Returns to On-the-job Soft Skills Training*

Achyuta Adhvaryu†
Namrata Kala‡
Anant Nyshadham§

December 2021

Abstract

We estimate productivity gains of 13.5 percent from on-the-job soft skills training among Indian garment workers. Productivity gains are most pronounced when trainees work on joint operations alongside other co-workers (particularly treated co-workers). Furthermore, they are mirrored among non-treated co-workers on the production line, consistent with gains being driven by improved teamwork and collaboration. Heterogeneous treatment effects indicate that improvements in the teamwork and collaboration skills of workers substitute for managerial attention, but that the training is complemented by the degree to which managers act autonomously to adjust production processes in response to issues raised by workers. Despite the large productivity gains and increased probability of promotion among treated workers, we find no effects on wages or worker retention, consistent with frictions in the low-wage labor market. Consequently, the net return to the firm was large: 256 percent eight months after program completion.

Keywords: soft skills, non-cognitive skills, worker training, productivity, India, teamwork, collaboration
JEL Codes: J24, M53, O15

*We are very grateful to Anant Ahuja, Chitra Ramdas, Raghuram Nayaka, Sudhakar Bheemarao, and Paul Ouseph for their coordination, enthusiasm, and guidance. Thanks to Dotti Hatcher, Lucien Chan, Noel Simpkin, and others at Gap, Inc. for their support and feedback on this work. We acknowledge funding from Private Enterprise Development in Low-Income Countries (PEDL) initiative. This research has benefited from discussions with Michael Boozer, Robert Gibbons, Paul Gertler, Markus Goldstein, Rocco Macchiavello, David McKenzie, Dilip Mookherjee, Claudia Olivetti, Antoinette Schoar, Tavneet Suri, Chris Udry, John Van Reenen, and Chris Woodruff, and seminar audiences at NBER, Penn, MIT, USC, Madrid, PEDL, IGC, World Bank, AEA, IPA, IFPRI, Columbia, Northeastern, NEUDC, U. Chicago, Michigan, McGill, Georgetown, UC Santa Barbara, UT Austin, and Copenhagen. Many thanks to Lavanya Garg, Robert Fletcher, and Aakash Mohpal for excellent research assistance. The views expressed herein do not represent PEDL, Gap, Inc., or Shahi Exports. All errors are our own.

†University of Michigan, BREAD, NBER; adhvaryu@umich.edu
‡MIT Sloan School of Management, CEPR, BREAD, NBER, JPAL; kala@mit.edu
§University of Michigan & NBER; nyshadha@umich.edu

1 Introduction

Soft skills – e.g., leadership, communication, teamwork and collaboration skills – are highly predictive of success in the labor market (Bassi et al., 2017; Borghans et al., 2008; Deming, 2015; Groh et al., 2015; Guerra et al., 2014; Heckman and Kautz, 2012; Heckman et al., 2006; Montalvao et al., 2017). Recent empirical evidence from personnel economics underscores that the nature and quality of interactions between co-workers is a strong determinant of productivity (Amodio and Martinez-Carrasco, 2018; Bandiera et al., 2010, 2013). Surveys of employers from around the world corroborate that soft skills are in great demand, but show that firms often struggle to find workers with high levels of these skills (Cunningham and Villaseñor, 2016).

It stands to reason then that employers might want to train workers in soft skills to offset an underemphasis on these skills in traditional schooling, particularly in developing countries like India which emphasize rote learning over collaborative work (Ajisuksmo and Vermunt, 1999; Kurtz, 1990; Kurtz et al., 1988).[1] But how malleable soft skills are in adulthood, and whether training programs that aim to increase the stock of these skills can indeed generate causal impacts on productivity, have only begun to be explored (Acevedo et al., 2017; Ashraf et al., 2017; Campos et al., 2017; Groh et al., 2012).[2] Moreover, when general training is delivered within the firm (as it often is[3]), it is imperative to know the firm's returns to training in addition to worker productivity effects. This impact, in turn, is governed by labor market structure.[4] That is, the employer's incentive to train workers in general skills depends on whether wages must rise proportionately with any increased productivity in order to retain trained workers. Since most soft skills are "general," the extent of labor market frictions thus likely polices the ability to deliver soft skills training through firms, even when training raises productivity.

The questions that motivate our study, then, are threefold. First, what are the causal impacts of on-the-job soft skills training on workplace productivity? Second, for which workers and under what conditions is productivity most impacted by training in soft skills, and what do these patterns tell us about which specific skills are most important for productivity in a manufacturing setting? Finally, when weighing impacts on wages and retention against productivity gains, does it pay for firms to provide on-the-job soft skills training to workers?

[1] As suggestive evidence of the potential value of group work practice, we use data from the Trends in International Mathematics and Science Study (TIMSS), which has been used in prior work to study the effect of pedagogical methods such as group work on outcomes like trust at the country level (Algan et al., 2013). We create a country-level average of a binary for whether each teacher reported working in groups in every or almost every lesson. This measure across 38 countries has a positive correlation with log GDP per capita of about 0.18.

[2] Studies from psychology and economics demonstrate that it is possible to inculcate soft skills in early childhood, via, for example, home-based stimulation and high quality preschool programs (Attanasio et al., 2014; Gertler et al., 2014; Grantham-McGregor et al., 1991; Ibarrarán et al., 2015). However, structural estimates of dynamic human capital accumulation models suggest that it may indeed be difficult to affect non-cognitive skill levels at later ages, particularly for those with low baseline stocks, due to dynamic complementarities (Aizer and Cunha, 2012; Cunha et al., 2010; Heckman and Mosso, 2014).

[3] See, e.g., Bassanini et al. (2007).

[4] In perfectly competitive markets, any firm that paid below marginal product would lose the newly trained workers to higher higher paying firms. As Becker (1964) noted, this implies that with perfect labor markets, even general training programs that generate large productivity returns may not be appealing investments for firms. On the other hand, if asymmetric information, slow employer learning, or search frictions play a role in the labor market, then the resulting wedge between workers' marginal products and their wages in equilibrium may create positive productivity rents from general training for firms (Acemoglu, 1997; Acemoglu and Pischke, 1998, 1999; Autor, 2001; Chang and Wang, 1996; Katz and Ziderman, 1990).

To answer these questions, we partnered with the largest ready-made garment export firm in India to evaluate an intensive, workplace-based soft skills training program. The initiative, which is named Personal Advancement and Career Enhancement (P.A.C.E.), aims to empower female garment workers via training in a broad variety of life skills, including modules on communication, time management, problem solving and decision-making, and effective teamwork. These skills are important inputs into production in the ready-made garments context. Workers produce in teams, and need effective communication, leadership, planning and problem-solving skills to both resolve throughput issues with other team members (e.g., identifying and working through bottlenecks in real time) and relay information in a productive way to supervisors (e.g., notifying them of machine malfunction, requesting help to complete operations or meet targets, etc.).

We conducted a randomized controlled trial (RCT) in five garment factories in urban Bengaluru, India. We enrolled female garment workers in a lottery for the chance to take part in the P.A.C.E. program and used a two-stage randomization procedure to assign workers to treatment. In the first stage, we randomized production lines to treatment. In the second stage, within treatment lines, we randomized workers who had enrolled in the lottery to either direct P.A.C.E. training or spillover treatment. We thus estimate treatment effects by comparing trained workers (on treatment lines) to control workers on control lines (who enrolled in the lottery but whose lines were assigned to control). We estimate spillovers by comparing control (untrained) workers on treatment lines to control workers on control lines.

Treated workers are more productive by 7.4 percentage points (13.5% higher than the control mean). Productivity gains are larger among workers with low stocks of leadership skills at baseline and are most pronounced when trainees work on operation teams alongside other co-workers (particularly treated co-workers). Untreated co-workers on the same lines as treated workers also exhibit large gains in productivity (with magnitudes more than 80% the size of the direct treatment effect). Taken together we interpret this evidence as indicating that the gains in productivity are driven largely by the teamwork and collaboration skills imparted by the curriculum.[5] Heterogeneous treatment effects by managerial quality suggest that improvements in the teamwork and collaboration skills of workers substitute for managerial attention toward production balancing and coordination responsibilities, but that the training is complemented by the degree to which managers act autonomously to adjust production processes in response to issues raised by workers.

Despite large and lasting productivity gains and an increased probability of promotion among treated workers, we find no treatment effects on wages or worker retention, consistent with the presence of labor market frictions that prevent workers from capturing more of the productivity rents from training (Acemoglu, 1997; Acemoglu and Pischke, 1999). Soft skills are largely unobserved in the hiring and wage-setting process in this context and therefore are not priced into the wage; this is consistent with hiring processes for frontline workers in other low-income country contexts (Bassi et al., 2017). This imperfect information (including potentially slow learning about higher productivity due to training) among both current and future potential employers likely generates the observed

[5] Consistent with this interpretation, despite the program being broad in scope and made up of 8 modules of differing content, we note that the largest productivity gains appear after the module covering teamwork and collaboration (Execution Excellence) is administered at the end of the curriculum.

difference in impacts on productivity as compared to wage.

We use our estimates of impacts on workplace outcomes along with program cost and accounting profit data to calculate the costs and benefits of the program to the firm. The net rate of return was 61% by the end of the program period. Eight months after program completion, due to post-program increases in productivity of both treated and spillover workers, the return climbed to 256%. These large returns are rationalized by the relatively low costs of the program combined with the accumulated effects on productivity, and are consistent with other recent interventions in similar settings (Menzel, 2017).

Our main contribution is to the study of soft skills in the labor market. We join a handful of recent studies that evaluate the causal impacts of soft skills training on economic outcomes (Acevedo et al., 2017; Ashraf et al., 2017; Campos et al., 2017; Edmonds et al., 2020; Groh et al., 2012; Schoar, 2014). These previous studies are largely focused on populations of unemployed (or not yet working-age) individuals, making the investigation of impacts on productivity in the workplace infeasible. We add to this work by studying training within the firm, which emphasizes estimating firms' returns and ties our work to the literature on the role of labor market frictions in firms' decisions to train their workers (Acemoglu, 1997; Acemoglu and Pischke, 1998, 1999; Autor, 2001).

Most importantly, we are able to directly estimate impacts on individual productivity, which is missing from previous work.[6] This individual-level analysis allows us to document for whom and under what conditions improved soft skills generate the largest impacts on productivity and, in doing so, to comment on *which* specific soft skills appear to matter. A few recent studies have shown that individuals exhibiting certain personality traits perform well on team tasks (Driskell et al., 2018; Weidmann and Deming, 2021). We build on this evidence by showing that productivity gains from soft skills training are most pronounced when teamwork and collaboration skills are most required. In this sense, our results highlight the productive value in the workplace of teamwork and collaboration skills, which are often underemphasized in traditional schooling systems like that in India (Ajisuksmo and Vermunt, 1999; Kurtz, 1990; Kurtz et al., 1988).

Previous work quantifying the productivity impacts of on-the-job training generally uses observational data on firms and workers in the United States and Western Europe (Barrett and O'Connell, 2001; Barron et al., 1999; Dearden et al., 2006; Konings and Vanormelingen, 2015; Mincer, 1962).[7] We add to this literature in three ways. First, we estimate causal effects by exploiting randomized assignment to training, which overcomes potential self-selection bias (Altonji and Spletzer, 1991; Bartel and Sicherman, 1998). Second, we estimate impacts on retention in addition to productivity; retention is crucial to understanding firms' overall returns to training but has not been examined thus far. Third, we carry out our experiment in a low-income country setting, where training frontline workers might have large potential gains given low levels of baseline skills. Indeed, as discussed above, we document that productivity gains from training are largest among workers with low baseline stocks of these skills.

[6] Campos et al. (2017) measure micro-enterprise profits, which of course are in part a function of productivity.

[7] These studies tend to find that training increases productivity, but there is disagreement on the magnitude of this increase (Blundell et al., 1999). Specifically, when endogeneity of training is accounted for (e.g., using matching methods), productivity returns become quite small (Goux and Maurin, 2000; Leuven and Oosterbeek, 2008).

2 Context, Program Details, and Experiment Design

2.1 Garment Manufacturing in Production Lines

Apparel is one of the largest export sectors in the world, and India is one of the world's largest producers of textile and garments. Garment manufacturing continues to be a labor-intensive process throughout the global industry. Women comprise the majority of the large workforce in garment factories, and new labor force entrants tend to be disproportionately female (Staritz, 2010).

Garments are sewn in production lines. Each line produces a single garment style at time for several weeks on average until the complete order for thousands of units of that garment style is completed and ready for shipment to the buyer. The average garment (e.g., a men's woven button-down shirt) will involve around 50 machine operations (e.g., left sleeve construction, right sleeve construction, right cuff, left cuff, cuff button hole, cuff button attachment, etc.), all together expected to take around 30-35 minutes in total to complete one entire garment. In our setting, production lines are made up of roughly 50-70 workers with almost every worker touching each garment produced by the line. This structure of course enables gains in efficiency (i.e., each worker specializes on a single operation, maximizing productivity by way of experience and momentum and minimizing extraneous movements between and recalibrations of machines). However, it also generates ways in which productivity depends crucially on communication and coordination amongst workers as well as between workers and supervisors charged with relieving bottlenecks and minimizing productivity imbalances so as to maximize line-level output of completed garments from the available labor.

Each worker spends more than 90% of their days on the "home line" to which they were assigned centrally by HR upon their hiring. Workers may be loaned out to other lines for the day, in response to idiosyncratic daily absenteeism across lines, or temporarily reassigned to other lines working on complex orders with short deadlines for important buyers. These instances account for the less than 10% of days for which a worker is observed working on a line other than her "home line" and are never initiated by the worker. Note, for the analysis, we statically assign a worker to her baseline production line used in the randomization. Managers' assignments to lines are fixed and do not vary over time in the data.

