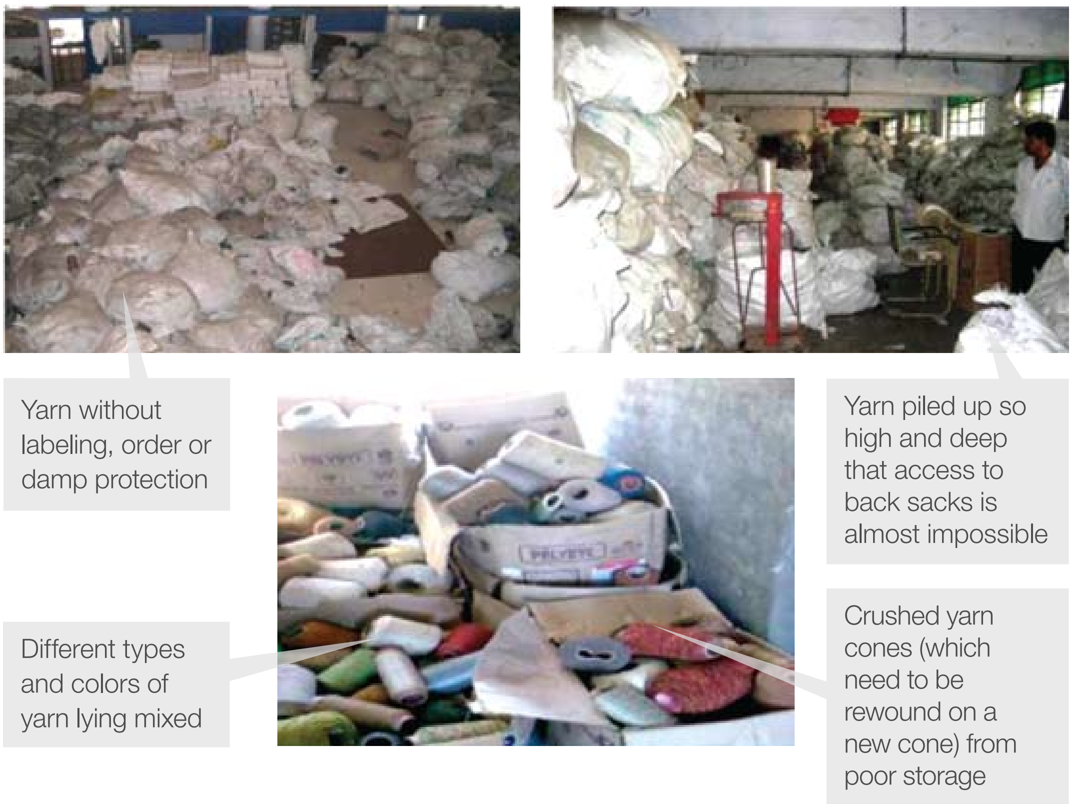


Exhibit 6: The parts stores were often disorganized and dirty

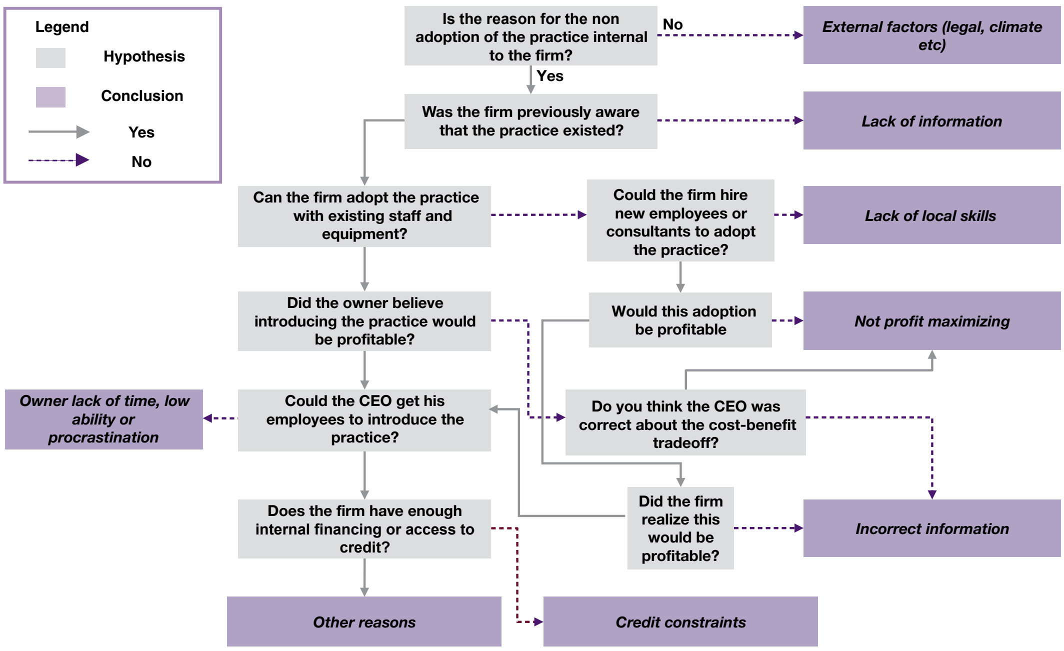

Spares without any labeling


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.

Bibliography

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

Autor, David, Lawrence Katz, and Mellissa Kearney (2008), 'Trends in US wage Inequality: Revising the Revisionists,' *Review of Economics and Statistics*, 90 (2), 300-323.

Banerjee, Abhijit and Esther Duflo (2000) 'Reputation Effects and the Limits of Contracting: a Study of the Indian Software Industry,' *Quarterly Journal of Economics*, 115(3), 989-1017.