2.1.1 Teamwork in the Production Process

Most of the workers on the line are assigned to machines completing sewing operations (one person to a machine) to cover the 50 or so operations required to make a full garment. The remaining workers perform complementary tasks to sewing, such as folding or aligning and ironing a seam before it is fed into a machine. These "helper" assignments are usually reserved for a few critical operations (i.e., more technical operations requiring more skill or experience, taking longer to complete, and, as a result, most likely to cause bottlenecks in line-level output of completed garments). These critical operations are also often staffed with more than one machine operator to balance production rates with the rest of the operations on the line. We observe these instances in our setting and can study the degree to which the productivity impacts of the soft skills training are amplified when workers are assigned to work together on the same operations.

The line is also subdivided into smaller groups of operations that produce subsections of the garment (e.g., collars or sleeves). These groups are separated by "feeding points" at which the prepared materials for each subsection of the garment are fed in bundles (e.g., materials for 20 pockets or collars of the current shirt will be fed at one point and materials for 40 sleeves will be fed at the next point).[8] Completed sections of garments pass between machine operators in these bundles, are attached to each other in additional operations along the way, and emerge at the end of the line as completed garments.

2.1.2 Supervisors

Each of these features create smaller subdivisions within the production line which amplify the need for teamwork or effective collaboration among workers. Additionally, the coordination between workers within these subdivisions and balancing of productivity across these subdivisions makes critical the role of the line supervisor. Though production line supervisors are explicitly responsible for monitoring production and relieving bottlenecks as well as making process adjustments to address needs and concerns raised by workers, related studies from this setting have documented that they vary substantially in the effort they devote to these responsibilities and as a result in the productivity they enable on their lines (Adhvaryu et al., 2019a,b).

It stands to reason that when managers are less attentive to their responsibility to coordinate and balance between workers and segments of the line, the ability of the workers to communicate and collaborate effectively among themselves becomes more valuable. By the same token, the confidence a worker feels and skill she has in communicating her needs or issues (e.g., regarding operation assignments or machine calibration) is likely complemented by the degree to the which the manager feels willing and able to make adjustments to the production process without having to consult with upper-level managers. We study below how improving teamwork and collaboration among workers interacts with managerial quality along these dimensions.

2.2 Program Details

The Personal Advancement and Career Enhancement (P.A.C.E.) program was designed and first implemented by Gap Inc. for female garment workers in low-income contexts. The intervention we study involved the implementation of the P.A.C.E. program in five factories in the Bengaluru area which had not yet adopted the program. The goal of this 80-hour program was to improve life skills such as time management, effective communication, and problem-solving and decision-making for its trainees.

2.2.1 Content

The program began with an introductory ceremony for participants, trainers, and firm management. The core modules of the program were Communication, Problem Solving and Decision-Making, and Time and Stress Management. Additional modules included Financial Literacy; Legal Literacy and Social Entitlements; Water, Sanitation and Hygiene; and General and Reproductive Health, but were

[8]This feature likely further contributes to the need for effective teamwork in this setting, but we are unfortunately not able to identify the feeding points or subsections of the line in our data in order to analyze impacts within or across these sections.

not considered core modules.[9] Table A1 provides an overview of the topics covered in each module and the content of each module is described in greater detail in section A.2 of the Appendix.

In the final month of the program, a module called Execution Excellence had participants apply the skills taught in the three core modules (e.g., effective and appropriate communication, identifying problems and potential solutions, implementing learning-by-doing, goal-setting, planning time-use, and task prioritization) to a teamwork exercise. The time-bound simulation had participants plan and execute an imaginary garment order as a group while being observed by the trainers. The debriefing following the exercise and remaining sessions in the module centered on the importance of teamwork for project success and the need for complementarity between the efforts and work of team members. Though most of the modules included group exercises to apply the learned skills to workplace scenarios, the explicit emphasis on how all of these skills contributed to effective teamwork and collaboration and the central importance of teamwork itself was strongest in this module in the final month of the program.[10]

After all modules had been completed, there were two review sessions (3 hours in total) reiterating concepts from early modules and discussing how participants would apply their learning to professional situations. At the close of the program there was a graduation ceremony.

2.2.2 Implementation

Workers participated in two hours of training per week. One hour of workers' production time a week was allocated to the training program, and workers contributed one hour of their own time each week. Training sessions were conducted at the beginning of the production day in designated classroom spaces in the factories, with workers assigned to groups corresponding to different days of the work week. Production constraints required that each day's group be composed of workers from across production lines so as not to produce large, unbalanced absences from any one line in the first hour of any production day.

Due to holidays and festivals (which are times of high absenteeism), in-practice sessions were conducted somewhat more flexibly with respect to timing. Catch-up sessions were conducted for workers who were unable to attend a session. This flexibility is reflected in average attendance (of non-attrited workers) of the core program modules, which was very high, ranging between 94 and 99 percent. With these adjustments, overall program implementation took about 12 months: the introductory ceremony was in July 2013, training was conducted between July 2013 and June 2014, and the closing ceremony in July 2014.[11]

2.3 Experimental Design

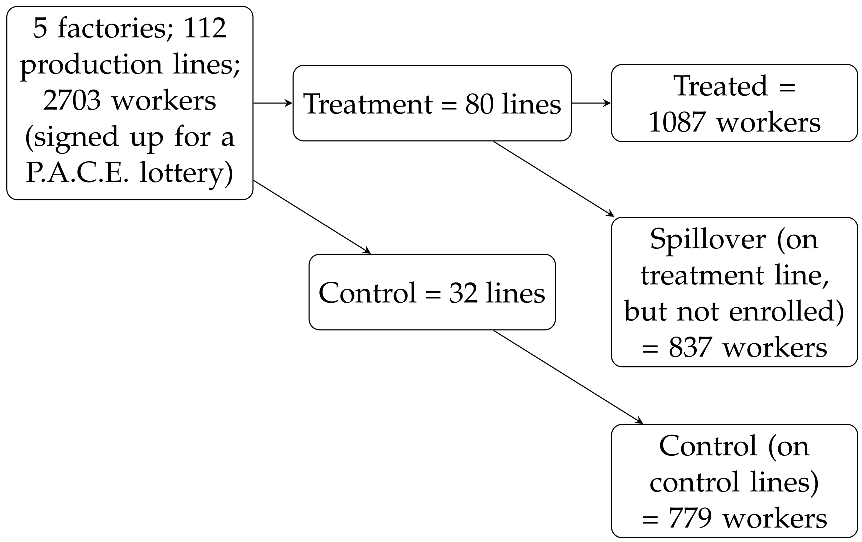
Participants were chosen from a pool of workers who expressed interest and committed to enroll in the program. The workers were informed that the training was oversubscribed and that a subset of work-

[9]Pre/post assessments were not conducted for ancillary modules such as sanitation.

[10]We note below that the heterogeneity in treatment effects, timing of productivity impacts, and spillover to untreated co-workers are all consistent with the teamwork skills emphasized in this Execution Excellence module contributing primarily to the treatment effects on productivity.

[11]See Appendix section A for more details regarding the program.

Figure 1: Experimental Design



Note: The retention rate in the worker-month panel is 65%.

ers would be chosen at random from a lottery to actually receive the training, with untreated workers granted the right to enroll in a later lottery for the next training batch.[12] Randomization was conducted at two levels: line level (stratified by factory, above- and below-median baseline efficiency, above- and below-median baseline attendance, and above- and below-median enrollment in the lottery), and then at the individual level within treatment lines. The five factories had 112 production lines in total. In the first stage of randomization, roughly two-thirds of production lines within each factory were randomized to treatment, yielding 80 treatment lines and 32 control lines across factories. In the second stage of randomization, within lines randomized to treatment, a fixed number of workers (13-14) from each treatment line were randomly chosen to take part in the P.A.C.E. program from the total set of workers who expressed interest by enrolling in the treatment lottery.[13]

Figure 1 presents a schematic diagram of the experimental design.[14] 2703 workers signed up for the treatment lottery, from which 1087 were chosen for treatment. Out of the 1616 untrained workers, 779 workers were in control lines, and the remainder, 837 workers, were in treatment lines. The former group (untrained workers in control lines) serves as the control. The latter group (untrained workers in treatment lines) is used to estimate treatment spillovers. Given the multifaceted nature of the program content, estimates of spillover effects on the productivity of untreated coworkers working alongside trained workers can help to inform the degree to which the program's emphasis on teamwork and collaboration skills was particularly valuable. Summary statistics and balance checks are discussed in

[12]Importantly, losers of the lottery were told that they would not necessarily receive the training in the next batch, nor would they be able to earn the right to be trained in any way, but rather that subsequent training batches would also be chosen at random via lottery.

[13]The decision to allocate a fixed number of workers to treatment per treatment line was due primarily to production constraints requiring a minimum manpower be present at all times during production hours.

[14]Additionally, Figure A1 in the Appendix presents the timeline of the experiment and data collection.

Section 3.4.[15]

3 Data

We leverage both administrative data from the factories and primary survey data to evaluate the program. Figure A2 presents an overview of the different data sources used in the evaluation, the frequency of data collection of each data type, and the availability of the data over time.

3.1 Production Data

Productivity data were collected using tablet computers assigned to each production line on the sewing floor. The employee in charge of collecting the data (the "production writer"), who was prior to our intervention charged with recording by hand on paper each machine operator's completed operations each hour for the line, was trained to input production data directly in the tablet computer instead. Importantly, from the perspective of the garment workers, production data were being recorded identically before, during, and after the intervention across treatment and control lines.[16]

3.1.1 Productivity and Operation Complexity

The key measure of productivity we study is efficiency. Efficiency is calculated as operations completed divided by the target quantity of operations per unit time. In order to calculate the worker-level daily mean of production from these observations, we average the efficiency of each worker over the course of the day (8 production hours).[17] It is worth emphasizing that due to the production process detailed in Section 2.1.1, although employees work on a production line, individual workers can increase productivity while others on the same line do not. At the worker-hour level, we define "pieces produced" as the number of garments that passed a workers' station by the end of that production hour. For example, if a worker was assigned to sew plackets onto shirt fronts, the number of shirt fronts at that workers' station that had completed placket attachment by the end of a given production hour would be recorded as that workers' pieces produced. The target quantity for a given operation is calculated using a measure of garment and operation complexity called the "standard allowable minute" (SAM). SAM is defined as the number of minutes required for a single garment of a particular style to be produced. That is, a garment style with a SAM of 30 is deemed to take half an hour to produce one complete garment. This measure at the line level is then decomposed into worker or operation specific increments. A line with 60 machine operators then would have an average worker-hourly SAM

[15] For the sake of brevity, we present only balance checks for treatment versus control workers, but balance holds across spillover versus control workers as well (results available upon request).

[16] Note that though productivity was being recorded prior to the program implementation, the worker level data was not kept prior to the introduction of the tablet computers for production writing but rather discarded after line-daily aggregate measures were input into the data server. Accordingly, line-daily aggregate data was all that was available at the time of treatment assignment, and as mentioned above, the first stage randomization of lines to treatment was stratified by line-level baseline efficiency.

[17] Completed operations recorded in the production data reflect only those which have passed quality checks, so our measure of efficiency actually reflects both quantity conditional on a minimum quality. In averaging across hourly quantities within the day, we expect that mis-measurement arising from re-worked (defective) pieces is minimized.

of 0.5. As the name suggests, this measure is standardized across the global garment industry and is drawn from an industrial engineering database.This measure may be amended to account for stylistic variations from the representative garment style in the database. Any amendments are explored and suggested by the sampling department, in which master tailors make samples of each specific style to be produced by lines on the sewing floor (for costing purposes). The target quantity for a given unit of time for a worker completing a particular operation is then calculated as the unit of time in minutes divided by the SAM. That is, the target quantity of pieces to be produced by a worker in an hour for an operation with a SAM of 0.5 will be $60/.5 = 120$. Hourly productivity data was available starting the month of treatment announcement. During the month of treatment announcement (June 2013) the tablets were introduced onto the production floors. Accordingly, June 2013 represents the pre-program baseline for all productivity analysis below.[18]

3.1.2 Other Variables

The production data include identifiers for the specific operation on which the worker is working as well as the garment style on which the line as a whole is working. In addition to the complexity of the operation to which a worker is assigned as measured by SAM and the achieved efficiency at the worker day level, these production data also allow us to calculate the number of workers and operations on a particular line on a given day, as well as the number and treatment status of other workers working together on the same operation. We use these additional variables to investigate heterogeneous productivity impacts below.

3.2 Human Resources Data: Attendance and Salary

We also obtain data on workers' daily data and monthly salary from 6 months prior to the training until 18 months after the completion of training. Data on demographic characteristics, attendance, tenure, and salary of workers are kept in a firm-managed database. The variables available in the demographic data include age, date on which the worker joined the firm, gender, native language, and education. Daily attendance data at the worker level includes whether a worker attended work on a given date, whether absence was authorized or not, and whether a worker was late to work on a given day (worker tardiness).

We also obtained monthly salary data which also indicates current grade level and designation (i.e., specific position title). Higher grade and/or designation workers are more technically skilled and more likely to be assigned technically complex operations (i.e., those with higher SAM, expected to take longer to complete) and the critical operations most likely to cause bottlenecks in line-level output of completed garments. Workers are compensated almost entirely by set monthly salaries. These salaries are benchmarked closely to the minimum wage, which in India varies by industry, state, urban zone, and "grade" (skill level). Accordingly, the main way that workers can earn higher pay is to be promoted to a higher grade (e.g., from a B to a B+ grade tailor) or designation (e.g., from helper to graded tailor or graded tailor to floater). Workers can request to have their grade reassessed

[18]The tablets were introduced for all lines in the five factories, so productivity was measured the same way for both treatment and control lines.

and/or supervisors can recommend workers for grade reassessment. These grade reassessments occur roughly annually at the same time for all workers who have been nominated or have requested one.[19]

Skill grades are defined for machine operating tailors with varying degrees of technical training. Some unskilled workers called helpers are ungraded, despite being included in the production line, and are primarily assigned to the complementary assignments mentioned above such as folding and ironing seems which do not require technical training. Above the highest grade tailor are additional ungraded designations such as floater. Workers with these designations still appear on the production line with targets and recorded efficiency, but no longer have grade levels recorded in their salary files. From these grade and designation variables, we construct an outcome variable for promotion taking value 1 if the grade or designation of the worker has changed upward since the start of the observation period up until that month's salary record, and 0 otherwise. We also study heterogeneous impacts on productivity by skill grade.

3.3 Survey Data

In addition to measuring workplace outcomes, a survey of 993 randomly chosen treated and control workers was conducted in June 2014, the month of program completion.[20] The survey covered, among other things, questions related to financial decisions (including savings and debt) and awareness of and participation in welfare programs (government or employer sponsored). It also measured personality characteristics (conscientiousness, extraversion, locus of control, perseverance, and self-sufficiency), mental health (hope/optimism, self-esteem, and the Kessler 10 module, which is used to diagnose moderate to severe psychological distress (Kessler et al., 2003)), and risk and time preferences elicited using lottery choices.[21] Finally, the survey covered workers' self-assessments relative to peers by asking them to imagine a six-step ladder with the lowest productivity workers on the lowest steps, and then asking them which step they would place themselves on; participation in skill development programs; production awards; and incentive programs on the job.

We also utilize data from a survey of production line supervisors meant to comprehensively capture various aspects of managerial quality. Adhvaryu et al. (2019b) estimate latent factors from the numerous measures in these survey data using a non-linear measurement system developed by Cunha et al. (2010). We study heterogeneity in productivity treatment effects and spillovers by the two factors capturing managerial style and practices: Attention and Autonomy. The Attention factor captures the degree of effort the manager puts forth in monitoring for and intervening to resolve productivity imbalances and coordination issues between workers.[22] The Autonomy factor captures the degree to

[19]In addition to a fixed monthly salary, workers are eligible for bonus pay for excess productivity assessed at the line level. In practice, the target productivity level for earning bonus pay is set extremely high, and workers rarely qualify for these bonuses as a result.

[20]Of the 993 surveyed, 403 were workers who underwent the soft skills training, 315 were control workers on control lines and the rest were untrained (control) workers on treated lines. We compare survey outcomes of treated workers in treated lines with those of control workers in control lines (N=363+258=621) to estimate the direct effects of the program, and compare outcomes of untrained workers in treated lines with control workers in control lines (total N=527) to estimate the indirect effects of the program.

[21]Risk and time preference modules were adapted from the Indonesian Family Life Survey. The other survey measures were measured using a rating scale of 5-10 statements measuring a particular outcome and assessing a workers' level of agreement with the statement. Survey questions are available on request.

[22]The Attention factor reflects contributions from five managerial practice measures: production monitoring frequency,

Table 1: Summary Statistics

	(1)		(2)		(3)	
	Control		Treated		Difference	
P.A.C.E. Treatment	**Control Workers in Control Lines**		**Treated Workers in Treatment Lines**			
Number of workers	779		1,087			
	Mean	SD	Mean	SD	Mean Difference	p value
Attendance Rate (Jan-May 2013)	0.898	0.117	0.903	0.103	0.005	0.380
High School	0.602	0.489	0.604	0.489	0.003	0.901
Years of Tenure	1.432	2.709	1.353	2.119	-0.079	0.500
Age	27.712	14.087	27.420	11.638	-0.292	0.637
1(Speaks Kannada)	0.657	1.560	0.671	1.156	0.014	0.834
High Skill Grade	0.685	0.840	0.702	0.679	0.017	0.645
log(Salary) (May 2013)	8.746	0.188	8.737	0.156	-0.009	0.258
Efficiency (Announcement Month)	0.586	0.587	0.556	0.426	-0.030	0.268
SAM (Announcement Month)	0.618	0.726	0.615	0.535	-0.003	0.928

Notes: Tests of differences calculated using errors clustered at the line level according to the experimental design.

which the manager's leadership style and decision-making practices reflect a willingness and ability to make production decisions and process adjustments without consulting upper level management.[23]

3.4 Summary Statistics and Balance Checks

Table 1 presents summary statistics of the main variables of interest, as well as balance checks for baseline values of attendance, high school completion, years of tenure with the firm, age, an indicator for median or above skill grade, and an indicator for speaking the local language (Kannada). Additionally, we check balance for several workplace outcomes: salary in the month before treatment announcement and productivity and operation complexity in the announcement month (the first month of observation for these outcomes).

Baseline variables for treated and control workers are balanced. The summary statistics and differences presented in Table 1 apply to the direct treatment comparison. Analogous balance checks for spillover comparisons are presented in Table C1 in the Appendix.[24]

4 Treatment Effects

The empirical analysis proceeds in several steps, beginning with testing the impact of the program on retention. Following this, we test for differences in productivity, pay, and promotion as well as in survey measures related to career advancement. We follow this section with a discussion of potential mechanisms in light of heterogeneity in treatment effects on productivity, spillover impacts on productivity, and interactions with managerial quality.

active personnel management, efforts to achieve production targets, communication, and issues motivating workers and overcoming resistance.

[23]The five measures informing the Autonomy factor captured leadership behaviors with respect to "initiating structure" and "consideration" (Stogdill and Coons, 1957), decentralized problem-solving practices adapted from the World Management Survey (Bloom and Van Reenen, 2007), and a self-assessment of managerial quality.

[24]The only significant difference is that spillover workers earn less than 2% less in salary at baseline than control workers.

4.1 Impacts on Retention and Removing Potential Bias from Selective Attrition

We estimate the following regression specification to test whether P.A.C.E. treatment impacts retention:

$$\begin{aligned} R_{wdmy} = \alpha_0 + \zeta_1 1[T_w] * 1[\text{Treatment Announced}]_{my} + \zeta_2 1[T_w] * 1[\text{During Treatment}]_{my} + \\ \zeta_3 1[T_w] * 1[\text{After Treatment}]_{my} + \psi_{ym} + \eta_w + \varepsilon_{wdmy} \end{aligned} \tag{1}$$

where the outcome is an indicator variable that takes the value 1 if worker w was retained on day d in month m and year y and 0 otherwise. The dummy variable $1[T_w]$ takes the value 1 if the worker is trained and on a treatment line and 0 if she is an untrained worker on a control line. It is interacted in the specification with dummies that take the value 1 for the month that the assignment to treatment was announced, the months during the treatment and the months post-treatment, respectively, thus allowing comparison relative to the pre-announcement period. Each regression includes year x month fixed effects ψ_{ym} (which absorb the main effects of the time dummies) and worker fixed effects η_w (which absorb the main effect of the treatment indicator). Standard errors are clustered at the production line level.[25]

The results of this analysis are presented in Table B1 in the Appendix, and show no significant impacts on retention during or after the program. We present a variety of additional tests in Appendix C and conclude that there is no discernible effect on the size or composition of the retained worker sample over the observation period. We discuss these tests and the implications for the analysis of conditionally observed outcomes used later in the analysis below.

When examining conditionally observed outcomes such as productivity (which are only observed if the worker is still at the firm and working that day), there is a potential for selective observation based on treatment, which could generate bias in the impact estimates. To test and account for this potential bias, we follow several approaches:

1. *Testing directly for treatment-induced changes in the relative size of treatment v. control groups:* Note that estimating the regression specification in equation 1 is a direct test for differential retention on average across treatment and control groups. As discussed above, the results presented in Table B1 indicate there was no differential retention on average during or after training.

2. *Balance tests by baseline characteristics at different points after program start:* Even if retention rates were similar on average between treatment and control groups during and after the program, the composition of retained workers may differ between treatment and control groups and bias estimates of impacts on conditionally observed outcomes. To test whether the retention across treatment and control is correlated with baseline characteristics, we present the results of balance tests by treatment and control one month after treatment (July 2014) as well as during the last month of data collection (December 2015). Results are presented in Table C1 and demonstrate that all baseline characteristics are balanced on means at both points in time.[26]

[25]While we did a two level randomized treatment assignment with the lower level of treatment at the worker level, we report line level clustering to be conservative in our estimation of confidence intervals.

[26]Tests conducted for other points in time are also balanced. Furthermore, graphs of retention at treatment announcement,

3. *Dynamic weighting of conditionally observed outcomes:* Despite not finding any evidence of differential retention on average after program start nor compositional differences in retained workers across treatment and control groups, in order to confidently recover population average treatment effects on conditionally observed outcomes throughout the observation period, we weight treatment and control groups by the probability of being observed at any intermediate point in the data.[27] We adapt the approach proposed in Wooldridge (2010) to accommodate any potential heterogeneous impacts of treatment by baseline characteristics of the workers and any differential dynamics in the onset or decay of treatment effects across time, via the following two steps. First, we estimate a probit specification for the probability of being observed on the treatment indicator interacted with month by year fixed effects and baseline characteristics (attendance, education, tenure, age, skill grade, productivity and operation complexity).[28] Second, we estimate equation 1 using the conditionally observed outcome variables on the left-hand side and the inverse of the predicted probabilities of observation from the first step as probability weights.[29]

4.2 Attendance

Along with retention, the attendance roster allows for estimation of treatment impacts on additional outcomes of interest such as daily presence, unauthorized absence, cumulative person days, and tardiness. We estimate impacts of treatment on these outcomes in the same specification presented in equation 1 above, weighting observations by inverse probability of retention as these attendance variables are only measured if the worker is still an employee of the firm. The results are also presented in Table B1 in the Appendix. We find no evidence of impacts on these outcomes.

4.3 Operation Complexity and Productivity

Next, we investigate treatment impacts on two key outcomes of interest from the productivity data: standard allowable minutes (SAM) and efficiency. As discussed above, SAM measures operation complexity, and efficiency is the industry standard measure of productivity calculated as actual pieces produced divided by target pieces (the latter being derived from SAM). Both of these variables are only measured if a worker is retained by the factory, and present in the factory that day. Accordingly,

program completion, and data collection endline at different points across the distributions of baseline characteristics (which provide a more stringent test than balance checks based on means) show no evidence of heterogeneity. Both sets of results are omitted here for brevity, but available on request.

[27] For example, there may exist differential attrition across treatment and control, say, six months into program implementation, even if this difference later equalizes. To ensure that we recover the population average treatment effect on any conditionally observed outcome (e.g., productivity or salary) at all subsequent points of observation, we weight all observations prior to that time by the probability of being able to measure the outcome at each point in time.

[28] The outcome is a dummy variable that takes the value 1 if the worker is in the sample on any given month and 0 otherwise (i.e., the retained dummy if studying impacts from the attendance or salary data and the working dummy combining retention and attendance if studying impacts from the production data). Since workers salaries are homogenous within skill grade level, grade proxies for skill level as well as salary.

[29] In practice, once worker fixed effects are included in all regressions, the weighting procedure has negligible effect on the results. We explored robustness to different weights, as well as the absence of weights altogether, but do not present these results for the sake of brevity as they are qualitatively similar. That is, for the remainder of the analysis we report results from weighted regressions as the technically correct approach, but the results generally differ negligibly from those obtained from unweighted regressions.

Table 2: Impacts of P.A.C.E. Treatment on SAM and Efficiency

	(1)	(2)	(3)	(4)	(5)	(6)
	SAM (Operation Complexity) Standard Allowable Minute			Efficiency Produced/Target		
After X P.A.C.E. Treatment	0.0410*	0.0316	-0.00101	0.0480***	0.0976**	0.0736**
	(0.0219)	(0.0291)	(0.0150)	(0.0176)	(0.0412)	(0.0355)
During X P.A.C.E. Treatment	0.0430**	0.0288	-0.00393	0.0379**	0.0519	0.0320
	(0.0188)	(0.0259)	(0.0108)	(0.0165)	(0.0347)	(0.0252)
Fixed Effects	Worker, Garment	Worker X Garment	Worker X Garment X Operation	Worker, Garment	Worker X Garment	Worker X Garment X Operation
Observations	263,161	260,984	255,877	263,161	260,984	255,877
Control Mean of Dependent Variable	0.565	0.565	0.565	0.542	0.542	0.542

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. In all columns observations are weighted by the inverse of the predicted probability of working (i.e., not yet attrited and present in the factory with non-missing data) in the sample that day from a probit regression of the working dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. All specifications include month-year fixed effects to account for nonlinear time trends. All samples are trimmed in these regressions to omit days in which the worker is observed for less than a full production day of 8 hours and days in which the worker is observed for more than 1 overtime hour as these are anomalous observations with imprecise production measures.

these conditionally observed outcomes are weighted in the analysis as discussed above.[30]

Because treatment may impact the types of garments and tasks assigned to workers, we present several different regression specifications in Table 2 to document how controlling for the specific garment style on which the line is working and the specific operation to which the worker is assigned change the estimated treatment effects. In the base specification in columns 1 and 4, we add only garment style fixed effects to the specification in equation 1 to account for the size of the order and any variation in style complexity not sufficiently captured by the SAM of the individual operations and the resulting target quantity.[31] The results in column 1 show that treated workers are assigned to more complex operations both during and after treatment. Operations to which they are assigned are expected to take about 2.5 seconds (0.041 minutes) more, or a roughly 7% increase over the control group mean.