Banerjee, Abhijit and Esther Duflo (2005) 'Growth Through the Lens of Development Economics', in Philippe Aghion and Stephen Durlauf (eds), *Handbook of Economic Growth, Volume 1*, Amsterdam.

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.

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 (2009a) 'Americans do IT Better: American Multinationals and the Productivity Miracle,' forthcoming *American Economic Review.*

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2009b) '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*.

Bresnahan, Timothy, Erik Brynjolfsson, and Lorin Hitt (2002) 'Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-level Evidence,' QJE, 117(1), 339-376.

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

Cameron, A 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 D Neumark (2001), 'Do 'High Performance' Work Practices Improve Establishment-Level Outcomes?' *Industrial and Labor Relations Review*, 54(4), 737-775.

Clark, Greg (1987), 'Why isn't the Whole World Developed?' *Journal of Economic History*, 141-173.

Conley, Tim and Christopher Udry, (2010), 'Learning about a New Technology: Pineapple in Ghana,' *American Economic Review*, 100(1), 35-69.

De Mel, Suresh, David McKenzie and Christopher Woodruff (2008), 'Returns to Capital in Microenterprises: Evidence from a Field Experiment,' *QJE*, 113(4), 1329-1372.

Delery, John and Harold Doty (1996), 'Modes of Theorizing in Strategic Human Resource Management: Test of Universalistic, Contingency and Configurational Performance Predictions', *Academy of Management Review*, 39(4), 802-835.

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

Duflo, Esther, Michael Kremer and Jonathan Robinson (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,' 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, Robert, Jeffrey H. Silber, Avital Cnaan and Paul R. Rosenbaum (2004), 'Randomization Inference with Imperfect Compliance in the ACE-inhibitor,' *Journal of the American Statistical Association*. 99(465), 7-15.

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(3), 400-422.

Ibragimov, R and UK Muller (2009), 't-statistic Based Correlation and Heterogeneity Robust Inference,' *Journal of Business and Economic Statistics*.

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.

Karlan, Dean and Martin Valdivia (2010) '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. *Handbook of Organizational Economics*.

Lucas, Robert (1978), 'On the Size Distribution of Business Firms,' *Bell Journal of Economics*, 508-23.

MacDuffie, John Paul (1995), 'Human Resource Bundles and Management Performance: Organizational Logic and Flexible Production Systems,' *Industrial and Labor Relations Review*, 48(2), 197-221.

McKenzie, David (2010a), 'Impact Assessments in Finance and Private Sector Development: What have we learned and what should we learn?' *World Bank Research Observer*, 25(2), 209-33.

McKenzie, David (2010b) 'Beyond Baseline and Follow-up: The Case for More T in Experiments,' Mimeo, World Bank.

McKinsey Global Institute (2001), *India: The Growth Imperative*.

Melitz, Marc (2003), 'The impact of Trade on Intra-industry Reallocations and Intra-industry Productivity,' *Econometrica*, 71(6), 1695-1725.

Milgrom, Paul and JohnRoberts (1990), 'The Economics of Modern Manufacturing: Technology, Strategy and Organization,' *American Economic Review*, 80 (3), 511-528.

Mookherjee, Dilip (2010), 'Incentives and Hierarchy,' in Robert Gibbons and John Roberts, editors, *Handbook of Organizational Economics*, Princeton University Press, forthcoming.

Mundlak, Yair (1961), 'Empirical Production Function Free of Management Bias,' *Journal of Farm Economics*, 43(1), 44-56.

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 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 (2004), 'Market Structure and Productivity: A Concrete Example', *Journal of Political Economy*, 112(6), 1181-1222.

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

Taylor, F (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), 265-288.

Womack, James, Daniel Jones and Daniel Roos (1991), *The Machine that Changed the World*, Harper Collins: New York.

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

About the International Growth Centre

The IGC offers independent advice on economic growth to governments of developing countries. Based at the London School of Economics and in partnership with Oxford University, the IGC is initiated and funded by the UK Department for International Development (DFID).

The IGC has active country programmes in Bangladesh, Ethiopia, Ghana, India, Pakistan, Sierra Leone, Tanzania, Mozambique, Zambia and Rwanda and supports over seventy individual research projects on issues of governance, human capital, agriculture, infrastructure, trade, firm capability, state capacity, macroeconomics and political economy.

The IGC is directed by a Steering Group consisting of an Executive Director (Gobind Nankani) in collaboration with a Deputy Executive Director (Mark Henstridge) and two Academic Directors, one from LSE (Robin Burgess) and one from Oxford University (Paul Collier). The Steering Group also includes Chang-Tai Hsieh from the University of Chicago, Timothy Besley at LSE and Stefan Dercon at Oxford University.

The organisational structure of the IGC spans a London hub, country offices in partner countries, a group of 10 research programmes with participation from academics in world-class institutions, a network of policy stakeholders in the developing world and a range of public, civil society and private sector partners.

Contact us

International Growth Centre
The London School of Economics and Political Science
4th Floor, Tower Two
Houghton Street
London WC2A 2AE
United Kingdom

General Office Tel: +44 (0)20 7955 6144

For enquiries about this paper, please contact Adam Green:
a.r.green@lse.ac.uk
+44 (0)20 7955 3665

Directed and Organised by

LSE
UNIVERSITY OF OXFORD

Initiated and funded by

UKaid from the Department for International Development

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about our work on our website
www.theigc.org

For media or communications enquiries, please contact
mail@theigc.org

Subscribe to our newsletter and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International Growth Centre

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid from the British people

Designed by soapbox.co.uk