In order to estimate the effect on productivity net of the treatment impacts on operation complexity shown in column 1, we replace the worker and garment style fixed effects with worker *by* garment style fixed effects in columns 2 and 5. The results in column 5 show that indeed this change to the specification impacts the estimated treatment effects on productivity, making them substantially larger. However, the estimates in column 2, though less precise than those in column 1, suggest that the treatment effects on operation complexity might not be fully accounted for with worker by garment style fixed effects. Accordingly, in columns 3 and 6 we include worker by garment style by operation fixed effects as our most saturated specification. The estimates in column 3 indicate that this specification

[30]The weights are obtained as discussed in section 4.1 using the working status dummy as the outcome which takes value 1 if the worker is retained as an employee *and* present that day in the factory, and 0 otherwise.

[31]In alternative specifications, we include as additional controls days that the style has been running on the production line and total order size to further account for learning dynamics, but find that these have little effect on the results once garment style fixed effects are included.

now fully accounts for treatment effects on operation complexity.

The estimates of treatment effects on productivity from this preferred specification are reported in column 6 of Table 2. The magnitudes indicate that treated workers are more efficient after the program (relative to the month of treatment assignment announcement) by 7.4 percentage points, a more than 13.5% increase over the control group mean.[32] Impacts on productivity are stronger after program completion, with the during treatment coefficient less than half the size of the after treatment coefficient.[33]

4.3.1 Impacts on sub-sample of retained workers and line-level estimates

To further address any remaining concerns regarding bias due to selective attrition specifically relating to impacts on productivity, we present two additional sets of estimates in Table B2 in the Appendix. First, we estimate worker-level productivity impacts of training for the sub-sample of workers who were retained until the end of the data collection period (column 1). The magnitude of the treatment effect is similar, about 9.5 percentage points higher efficiency after the treatment, supporting the notion that productivity impacts are not driven by changing composition of the sample over time, though the effects are noisier due to lower sample size.[34] We also present in Table B2 results for productivity and operation complexity at the line level, including all workers on the production lines, not just individual workers who were included in the experimental sample. Line level results are also consistent with individual-level results. Note that we would expect smaller effects at the production line level, given that only a fraction of workers on each line were treated, but we still find a significant productivity increase of about 8% of the mean (about 5.7 percentage points in efficiency).

4.4 Salary and Career Advancement

In addition to worker presence and productivity, we study career advancement within the firm. To estimate the impacts of treatment on career advancement, we consider both whether the worker was given a raise or promotion using monthly payroll data as well as worker-reported measures of expectations of promotion; whether they recently requested (and received) skill development training; earned production incentives; and finally, how they assess their own ability relative to all workers on their production line, and relative to workers of the same technical skill grade as them. Unlike the salary

[32]This magnitude is consistent with recent studies estimating treatment effects of light touch, soft skills related interventions on workplace performance. Menzel (2017) finds that an intervention in which supervisors who have produced a garment style before communicated for up to 30 minutes with supervisors who are about to produce the same style increases productivity by five productivity points (about 12% of mean productivity) in the first two days of production, which indicates a substantially faster learning curve. Prada et al. (2019) find that communication and leadership training for one day per week over five weeks amongst sales associates and managers in a large retailer in Chile increased sales by about 12% up to three months later.

[33]The fact that workers are absent from the production line for one hour per week during the program for training raises the concern that productivity gains from the program may arise because workers may be happy with the reduction in working time, or necessarily more efficient in the remainder of the time in order to meet targets despite a shorter work day. However, the largest productivity gains appear after the completion of the training, when workers are not receiving these breaks anymore, and persist for 8 months after the program (when the collection or productivity data ended).

[34]In additional results, omitted here for brevity, we test differences between the productivity gains for the available sample at each point in time and the sub-sample of workers who are retained at the end of the observation period. We cannot reject that the coefficients are the same in any month.

and promotion outcomes which are observed at the monthly level for each worker, the self-reported measures are from the worker-level survey conducted in the month of program completion and vary only cross-sectionally.

Subjective expectations of promotion were measured by a binary variable for whether the worker expects to be promoted in the next six months. The request for skill development was measured by asking workers whether they have undergone technical skill development training in the last six months. Self-reported performance was measured by asking whether workers have received production awards or incentives in the last 6 months. Finally, we measured two kinds of self-assessment. Both asked the worker to imagine a ladder with six steps representing the worst to best workers on their production line (6 being the best). The first self-assessment asked workers where they would place themselves relative to all the workers on their line, and the second where they would place themselves relative to other workers of their technical skill grade.

For salary and promotion outcomes, we first estimate the retention probability weights as detailed in section 4.1, and then estimate equation 1 using those inverse probability weights, with the log of gross salary as the outcome.[35] Since the variation in the survey variables is only cross-sectional, we regress these outcomes on a binary variable for treatment or control, and include factory fixed effects, as well as controls for age, tenure with the firm, and education of the worker. In survey outcome regressions, we employ weights obtained from the retention probit using attendance data matched to the date of survey.

Column 1 of Table 3 presents the results of the estimation comparing treatment workers to control workers during the treatment assignment announcement month, and during and after the treatment (relative to before the treatment assignment announcement month). Treatment workers receive an insignificant half a percent more wages in the period after the program completion, which translates to roughly 45 INR (.75 USD) a month. Thus, despite being assigned to more complex operations and being substantially more productive, treated workers are not paid meaningfully higher wages over the 18 months following program completion.

Column 2 of Table 3 presents impacts on promotion, measured in the same monthly salary data. We find that treated workers are 4.7 percentage points more likely to be promoted over the year and a half following program completion than are control group workers (13.4% of whom were promoted by the end of the observation period). Taken together, the results in columns 1 and 2 indicate that the firm rewards treated workers for their greater productivity with a title promotion, but that this promotion does not come along with a substantial increase in pay. This pattern is consistent with the wage compression in this setting. Pay schedules are closely pinned to minimum wages which vary along more coarse levels of skill than do grade and designation within the firm. Accordingly, it is quite possible in this context to be promoted to a higher skill grade or designation but be subject to the same minimum wage as before the promotion, resulting in little or no increase in pay.[36]

Columns 3-7 of Table 3 present the results from analyses of related survey outcomes. Treatment

[35] Note that the administrative salary data is at the monthly level for each worker rather than the daily-level.

[36] Indeed in the month a worker was promoted, the median increase in salary is less than 2%; the mean is about 3%; and for 25% of the sample, pay does not increase at all despite a title promotion. Nevertheless, the title promotion is still valued by workers as higher designation workers get more respect from their peers; pressure from supervisors is likely less intense; and the operations to which higher skill workers are assigned tend to be less tedious.

Table 3: Impacts of P.A.C.E. Treatment on Salary and Career Advancement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Log(Gross Salary)	Promoted	Expect Promotion Next 6 Mos	Skill Development Training	Production Award or Incentive	Skill Peer Self-Assessment	Co-Worker Self-Assessment
	Salary Data		***Survey Data***				
After X P.A.C.E. Treatment	0.00554	0.0468**					
	(0.00355)	(0.0202)					
During X P.A.C.E. Treatment	0.00147	0.0107					
	(0.00109)	(0.00975)					
Announced X P.A.C.E. Treatment	0.000370	-0.00164					
	(0.000873)	(0.00331)					
P.A.C.E. Treatment			0.0852**	0.157***	0.0298	0.118*	0.0628
			(0.0415)	(0.0468)	(0.0184)	(0.0651)	(0.0666)
Observations	34,480	32,596	621	621	621	621	621
Control Mean of Dependent Variable	9.021	0.134	0.563	0.249	0.032	5.337	5.298

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. Observations are weighted in regressions by the inverse of the predicted probability of being retained (i.e., not yet attrited with non-missing data) in the sample that day from a probit regression of the retained dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. Specifications in columns 1 and 2 include month-year fixed effects to account for nonlinear time trends, as well as worker fixed effects. Controls in regressions for survey outcomes in columns 3 through 6 include demographic baseline variables from Table 1 (i.e., dummies for education levels, dummies for deciles of the age distribution, and dummies for tenure in integer years) as well as unit fixed effects.

workers are about 8.5 percentage points more likely to report that they expect a promotion within the next six months (roughly 15% of the control group mean), and are nearly 16 percentage points more likely to request skill development training (63% of the control group mean). They are not significantly more likely to report having received a production incentive or award, but rate themselves higher relative to peer co-workers. Specifically, when asked to rank themselves relative to workers the same technical skill grade, they are significantly more likely to rate themselves at a higher level (as shown in column 6).

5 Mechanisms

Though the program was broad and multifaceted in its content, the core modules focused on communication skills, problem solving, time-use planning and task prioritization. These modules involved both lectures and team-based exercises to practice applying these skills to work scenarios. In this sense, the program as a whole emphasized both the skills underlying effective teamwork and group-based application of these skills to team exercises. As mentioned in section 2.2, the final month of the program included a module (Execution Excellence) in which this emphasis on the importance of teamwork and how the skills learned up till that point contribute to effective collaboration at work was made explicit. The fact that the largest productivity impacts appear at the end of the program is consistent with this emphasis on teamwork contributing strongly to the gains in productivity.

In this section we explore heterogeneity in productivity impacts as well as spillover impacts on the productivity of untrained co-workers working alongside treated workers. We discuss the degree to which these additional results support the interpretation that productivity impacts are largely driven by teamwork and collaboration skills. We also investigate the degree to which improvements in these skills substitute for managerial attention on coordination and productivity balancing responsibilities and/or are complemented by autonomy on the part of managers in making adjustments to production processes. We also review several alternative interpretations (e.g., reciprocity, sheepskin effects, general worker satisfaction, and social capital) and discuss the plausibility of each in light of the additional analysis presented here and in the Appendix.

5.1 Heterogeneous Productivity Impacts: Skill Level, Teamwork, and Leadership

We first investigate the degree to which productivity gains from training are heterogenous by the skill grade of the worker, various measures of the degree of teamwork involved in the production process, and the workers' baseline stock of leadership skills. We do so by estimating a modified version of equation 1 in which we add terms for interactions between the dimension of heterogeneity and the treatment by time period terms, as well as all double interactions and the main effect of the dimension of heterogeneity when not absorbed by the fixed effects in the specification.[37]

For the skill grade of the worker, we create a dummy variable taking value 1 if the worker was at least a B grade operator in May 2013 (before training), and 0 otherwise. The results in column 1 of Table

[37] The specifications in columns 1 through 3 of Table 4 correspond to that in column 4 of Table 2 as the dimensions of heterogeneity studied do not vary sufficiently within worker-operation assignments in the data; while the specifications in columns 4 and 5 of Table 4 correspond to that in column 6 of Table 2.

Table 4: Heterogeneous Impacts of P.A.C.E. Treatment on Productivity

	(1)	(2)	(3)	(4)	(5)
			Efficiency Produced/Target		
After X P.A.C.E. Treatment	0.00325	-0.00460	0.0438**	0.0697*	0.145**
	(0.0193)	(0.0274)	(0.0183)	(0.0356)	(0.0572)
During X P.A.C.E. Treatment	0.00191	-0.0135	0.0349**	0.0297	0.0554
	(0.0168)	(0.0253)	(0.0170)	(0.0253)	(0.0535)
Skill Grade B or Above X After X P.A.C.E. Treatment	0.0603***				
	(0.0214)				
Skill Grade B or Above X During X P.A.C.E. Treatment	0.0490***				
	(0.0166)				
Above Median Workers per Operation X After X P.A.C.E. Treatment		0.0731**			
		(0.0291)			
Above Median Workers per Operation X During X P.A.C.E. Treatment		0.0703**			
		(0.0273)			
Size of Operation Team X After X P.A.C.E. Treatment			0.0372**		
			(0.0180)		
Size of Operation Team X During X P.A.C.E. Treatment			0.0264		
			(0.0196)		
Other PACE Trainee in Operation Team X After X P.A.C.E. Treatment				0.101***	
				(0.0373)	
Other PACE Trainee in Operation Team X During X P.A.C.E. Treatment				0.0385	
				(0.0339)	
Baseline Leader Grade X After X P.A.C.E. Treatment					-0.0488**
					(0.0244)
Baseline Leader Grade X During X P.A.C.E. Treatment					-0.0191
					(0.0212)
Observations	263,161	263,161	263,161	255,877	225,862
Control Mean of Dependent Variable	0.543	0.542	0.542	0.542	0.541

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. In all columns observations are weighted by the inverse of the predicted probability of working (i.e., not yet attrited and present in the factory with non-missing data) in the sample that day from a probit regression of the working dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. All specifications include month-year fixed effects to account for nonlinear time trends. All samples are trimmed in these regressions to omit days in which the worker is observed for less than a full production day of 8 hours and days in which the worker is observed for more than 1 overtime hour as these are anomalous observations with imprecise production measures. Specifications in columns 1 through 3 incldue worker and garment fixed effects like that in column 1 of Table 2; while the specifications in columns 4 and 5 include worker by garment by operation fixed effects and are identical to that in column 4 of Table 2. "Skill Grade B or Above" is a dummy indicating if the worker has a skill grade of at least B or higher. "Above Median Workers per Op" is a dummy indicating if the ratio of workers per operation for that line day is above the median value in the data of 1.04. The number of workers per operation ranges from 1 (for 43.5% of observations) to 1.5 (99th percentile) mostly. "Size of Operation Team" is the number of other workers working on the same operation and ranges mostly from 0 to 2 (99th percentile). "Other PACE Trainee in Operation Team" is a dummy taking value 1 if any of the other workers working on the same operation along with the worker is also PACE treated, and 0 if none of the other workers working on the same operation are treated or if there are no other workers working on the same operation. "Baseline Leader Grade" was assessed by supervisors and HR representatives prior to the randomization for nearly 90% of the workers enrolled in the lottery and ranges 0 to 3.

4 indicate that the impacts on productivity are concentrated entirely among workers with at least a skill grade of B. Higher skill operators can perform more technical operations, and are, therefore, more likely to be assigned to those critical operations for a style, which are most likely to cause bottlenecks in the line-level output of completed garments. As discussed in section 2.1.1 these critical operations are more likely to have multiple operators and/or helpers assigned to them. In this sense, the results in column 1 may reflect the degree to which higher skill operators are more likely to be working in operation teams within the line.

To investigate this further, we calculate at the line-day level the ratio of the number of workers to the number of operations. We then create a dummy variable for whether the ratio is above the median value of 1.04.[38] Once again we find in column 2 of Table 4 that the productivity impacts are concentrated entirely among line-day observations for which the ratio of workers to operations indicates production involves teams of workers working together on the same operation.

We then calculate for each worker-day the number of additional workers working alongside on the same operation. This value ranges from 0 (indicating that the operation was being performed alone rather than by a team) to a 99th percentile of 2 additional workers (corresponding to a team of 3 total workers). The results in column 3 of Table 4 show that the magnitude of the treatment effect increases by roughly 85% from 4.38 to 8.1 percentage points when moving from working alone to working on an operation team with 1 co-worker, and grows to 11.8 percentage points when the treated worker is working with 2 other co-workers on the same operation.

In column 4 of Table 4, we investigate whether productivity impacts are larger when working alongside another treated co-worker. That is, we construct a dummy variable taking value 1 if the worker is working alongside at least 1 other trained worker, and 0 otherwise. Note that the worker-day observations for which this variable is 0 include both instances in which the treated worker is working alone and instances in which she is working on a team but none of the other co-workers are treated. The results in column 4 show that the treatment effect on productivity is nearly 2.5 times the size when the treated worker is working alongside another treated worker on the same operation.

Finally for baseline stock of leadership skills, we asked factory HR representatives to rank participants in the training lottery (before treatment assignment) into 4 levels (ranging from 0 to 3) of baseline leadership skill (defined broadly as confidence and ability to effectively communicate with and motivate co-workers). The results, presented in column 5 of Table 4, show that productivity gains from training were strongest (nearly twice the size of the average effects reported in column 6 of Table 2) among those assessed by factory HR representatives as having the lowest baseline stock of leadership skills. The magnitude of the interaction coefficient indicates that workers assessed to already have the highest possible stock of leadership skills at baseline (with a grade of 3) exhibited no productivity gains from training at all.

This result suggests that the training in soft skills was indeed a substitute for baseline stocks rather than a complement. That is, it was not clear prior to the experiment whether the training would be most impactful for workers with deficiencies in those skills at baseline or rather would require some foundational stock of skills upon which to build. Structural estimates of dynamic human capital

[38] 43.5% of line-day observations have a ratio of 1 and the 99th percentile is 1.5.

accumulation models suggest dynamic complementarities in the productive value of non-cognitive skills, such that older children with low stocks of these skills benefit less than those who accumulated greater foundational stocks of these skill at earlier ages (Aizer and Cunha, 2012; Cunha et al., 2010; Heckman and Mosso, 2014). However, whether this translates into similar patterns among adults with varying stocks of baseline skills is unclear. Our results suggest that targeting those who lack soft skills necessary for their current job has high returns.

5.2 Productivity Spillovers and Interaction with Managerial Quality

Taken together, the results in Table 4 strongly support the interpretation that the treatment effects on productivity were driven largely by improvements in teamwork. If indeed improvements in teamwork are driving the productivity gains, we should expect that the productivities of untrained co-workers working alongside trained workers also reflect the improved teamwork. Recall from the description in section 2.3 that the experiment was specifically designed to capture spillovers within production lines through a two-stage randomization procedure, in which lines were first randomized to treatment or control, and then within treatment lines, workers who had enrolled in the P.A.C.E. lottery were randomized to treatment or to the spillover group. To estimate the effects on untrained workers who interact with trained workers, we re-run the specification reflected in column 6 of Table 2, replacing the binary treatment variable with the binary spillover treatment variable. This variable compares untrained workers in treatment lines (workers who enrolled in the lottery but did not receive the program and who work in production lines with workers who were treated with the program) with control workers in control lines (workers who enrolled in the lottery but did not receive the program and who work in production lines without any treated workers).[39]

The results, presented column 1 of Table 5, show that untrained co-workers on treatment lines do indeed become significantly more productive after the program. A comparison of the magnitudes between column 1 of Table 5 and column 6 of Table 2 indicates that more than 80% of the direct effect of treatment on productivity spills over onto the productivities of untrained workers on the same lines. This result is again consistent with productivity gains largely reflecting improvements in teamwork (from which untrained co-workers would benefit) rather than other elements of the curriculum or other mechanisms (discussed below) which would only be reflected in the productivity of directly trained workers. Additional spillover results, presented in Table B3 in the Appendix, show little evidence of substantial spillover effects on other attendance and advancement outcomes.[40]

Given the evidence in support of improved teamwork among co-workers driving the productivity impacts, we next investigate the degree to which this improved collaboration interacts with the effort and leadership style of production line supervisors. We estimate heterogeneity in both treatment and spillover effects in a similar approach to that employed in Table 4 (with specifications analogous to that in column 6 of Table 2). In particular as described in section 3.3, we leverage data from a survey among

[39]Probability weights are calculated exactly as they are in the treatment effect estimation, using spillover treatment indicators in place of direct P.A.C.E. training.

[40]Estimates are only significant for 2 of the 10 additional outcomes studied. Impacts on salary are very precisely estimated but small in magnitude (1.25%). We also note that there was a small significant difference in salary (-1.8%) between the spillover and control samples at baseline which might be reflected in this result.

Table 5: Productivity Spillovers and Interaction with Managerial Quality

	(1)	(2)	(3)	(4)	(5)
			Efficiency Produced/Target		
After X Spillover	0.0642*	0.104***	0.0848***		
	(0.0372)	(0.0369)	(0.0319)		
During X Spillover	0.00173	0.0100	0.00540		
	(0.0291)	(0.0364)	(0.0242)		
Managerial Attention (Std) X After X Spillover		-0.0834**			
		(0.0413)			
Managerial Attention (Std) X During X Spillover		-0.0235			
		(0.0359)			
Managerial Autonomy (Std) X After X Spillover			0.0627*		
			(0.0363)		
Managerial Autonomy (Std) X During X Spillover			0.0103		
			(0.0221)		
After X P.A.C.E. Treatment				0.106***	0.108***
				(0.0353)	(0.0330)
During X P.A.C.E. Treatment				0.0332	0.0422*
				(0.0338)	(0.0245)
Managerial Attention (Std) X After X P.A.C.E. Treatment				-0.0979**	
				(0.0432)	
Managerial Attention (Std) X During X P.A.C.E. Treatment				-0.0473	
				(0.0376)	
Managerial Autonomy (Std) X After X P.A.C.E. Treatment					0.0784**
					(0.0368)
Managerial Autonomy (Std) X During X P.A.C.E. Treatment					0.0314
					(0.0227)
Observations	213,499	150,410	150,410	168,585	168,585
Control Mean of Dependent Variable	0.548	0.577	0.577	0.569	0.569

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. In all columns observations are weighted by the inverse of the predicted probability of working (i.e., not yet attrited and present in the factory with non-missing data) in the sample that day from a probit regression of the working dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. All specifications include month-year fixed effects to account for nonlinear time trends. All samples are trimmed in these regressions to omit days in which the worker is observed for less than a full production day of 8 hours and days in which the worker is observed for more than 1 overtime hour as these are anomalous observations with imprecise production measures. All specifications include worker by garment by operation fixed effects and are identical to that in column 6 of Table 2. "Managerial Attention" is a factor obtained from survey measures of supervisors' reported monitoring frequency, various personnel management activities, communication with workers and upper level management, as well as efforts to meet targets and motivate workers. "Managerial Autonomy" is a factor obtained from survey measures measuring leadership style (consideration and initiation structure) and willingness to solve production problems and respond to worker concerns and issues without relying on upper management. The measurement system from which these factors are obtained is developed in Adhvaryu, Nyshadham, and Tamayo (2019d). Both factors are standardized to be mean 0 and SD 1 in the sample so as to interpret the main effects of the treatment by time regressors as the effects on lines with managers of average quality along these dimensions and the interactions as the incremental effects on lines with supervised by managers with 1 SD greater Attention or Autonomy.

production line supervisors meant to comprehensively capture different dimensions of managerial quality. We focus on two dimensions: Attention and Autonomy.

The Attention factor captures the effort managers devote toward their responsibilities of coordinating between and balancing productivity across workers and segments of the lines so as to avoid bottlenecks in line-level output of completed garments. To the degree that substantial managerial effort of this type obviates, at least in part, the need for significant communication and coordination among co-workers, we might expect that productivity gains from improving workers' teamwork skills are smaller on lines supervised by more attentive managers. The results in columns 2 and 4 of Table 5 show that, indeed, spillover and treatment effects are large on lines managed by supervisors exhibiting average levels of Attention, but are 80-90% smaller on lines managed by supervisors exhibiting Attention 1 SD above the mean.

The Autonomy factor captures the degree to which the manager is willing and able to make changes to production processes without having to consult upper level management. If the treatment effects on productivity work to some degree through a greater confidence and ability among workers to effectively communicate needs and issues (e.g., machine calibration issues, need for a helper to complete operations, etc.) to their line supervisor, then we might expect that these improvements would be complemented by the degree to which the supervisor is willing and able to meet those needs or resolve those issues in a timely manner. The results in columns 3 and 5 of Table 5 show that, indeed, spillover and treatment effects are again large on lines managed by supervisors exhibiting average levels of Autonomy, but are about 75% larger on lines managed by supervisors exhibiting Autonomy 1 SD above the mean.

5.3 Survey Outcomes

In the Appendix (Table B4), we present additional results reflecting treatment effects on a broad set of survey outcomes meant to provide some evidence of first stage impacts on the acquisition of knowledge and resulting stock of non-cognitive skills. As mentioned previously, the survey measures are cross-sectional. The regression specification is thus the same as for the survey outcomes in Table 3.[41] The results are presented in Table B4 in the Appendix.

5.3.1 Financial Behaviors and Attitudes

The results in Panel A of Table B4 show a large positive impact (about 30% increase from the control group mean) on saving for own and children's education, consistent with personal goal setting, prioritizing actions in service of these goals, and mapping workplace motivation to the pursuit of these goals at the center of the Time and Stress Management core module.

[41] We regress the outcome on the binary treatment variable and include factory fixed effects, as well as controls for age, tenure with the firm, and education of the worker. and use retention weights from the attendance data matched by survey date.

5.3.2 Government and Firm Entitlements

The results in Panel B show that treated workers are substantially more likely to avail themselves of government pension and government subsidized healthcare programs (both covered in the Legal Literacy and Social Entitlements module of the program). These impacts could reflect both improved knowledge on the topics and increased effectiveness of information acquisition.

5.3.3 Personality Traits

Impacts on key personality traits (shown in Panel C) are imprecisely estimated, likely reflecting the challenge of translating these concepts into several local languages and the degree to which these concepts are novel or foreign to both field surveyors and factory workers. Nevertheless, P.A.C.E. treatment does have a large positive and statistically significant impact on extraversion.[42] This result is consistent with both the theme of assertive and effective communication emphasized in the Communication core module as well as the practice of role-playing and participation in teamwork activities emphasized throughout the training, particularly in the Execution Excellence module in the final month of the program.

5.3.4 Mediation Analysis

In the Appendix, we follow Heckman et al. (2013) and Huber (2014) in conducting a mediation analysis to calculate the contribution of the estimated changes in personality characteristics presented in Panel C of Table B4 to the productivity impacts estimated in Table 2.[43] The results support a role for extraversion in mediating the productivity gains (explaining roughly 30% of the overall treatment impact on productivity). We interpret this pattern as consistent with improved communication and leadership skills driving the productivity gains, but also note that the mediation analysis shows that a majority of the productivity gains are unexplained by these survey measures. Accordingly, we rely more heavily on the above heterogeneity analysis to inform the interpretation of mechanisms of impact.

5.3.5 Mental Health and Aspirations

Finally, Panel D shows that, in general, outcomes associated with psychological well-being (self-esteem, optimism, and mental distress) are unaffected by P.A.C.E. treatment, but aspirations for children's education rise substantially in relation to the control group mean, consistent with the result on saving for

[42]In order to measure extraversion, workers were asked about the extent (measured on a 5-point scale of agreement ranging from Strongly Agree to Strongly Disagree) to which 10 statements described themselves, five of which were positively related to extraversion (e.g., "I take charge and "I make friends easily) and five of which were negatively related (e.g., "I have difficulty expressing my feelings and "I hold back my opinions). The score from each variable was added up for the negative statements and the score from the negative statements was then subtracted from the score for positive statements. We present results in Table C2 of the Appendix which correct for multiple hypothesis testing. The corrected q-value (following (Anderson, 2008)) for this impact on extraversion is .108, just shy of statistical significance at conventional levels but still indicative of a treatment impact.

[43]We do so by combining treatment effects on mediators and productivity with estimated heterogeneity in productivity impacts by these mediators. We employ inverse probability weighting to account for endogeneity in mediators, as in Huber (2014).

education in Panel A and, as mentioned above, the core concept of personal goal setting as a source of workplace motivation emphasized in one of the core modules.

5.3.6 Additional Propensity Score Matched Evidence on Soft Skills

Panels C and D present an admittedly small set of personality measures and non-cognitive skills. To supplement this analysis, we fielded a subsequent survey of additional soft skills and personality measures on a separate propensity score-matched sample of trained and untrained workers.[44] The results are presented in Figure B2 in section B.6 of the Appendix and are consistent with P.A.C.E. training being associated with increases in a broad array of non-cognitive skills; however, the lack of randomized training assignment in this sample makes these results at most suggestive. Taken together we interpret impacts on survey outcomes as consistent with first stage impacts on the acquisition of knowledge and non-cognitive skills as a result of the program.[45]

5.4 Alternative Mechanisms

We presented evidence in Tables 4 and 5 that the productivity impacts are most pronounced when treated workers work on common operations along with other co-workers (particularly treated co-workers) and that untreated co-workers on the same lines as treated workers also exhibit large gains in productivity (with magnitudes more than 80% the size of the direct treatment effect). We also show that productivity gains are concentrated among workers with low baseline stocks of leadership skills and on lines supervised by managers who are less attentive to their responsibilities of coordinating and balancing production between workers and segments of the line. Finally, we document evidence consistent with a first stage impact of P.A.C.E. training on stocks of non-cognitive skills. Taken together, we interpret this pattern as indicative of improvements in teamwork and collaboration skills (e.g., communication, leadership, and planning) from the training being a primary driver of the gains in productivity. However, several alternative interpretations are possible and we now discuss each in light of the full set of results discussed above.

[44]We were unable to collect a larger set of measures during the experiment as factory management imposed a ceiling on the duration of survey enumeration. Unfortunately, the original sample of workers had mostly left the factory by the time we fielded this second survey.

[45]Additional suggestive evidence on gains in stocks of soft skills comes from pre- and post-module assessments built into the program. These assessments were only administered to program participants, and thus we cannot compute a treatment vs. control difference, rather only a post vs. pre-module difference for treated workers. The patterns suggest that P.A.C.E. participants had low baseline stocks of soft skills and improved their stocks of these skills through the training. The changes are all in the neighborhood of 85-110 percent, with the largest changes (in percent terms) for Communication, Problem Solving/Decision-Making, Legal Literacy, and Execution Excellence. The largest raw difference is in the Time and Stress Management module. These results support the notion that workers absorbed the skills taught in each of the core modules, such that the stock of skills increased, at least when measured in the short-term; however, we should note some caveats in interpreting these changes. First, as described above, control workers were not given the assessments, so we are not able to estimate impacts by comparing treatment v. control. Second, we are measuring skill retention directly after module completion; this does not necessarily reflect long-term skill retention.

5.4.1 Reciprocity

First, we address the potential importance of reciprocity / gift exchange (an increase in effort provision in response to the employer "gifting" the worker access to the program). While it is indeed plausible that reciprocity explains some part of the observed impacts of P.A.C.E., we believe it is unlikely that the majority of impacts are due to this mechanism, for two reasons. First, the heterogeneity results show that productivity gains are largest when treated workers work on common operations along with other co-workers. Column 2 of Table 4 indicates that the treatment effects on productivity are entirely driven by instances in which the line has more than 1 worker per operation and column 3 indicates that the magnitude of the productivity effect is significantly increasing in the number of co-workers with which a treated worker is working alongside on the same operation. It is difficult to think of a reason for why reciprocity would only manifest when the worker is working on a team but not when she is working alone on an operation. Second, one would expect that reciprocity would be strongest during the training program and weaken or dissipate altogether in the 8 months after program completion over which we observe the participants. However, the pattern of impacts is precisely the opposite (i.e., small impacts during the program, and larger persistent impacts post-program completion). For both of these reasons we do not believe the reciprocity motive is the primary driver of the productivity gains. This indirect evidence is in line with recent, more direct tests of the role of reciprocity in workplace settings as well (DellaVigna et al., 2016).

5.4.2 Certification Effects

Next, we evaluate the possibility that the results on productivity and operation complexity were due to sheepskin effects; i.e., taking part in P.A.C.E. "certified" workers as high quality from the perspective of management, and this led to the improvements in workplace outcomes we observe. First, we note that productivity gains among untrained workers working alongside treated workers (more than 80% the magnitude of direct effects) are inconsistent with a sheepskin effect mechanism, as this spillover sample would not be "certified" the way treatment workers could be. Additionally, the pattern of heterogeneity once again is inconsistent with a sheepskin mechanism as it is difficult to think of a reason for why sheepskin effects would appear when the worker is working on a team but not when she is working alone on an operation. Finally, managers were aware that training assignment was done via a lottery (i.e., selection into the program based on "high quality" unobservables was explicitly ruled out).

5.4.3 Increases in Well-Being

Third, it is possible that workers found the classes enjoyable and they improved workers' subjective wellbeing, which in turn made workers more productive. Once again, the fact that productivity impacts are smaller during the program (when such enjoyment would be most salient), and large and persistent after the program (when any derived enjoyment has presumably ceased), is inconsistent with this interpretation. In addition, the results reported in Panel D in Table B4 show that levels of moderate psychological distress, which might reflect this subjective wellbeing to some degree, are not statistically different by treatment status.[46]

5.4.4 Increased Social Connectivity

Finally, we consider the idea that increased social connectivity among co-workers on a line drives the results on productivity. The argument is that it is possible that the time spent together in the training sessions improved the social ties among co-workers, which could generate higher productivity on their production lines given the reliance on teamwork and collaboration. This mechanism is closest to our preferred interpretation, but differs in that no specific skills (e.g., communication, leadership, or planning) need be learned to achieve the gains in productivity if time spent together is enough to improve the quality of teamwork. Due to production constraints which dictated that the number of workers from the same production line who could leave at the same time for a P.A.C.E. session be minimized, co-workers on the same line were placed in different sessions conducted on different days of the week. We believe this likely limited the increase in within-line social connectivity as a result of shared time in the classroom.[47] Furthermore, spillover workers did not experience any changes to connectivity but still exhibited increased productivity, lowering the likelihood that time spent together during training is driving the productivity effects.

6 Return on Investment

To quantify the profit return to the firm, we combine our treatment effect estimates on productivity and wages for both treated and spillover workers with costing data obtained from the program administrators. We report in Table 6 calculations of the net present value of costs and benefits. Benefits are calculated in terms of incremental productivity from both treated and spillover workers using estimates from the randomized evaluation.[48] Costs involve programmatic overhead costs, lost productivity due to training time, and increased wages.

Table 6 first outlines costs of the program, both overhead costs and variable costs. The overhead costs are given by the costs of hiring two full-time trainers per factory for the 12 months of the program, additional support time from HR personnel, printed materials, food, and equipment (e.g., PA system). The variable cost is from lost production hours and increased wages. For the 1087 treated workers, total program costs are approximately $102,000, about $56,000 of which are overhead costs, and the remainder variable costs.

Details on revenue per garment from additional productivity and the corresponding profit margins were obtained from the firm. The benefits of the program are generated by higher productivity of both treated and spillover workers. At the end of the program period, the NPV of these benefits is just over

[46] Results are unchanged if severe mental distress is used as an outcome instead of moderate mental distress.

[47] Note that this feature does not preclude social connectivity from being impacted by the program; it simply lowers the likelihood that the time co-workers spent together in the classroom contributed largely to any improvements in connectivity that are reflected in productivity. Indeed the observed treatment effects on extraversion likely induced better communication and thus greater social ties – something that would be part of our teamwork and collaboration skills interpretation of the program's impacts.

[48] We omit gains from any additional person days (i.e., attendance and retention) as the estimated impacts are not statistically significant.

Table 6: Return on Investment Calculations (Costs and Benefits to Firm)

Sewing Department Only (1087 Treated Workers)	
P.A.C.E. Training Overhead Cost (Trainers, HR Oversight, Materials, and Food for 12 Mos)	-$55,732.59
P.A.C.E. Training Variable Cost (Lost Garments from Lost Production Hours)	-$36,200.12
Increased Wages (P.A.C.E. Trainees + Spillover Workers)	-$10,346.79
Total Cost *(All numbers in present value)*	-$102,279.50
1 Year After Program Announcement	
Additional Productivity (P.A.C.E. Trainees)	$148,184.50
Additional Productivity (Spillover Workers)	$5,998.64
Net Present Value of Subtotal	$154,183.10
Net Rate of Return	**61%**
20 Mos After Program Announcement	
Additional Productivity (P.A.C.E. Trainees)	$278,311.40
Additional Productivity (Spillover Workers)	$85,664.76
Net Present Value of Subtotal	$363,976.10
Net Rate of Return	**256%**
Assumptions	
Garments per Production Hour	1.0
Revenue per Additional Garment	$7.00
Labor Contribution to Cost ("Cut to Make")	25%
Profit Margin on Revenue from Additional Productivity	20.00%
Profit Margin on Lost Garments from Lost Hours	5%
Interest Rate	8%
INR per 1 USD	60

Notes: Trainer salaries were 17,000 INR per month for each trainer. There were 2 trainers for each of the 5 factories; 10 trainers in total. Additional HR personnel time for program oversight amounted to 6,659 INR per month across all 5 factories. Materials and equipment costs amounted to 26689 INR per month across all 5 factories, and food costs amounted to 27,175 INR per month across all 5 factories. Garments per hour is calculated by dividing the average worker level SAM (minutes to complete the operation on a single garment) by the line level SAM (minutes to complete a full garment for the line) and multiplying by 60 minutes in a work hour. All additional productivity and wage coefficients are taken from the estimates in the main results, assumted to be constant over the months within the during and after periods, and appropriately scaled by the number original sample workers remaining in the factory in each month. Revenue per additional garment is taken from the accounting department of the firm, as is the "Cut to Make" or labor percent contribution to total production cost. Profit margin on additional revenue generated through improved efficiency is calculated as 80% of the "Cut to Make" cost as instructed by the accounting office of the firm and the profit margin on additional revenue from an additional man day is equivalent to the average profit margin of the firm. The monthly interest rate is the average interest rate that prevailed during the study time period. Similarly, the exchange rate is the average from the study period.

$154,000. At the end of the productivity tracking period (8 months after program completion), total benefits are substantially higher, more than $360,000.[49]

The net rate of return at the end of the program period is thus 61% (i.e., at program end, costs had been entirely recouped by the firm, plus 61 percent additional returns). Eight months after program completion, flow benefits from post-program productivity impacts help generate a net rate of return of 256%.

7 Conclusion

We study the impacts of soft skills training on labor market and workplace outcomes. We combine randomized placement into an on-the-job soft skills training program for female garment workers in India with detailed measurement of productivity, retention, wages, and other survey outcomes, to characterize the effects of this training on workers as well as on the firm. We find that soft skills improvements generate large and persistent productivity impacts, seemingly by way of improved teamwork and collaboration, but have negligible effects on wages and turnover. These results are consistent with theories of labor market imperfections, and suggest that the firm captures most of the gains from the increased marginal productivity of labor.

Given the on-the-job setting and two-stage randomization among workers within the firm, we are able to directly estimate impacts on individual productivity as well as spillover impacts on untrained co-workers. This individual-level analysis allows us to document for whom and under what conditions improved soft skills generate the largest impacts on productivity. We show that productivity gains are most pronounced when trainees work on operation teams alongside other co-workers (particularly treated co-workers) and are mirrored among non-treated co-workers on the line, consistent with gains being driven by improved teamwork and collaboration. In this sense, our results highlight the productive value in the workplace of teamwork and collaboration skills, which are often underemphasized in traditional schooling systems like that in India (Kurtz, 1990; Kurtz et al., 1988). We also find that training is most impactful for workers identified by factory HR representatives to be lacking in leadership skills (defined broadly as confidence and ability to effectively communicate with and motivate co-workers) at baseline, indicating that training in soft skills is a substitute for baseline stocks of skill. Relatedly, we find that these improved teamwork and collaboration skills among workers appear to substitute for managerial attention (e.g., more intensive monitoring of and intervention to resolve production bottlenecks and imbalances), but are complemented by managerial autonomy (i.e., the degree to which production line supervisors make process adjustments in response to production issues raised by workers without consulting or acquiescing to upper level management).

Growing interest in active labor market policies (Card et al., 2017; Heckman et al., 1999; McKenzie, 2017), including in low-income countries (McKenzie, 2017) has spurred study of the impacts of vocational training programs, which often include a soft skills training component (Betcherman et al., 2004).

[49]Note that though the spillover treatment effects on productivity after the program are nearly 80% the magnitude of the direct effects, during the program spillover effects are negligible while direct effects are still economically meaningful. Accordingly, in combination with the smaller number of spillover workers as compared to the number of directly trained workers in the experimental sample, spillover productivity contributes little to program benefits 1 year after program announcement, but accounts for nearly a quarter of total benefits by the end of the observation period.

In general, estimates of the labor market benefits of training alone (as opposed to training plus asset or cash transfers) do not yield consistent evidence of impact (McKenzie, 2017). Interventions focused on young women may be one area of exception – see, e.g., recent work by Buvinić and Furst-Nichols (2016) and Acevedo et al. (2017). This recent work, along with our findings, indicate that greater concentration on active labor market interventions focused on women workers may yield high returns.

Finally, our work is relevant to the literature on female labor force participation (LFP) and employment outcomes, particularly in low-income country contexts (Heath and Jayachandran, 2016). This policy question of how to increase the LFP and career growth of women is especially salient in India, where the level of female LFP is not only unusually low considering India's level of development (India ranks 120th out of 131 countries in female LFP (Chatterjee et al., 2015)), but has substantially decreased in rural areas between 1987 and 2009, despite a fertility transition and relatively robust economic growth (Afridi et al., 2016). Studying improvements in career prospects for women, via managerial training and promotion as Macchiavello et al. (2015) do, or via soft-skills training and resulting productivity enhancements and promotions as we do, can contribute to our understanding of determinants of female labor force participation that are amenable to policy intervention.

References

Acemoglu, D. (1997). Training and innovation in an imperfect labour market. *The Review of Economic Studies*, 64(3):445–464.

Acemoglu, D. and Pischke, J.-S. (1998). Why do firms train? Theory and evidence. *The Quarterly Journal of Economics*, 113(1):79–119.

Acemoglu, D. and Pischke, J.-S. (1999). Beyond Becker: Training in imperfect labour markets. *The Economic Journal*, 109(453):112–142.

Acevedo, P., Cruces, G., Gertler, P., and Martinez, S. (2017). Living up to expectations: How job training made women better off and men worse off. Technical report, National Bureau of Economic Research.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2019a). Management and shocks to worker productivity. Technical report, National Bureau of Economic Research.

Adhvaryu, A., Nyshadham, A., and Tamayo, J. (2019b). Managerial quality and productivity dynamics. Technical report, Mimeo, University of Michigan, Boston College and Harvard Business School.

Afridi, F., Dinkelman, T., and Mahajan, K. (2016). Why are fewer married women joining the work force in India? A decomposition analysis over two decades. Technical report, Institute for the Study of Labor (IZA).

Aizer, A. and Cunha, F. (2012). The production of child human capital: Endowments, investments and fertility. *Unpublished paper, Brown University*.

Ajisuksmo, C. R. and Vermunt, J. D. (1999). Learning styles and self-regulation of learning at university: An indonesian study. *Asia Pacific Journal of Education*, 19(2):45–59.

Algan, Y., Cahuc, P., and Shleifer, A. (2013). Teaching practices and social capital. *American Economic Journal: Applied Economics*, 5(3):189–210.

Altonji, J. G. and Spletzer, J. R. (1991). Worker characteristics, job characteristics, and the receipt of on-the-job training. *Industrial & Labor Relations Review*, 45(1):58–79.

Amodio, F. and Martinez-Carrasco, M. A. (2018). Input allocation, workforce management and productivity spillovers: Evidence from personnel data. *The Review of Economic Studies*, 85(4):1937–1970.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, pages 1481–1495.

Ashraf, N., Low, C., and McGinn, K. (2017). Negotiating a better future: Communication skills and inter-generational investment in Zambia. Technical report, mimeo.

Attanasio, O. P., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., and Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: Cluster randomized controlled trial. *BMJ*, 349:g5785.

Autor, D. H. (2001). Why do temporary help firms provide free general skills training? *Quarterly Journal of Economics*, pages 1409–1448.

Bandiera, O., Barankay, I., and Rasul, I. (2010). Social incentives in the workplace. *The review of economic studies*, 77(2):417–458.

Bandiera, O., Barankay, I., and Rasul, I. (2013). Team incentives: Evidence from a firm level experiment. *Journal of the European Economic Association*, 11(5):1079–1114.

Barrett, A. and O'Connell, P. J. (2001). Does training generally work? The returns to in-company training. *Industrial & Labor Relations Review*, 54(3):647–662.

Barron, J. M., Berger, M. C., and Black, D. A. (1999). Do workers pay for on-the-job training? *Journal of Human Resources*, pages 235–252.

Bartel, A. P. and Sicherman, N. (1998). Technological change and the skill acquisition of young workers. *Journal of Labor Economics*, 16(4):718–55.

Bassanini, A., Booth, A. L., Brunello, G., De Paola, M., and Leuven, E. (2007). Workplace training in Europe. *IZA Discussion paper*.

Bassi, V., Nansamba, A., and Liberia, B. (2017). Information frictions in the labor market: Evidence from a field experiment in Uganda. Technical report, mimeo.

Becker, G. S. (1964). *Human Capital*. New York: Columbia University Press.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300.

Betcherman, G., Dar, A., and Olivas, K. (2004). Impacts of active labor market programs: New evidence from evaluations with particular attention to developing and transition countries. *Social Protection, World Bank*.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics*, 122(4).

Blundell, R., Dearden, L., Meghir, C., and Sianesi, B. (1999). Human capital investment: The returns from education and training to the individual, the firm and the economy. *Fiscal Studies*, 20(1):1–23.

Borghans, L., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059.

Buvinić, M. and Furst-Nichols, R. (2016). Promoting women's economic empowerment: What works? *The World Bank Research Observer*, 31(1):59–101.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H. C., McKenzie, D., and Mensmann, M. (2017). Teaching personal initiative beats traditional training in boosting small business in West Africa. *Science*, 357(6357):1287–1290.

Card, D., Kluve, J., and Weber, A. (2017). What works? A meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931.

Chang, C. and Wang, Y. (1996). Human capital investment under asymmetric information: The Pigovian conjecture revisited. *Journal of Labor Economics*, pages 505–519.

Chatterjee, U., Murgai, R., and Rama, M. (2015). Job opportunities along the rural-urban gradation and female labor force participation in India. *World Bank Policy Research Working Paper*, (7412).

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Cunningham, W. V. and Villaseñor, P. (2016). Employer voices, employer demands, and implications for public skills development policy connecting the labor and education sectors. *The World Bank Research Observer*, 31(1):102–134.

Dearden, L., Reed, H., and Van Reenen, J. (2006). The impact of training on productivity and wages: Evidence from British panel data. *Oxford Bulletin of Economics and Statistics*, 68(4):397–421.

DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2016). Estimating social preferences and gift exchange at work. Technical report, National Bureau of Economic Research.

Deming, D. J. (2015). The growing importance of social skills in the labor market. Technical report, National Bureau of Economic Research.

Driskell, J. E., Salas, E., and Driskell, T. (2018). Foundations of teamwork and collaboration. *American Psychologist*, 73(4):334.

Edmonds, E. V., Feigenberg, B., and Leight, J. (2020). Advancing the agency of adolescent girls. Technical report, National Bureau of Economic Research.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M., and Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, 344(6187):998–1001.

Goux, D. and Maurin, E. (2000). Returns to firm-provided training: Evidence from French worker–firm matched data. *Labour Eonomics*, 7(1):1–19.

Grantham-McGregor, S. M., Powell, C. A., Walker, S. P., and Himes, J. H. (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican study. *The Lancet*, 338(8758):1–5.

Groh, M., Krishnan, N., McKenzie, D. J., and Vishwanath, T. (2012). Soft skills or hard cash? The impact of training and wage subsidy programs on female youth employment in Jordan. *World Bank Policy Research Working Paper*, (6141).

Groh, M., McKenzie, D., and Vishwanath, T. (2015). Reducing information asymmetries in the youth labor market of Jordan with psychometrics and skill based tests. *The World Bank Economic Review*, 29(suppl 1):S106–S117.

Guerra, N., Modecki, K., and Cunningham, W. (2014). Developing social-emotional skills for the labor market: The practice model. *World Bank Policy Research Working Paper*, (7123).

Heath, R. and Jayachandran, S. (2016). The causes and consequences of increased female education and labor force participation in developing countries. Technical report, National Bureau of Economic Research.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *The American Economic Review*, 103(6):1–35.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.

Heckman, J. J., LaLonde, R. J., and Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of Labor Economics*, 3:1865–2097.

Heckman, J. J. and Mosso, S. (2014). The economics of human development and social mobility. Technical report, National Bureau of Economic Research.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Technical report, National Bureau of Economic Research.

Huber, M. (2014). Identifying causal mechanisms (primarily) based on inverse probability weighting. *Journal of Applied Econometrics*, 29(6):920–943.

Huber, M., Lechner, M., and Mellace, G. (2016). The finite sample performance of estimators for mediation analysis under sequential conditional independence. *Journal of Business & Economic Statistics*, 34(1):139–160.

Ibarrarán, P., Kluve, J., Ripani, L., and Shady, D. R. (2015). Experimental evidence on the long-term impacts of a youth training program. *Institute for the Study of Labor (IZA) Discussion Papers 9136*.

Katz, E. and Ziderman, A. (1990). Investment in general training: The role of information and labour mobility. *The Economic Journal*, 100(403):1147–1158.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., et al. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, 60(2):184–189.

Konings, J. and Vanormelingen, S. (2015). The impact of training on productivity and wages: Firm-level evidence. *Review of Economics and Statistics*, 97(2):485–497.

Kurtz, B. E. (1990). Cultural influences on childrens cognitive and metacognitive development. In *Interactions among aptitudes, strategies, and knowledge in cognitive performance*, pages 177–199. Springer.

Kurtz, B. E., Borkowski, J. G., and Deshmukh, K. (1988). Metamemory and learning in maharashtrian children: Influences from home and school. *The Journal of genetic psychology*, 149(3):363–376.

Leuven, E. and Oosterbeek, H. (2008). An alternative approach to estimate the wage returns to private-sector training. *Journal of Applied Econometrics*, 23(4):423–434.

Macchiavello, R., Menzel, A., Rabbani, A., and Woodruff, C. (2015). Challenges of change: An experiment training women to manage in the Bangladeshi garment sector. Technical report.

McKenzie, D. (2017). How effective are active labor market policies in developing countries? A critical review of recent evidence. *The World Bank Research Observer*, 32(2):127–154.

Menzel, A. (2017). Knowledge exchange and productivity spill-overs in Bangladeshi garment factories. *CERGE-EI Working Paper Series*, (607).

Mincer, J. (1962). On-the-job training: Costs, returns, and some implications. *The Journal of Political Economy*, pages 50–79.

Montalvao, J., Frese, M., Goldstein, M. P., and Kilic, T. (2017). Soft skills for hard constraints: Evidence from high-achieving female farmers. *World Bank Policy Research Working Paper*, (8095).

Prada, M. F., Rucci, G., and Urzúa, S. (2019). Training, soft skills and productivity: evidence from a field experiment in retail. Technical report, IDB Working Paper Series.

Schoar, A. (2014). The importance of being nice: Supervisory skill training in the Cambodian garment industry. Technical report, working paper.

Staritz, C. (2010). *Making the Cut?: Low-income Countries and the Global Clothing Value Chain in a Post-quota and Post-crisis World*. World Bank Publications.

Stogdill, R. M. and Coons, A. E. (1957). Leader behavior: Its description and measurement.

Weidmann, B. and Deming, D. J. (2021). Team players: How social skills improve team performance. *Econometrica*, 89(6):2637–2657.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel data*. MIT press.

APPENDIX: NOT FOR PUBLICATION.

A Experiment Details

A.1 Experiment and Data Timeline

Figure A1: Timeline of Experiment and Data Collection

January 2013	Salary and Attendance Data Collection Starts
June 2013	Treatment Assignment Announcement and Productivity Data Collection Starts
July 2013	Training Program Starts (Pre and Post Module Testing During Training)
June 2014	Training Program Ends and Worker Survey Conducted
November 2014	Supervisor Survey Conducted
February 2015	Production Data Collection Ends
December 2015	Attendance and Salary Data Collection Ends

Figure A2: Data Type and Availability

Attendance & Retention	Daily (January 2013-December 2015)
Productivity	Daily (June 2013-February 2015)
Salary	Monthly (January 2013-December 2015)
Survey Outcomes	Cross-sectional (June 2014)

A.2 P.A.C.E. Modules

A.2.1 Training Module Detailed Description

Table A1 presents an overview of the modules included in the P.A.C.E. training program. The program spanned about 80 hours of training, but involved additional meetings for review sessions as well as introduction and conclusion sections. The core content sessions covered content regarding communication, problem-solving and decision-making, time and stress management, sanitation and hygiene, financial literacy, general and reproductive health, legal literacy and social entitlements, and execution excellence.

Below we provide a detailed description of the core training modules (the Problem Solving and Decision-Making module, the Communication module and the Time and Stress Management module) and the supplementary modules.

- **Problem Solving and Decision-Making**: This was the longest module (13 hours). The first session in this module was about 6 hours long, and included basic problem- solving skills training, including group discussions and role plays on how the group would solve a particular problem, and how this highlighted various approaches to problem-solving (self-reliance vs. reliance on

Table A1: P.A.C.E. Training Modules and Duration

Module Name	(Non-Exhaustive) Overview of Topics Covered	Aproximate Duration (hours)
Introductory Session	Ice-breaking games, overview of program topics and importance, program background and importance.	5
Communication	Basics and importance of communication, gender dynamics and bairriers in communication, communication in the workplace, home, and community.	9.5
Problem Solving and Decision Making (PSDM)	Basic concepts in PSDM, problem analysis and solution finding, creative thinking for solutions, problem-solving in groups and accountability, consensus-building at work, home, and in the community.	13
Time and Stress Management	Time management, stress management (including some exercises for stress management), positive thinking	12
Water, Sanitation, and Hygiene (WASH)	Sanitary practices, the importance of clean water to health, rights of access to water	6
Financial Literacy	Importance of savings, financial planning tools, savings options	4.5
General and Reproductive Health	Nutrition, reproductive health, mental and emotional health	10
Legal Literacy and Social Entitlements	Basics of the legal system and structure, womens' legal rights	8.5
Execution Excellence	Important aspects of workplace excellence like attention to quality, teamwork, and timeliness.	5
Two Consolidation Sessions of 90 minutes each	Review sessions	3
Closing Session	Celebratory conclusion of the program	5

others etc.). The trainers then emphasized that these approaches are complementary. The session also included skills training such as identifying a problem statement, identifying the cause of the problem, considering all possible solutions, and implementing learning by doing, followed by a group exercise to implement these steps. Finally, there were three application modules, one on the dynamics of problem solving, decision making and consensus building at work, a second on these applications for problem-solving at home, and the third on the same in the community.

- **Time and Stress Management**: This was also a long module (12 hours). In time management, the training began with an overview of the importance of time management. This was followed by exercises involving making a time-use chart, and discussing it with other participants and getting feedback, as well as giving feedback on other participants' time charts. This also involved reflection on what changes the participant could make to have their time allocation be closer to their desired time allocation. Following this, there was a goal-setting module (in which participants chose goals from a variety of different settings, such as a savings or workplace goal, and planned activities required to reach the goals) and a prioritization module (where they learned to classify tasks by priority). There were standalone sessions on goal-setting and prioritization separately as well, which included more in-depth training to apply the skills they learned before in the training. In the stress-management training, the first session focused on identifying stress, as well as its ubiquity. There was an exercise and a group discussion that focused on identifying stress in a situation, as well as healthy coping mechanisms for stress. The second session focused on positive thinking and the benefits of personal time, and several additional sessions included stress management exercises.

- **Communication**: This module was one of the three core modules (in addition to the Time and Stress Management module, and Problem-Solving and Decision-Making module). It included various role plays where participants were in turn assigned to practice different types of communication techniques (such as submissive relative to assertive communication), and also observe other participants and provide and receive feedback on which aspects of different communication seemed more effective. Additional exercises involved role-playing different situations where communication may be difficult, and brainstorming different communication techniques that might be effective. A third session focused on power dynamics in communication via role-playing, and three final sessions focused on the application of the techniques discussed in the workplace, at home, and in the community, respectively.

- **Execution Excellence**: The module began with an introductory discussion on the importance of factors affecting the quality of task completion – these comprised internal motivation, teamwork and effective workflow processes. This was followed by a time-bound, team exercise while being observed by the trainers, which was simulating the planning and execution of an imaginary garment order. After the exercise, there was a debriefing where workers reflected on the strengths of their teamwork and workflow processes that they had set up, as well as things they would do differently if they had to re-do the task. This debrief also included feedback from the trainers. Finally, there was a wrap-up discussion underscoring how high-quality work can improve workers' career outcomes as well as benefit the firm and the customers, and the importance of internal motivation in executing tasks well. There was also a discussion of how teamwork and effective processes can affect project success, and how successful teamwork involves complementing team members efforts and work.

- **Financial Literacy**: This was a relatively short (4-hour) module. The module began with a discussion of income relative to common expenditures, and how these expenditures may vary by income. It continued with emphasizing that financial literacy is the capacity to financially plan (expenditure and savings) for a secure future. The training concluded with a discussion of the importance of saving in helping cope with household shocks, and the importance of cultivating a habit of saving.

- **Health**: This module included a session on food and nutrition, as well as three sessions on reproductive and maternal health (one on reproductive health and planned pregnancy, another on staying health during a pregnancy and postpartum care, and the third on reproductive system diseases and associated stigmas). There final module was on mental health, and focused on the importance of mental health, that stigma could impact care-seeking, and that once overcome, seeking help for mental health issues could significantly impact a person's quality of life.

- **Legal Literacy**: The module began with an overview of basic laws and a session on how participants could seek basic legal help (such as file a police complaint). This was followed by a session on an overview of the marriage registration process as well as marriage laws and its protections for women, including in the cases of separation or divorce. The was also a session on domestic violence and child custody laws and another on sexual violence and child abuse laws. The overall goal of this module was to increase awareness of relevant laws and empower participants to seek the appropriate legal help as required.

- **Water, Sanitation, and Hygiene**: The first session emphasized the importance of clean water for health and a discussion of waterborne diseases, and a demonstration of rainwater harvesting. The session also focused on several techniques to make water safe for consumption, such as boiling and adding chlorine tablets. It also discussed appropriate techniques for waste disposal. The third session discussed personal hygiene practices such as hand-washing, and menstrual hygiene. The final session focused on increasing participants' awareness of safety issues around accessing clean water and sanitation, including information on government initiatives that facilitate this access (such as community initiatives for water pumps or toilets).

The dates spanned by each of the major modules is listed below (note that these dates differed slightly in each factory):

- **Communication**: July 7, 2013 to August 23, 2013

- **Problem-solving and decision-making**: August 30, 2013 to November 15, 2013

- **Time and stress management**: November 22, 2013 to January 18, 2014

- **Financial literacy**: February 3, 2014 to February 21, 2014
- **Health**: February 24, 2014 to March 28, 2014
- **Execution excellence**: April 11, 2014 to May 2, 2014
- **Legal literacy and social entitlements**: May 11, 2014 to June 1, 2014
- **Review Sessions**: June 8, 2014 to June 30, 2014
- **Closing Ceremony**: July 7, 2014 to July 31, 2014

B Additional Results

B.1 Retention

To estimate the impact of treatment on the additional number of days the firm receives from the worker, we first construct a binary working variable that is 1 if the worker was retained *and* is present in the factory on a given day and 0 otherwise. We then calculate the number of cumulative person days as measured by the cumulative running sum of this binary, defined at the daily level for each worker. We estimate impacts on this outcome by replacing retention on the left-hand side of equation 1 with cumulative person days.

The results from this analysis reveal no clear evidence of significant impacts on retention during or after training. We present these results here (Table B1).

B.2 Line-Level Productivity and Operation Complexity Results

As a further test of robustness of our main results, we present regression results using daily productivity and operation complexity at the production-line level instead of the individual-level.[50] Results are presented in Table B2. They are less precise since they include all workers on the line, not just treated workers, but are consistent with the individual-level results. The treatment effects for efficiency is statistically significant at the 10% level after treatment. The magnitude of the line-level treatment effect for efficiency is nearly 80% of the direct treatment effect. These results provide further evidence that the main results are not driven by differential attrition rates by treatment. Furthermore, they indicate that the firm gains not only higher individual-level productivity from training the treated workers, but that these workers enable the entire production lines on which they produce to become more productive.

B.3 Treatment Spillovers

Additional results on treatment spillovers are presented in Table B3.

B.4 Survey Outcomes

Table B4 presents estimates of the impact of P.A.C.E. treatment on four categories of survey outcomes.

[50]Note that these results include all workers on the production line, not just those that signed up for the program.

Table B1: Impacts of P.A.C.E. Treatment on Retention and Attendance

	(1)	(2)	(3)	(4)	(5)
	Retained	Cumulative Person Days	Present	Unauthorized Absent	Tardy
	1(Worker Still on ***Attendance*** Roster)	Sum of Days Working for Each Worker to Date	1(Worker Present in Factory Today if Stilll on Attendance Roster)	1(Worker Absent without Leave Today if Still on Attendance Roster)	1(Worker Arrived Late Today Relative to Other Workers on Line)
After X P.A.C.E. Treatment	0.00999	13.60	0.00426	-0.00588	-0.0117
	(0.0239)	(11.89)	(0.00789)	(0.00695)	(0.0219)
During X P.A.C.E. Treatment	0.0322	6.009	0.00712	-0.00515	-0.00271
	(0.0234)	(3.694)	(0.00637)	(0.00624)	(0.0183)
Announced X P.A.C.E.. Treatment	0.00833	0.713	0.0128	-0.0135	0.000793
	(0.0137)	(1.885)	(0.0112)	(0.0111)	(0.0149)
Observations	2,004,671	1,841,394	859,916	859,916	574,164
Control Mean of Dependent Variable	0.51	237.43	0.89	0.09	0.36

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. Retained dummy and Cumulative Person Days are defined for every worker date observation in the data and therfore the regressions do not require any weighting. For columns 3 through 5 observations are weighted in regressions by the inverse of the predicted probability of being retained (i.e., not yet attrited with non-missing data) in the sample that day from a probit regression of the retained dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. All specifications include month-year fixed effects to account for nonlinear time trends, as well as worker fixed effects.

Table B2: Impact of P.A.C.E. Treatment on Line-Level Daily Productivity and Operation Complexity

	(1)	(2)	(3)	(4)
	Efficiency	SAM (Operation Complexity)	Efficiency	SAM (Operation Complexity)
	Produced/Target	Standard Allowable Minute	Mean(Produced/Target)	Mean(Standard Allowable Minute)
	Retained Workers Only (still in factory in December 2015)		*Line-level (including all workers on line)*	
After X P.A.C.E. Treatment	0.0955	-0.00216	0.0573*	-0.00318
	(0.0623)	(0.0262)	(0.0308)	(0.0293)
During X P.A.C.E. Treatment	0.0525	-0.0200	0.0124	0.00470
	(0.0495)	(0.0231)	(0.0211)	(0.0210)
Fixed Effects	Worker X Garment X Operation	Worker X Garment X Operation	Line X Garment	Line X Garment
Weights	None			
Observations	106,511	106,511	81,258	81,258
Control Mean of Dependent Variable	0.542	0.565	0.513	0.573

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. Sample in columns 1 and 2 is restricted to only workers still retained in the factory by the end of observation. All samples are trimmed in these regressions to omit days in which the worker is observed for only a half a production day or less or days in which the worker is observed for more than 2 overtime hours as these are anomalous observations with imprecise production measures. These outliers make up only around 5% of the work-day observations. Line-level regressions in 3 and 4 include all workers on the line, even those who did not sign up for the lottery and those who were not trained. All specifications include month-year fixed effects to account for nonlinear time trends. All samples are trimmed in these regressions to omit days in which the worker is observed for less than a full production day of 8 hours and days in which the worker is observed for more than 1 overtime hour as these are anomalous observations with imprecise production measures.

Table B3: Spillovers on Co-Workers (Attendance, Productivity, and Career Advancement)

	(1)	(2)	(3)	(4)	(5)
Panel A: Retention, Production, and Advancement	Retained	Cumulative Person Days	SAM	Log(Gross Salary)	Promoted to Date
After X Spillover	-0.0105	4.416	-0.0135	0.0125***	0.0168
	(0.0249)	(12.80)	(0.0166)	(0.00387)	(0.0232)
During X Spillover	0.0184	5.982	-0.0103	0.00268**	0.0160
	(0.0258)	(4.328)	(0.0123)	(0.00115)	(0.0142)
Announced X Spillover	0.0116	3.371			-0.00214
	(0.0151)	(2.146)			(0.00414)
Observations	1,735,518	1,596,898	213,499	29,177	27,550
Control Mean of Dependent Variable	0.513	237.429	0.565	9.021	0.134
Panel B: Career Advancement	Expect Promotion Next 6 Mos	Skill Development Training	Production Award or Incentive	Skill Peer Self-Assessment	Co-Worker Self-Assessment
Spillover	-0.0239	0.0241	0.0203	0.112	0.138*
	(0.0516)	(0.0605)	(0.0240)	(0.0689)	(0.0772)
Observations	527	527	527	527	527
Control Mean of Dependent Variable	0.566	0.244	0.0310	5.287	5.267

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. All regressions are for sewing department workers only as spillover sample is not defined for non-sewing workers. Retained and cumulative man days are defined for every worker date observation in the data and therfore regressions do not require any weighting. Observations all other regressions are weighted by the inverse of the predicted probability of being retained (i.e., not yet attrited with non-missing data) in the sample that day from a probit regression of the retained dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. All specifications in Panel A include month-year fixed effects to account for nonlinear time trends, as well as worker fixed effects. The specification in column 3 of Panel A includes worker by item by operation fixed effects, corresponding to the results in column 1 of Table 4. Controls for survey outcome regressions in Panel B include demograhpic baseline variables from Table 1 (i.e., dummies for education levels, dummies for deciles of age distribution, and dummies for tenure in integer years), as well as unit fixed effects.

Table B4: Impacts of P.A.C.E. Treatment on Survey Outcomes

	(1)	(2)	(3)	(4)	(5)
Panel A: Financial Behaviors and Attitudes	Saving for Education	Saving for Other Reasons	Risk Preference Index	Time Preference Index	Insurance or Informal Risk-Sharing
P.A.C.E. Treatment	0.0794**	-0.0454	0.167*	-0.106	0.0617*
	(0.0312)	(0.0336)	(0.0871)	(0.0933)	(0.0350)
Control Group Mean of Dependent Variable	0.265	0.272	-0.0518	0.0194	0.628
Control Group Standard Deviation of Dependent Variable	0.442	0.446	0.933	0.955	0.484
Panel B: Government and Firm Entitlements	Gov. Pension	Gov. Subsidized Healthcare	Other Gov. Subsidy	Firm Entitlements	Community Self Help Group
P.A.C.E. Treatment	0.0250*	0.0230**	0.0135	-0.0272	-0.0282
	(0.0141)	(0.00960)	(0.0315)	(0.0351)	(0.0300)
Control Group Mean of Dependent Variable	0.0388	0.00647	0.120	0.142	0.152
Control Group Standard Deviation of Dependent Variable	0.194	0.0803	0.325	0.350	0.360
Panel C: Personality	Conscientiousness	Locus of Control	Perserverance	Extraversion	Self-Sufficiency
P.A.C.E. Treatment	0.0205	0.0253	-0.126	0.160**	0.0438
	(0.0732)	(0.0775)	(0.0772)	(0.0701)	(0.0883)
Control Group Mean of Dependent Variable	-0.0467	-0.0403	0.0202	-0.0713	-0.0634
Control Group Standard Deviation of Dependent Variable	0.931	0.947	0.974	0.971	1.029
Panel D: Mental Health and Aspirations	Self-Esteem	Hope/Optimism	Moderate Distress	Child's Expected Age at Marriage	Child Educated Beyond College
P.A.C.E. Treatment	-0.176*	-0.0607	-0.0413	0.0472	0.0887***
	(0.105)	(0.0824)	(0.0385)	(0.165)	(0.0282)
Control Group Mean of Dependent Variable	0.0479	0.0155	0.0939	23.43	0.117
Control Group Standard Deviation of Dependent Variable	1.016	0.964	0.292	2.084	0.321
Observations	621	621	621	621	621

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. Obersvations are weighted in regressions by the inverse of the predicted probability of being retained (i.e., not yet attrited with non-missing data) in the sample on the day of the survey from a probit regression in the attendance roster of the retained dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. Controls include demographic baseline variables from Table 1 (i.e., dummies for education levels, dummies for deciles of the age distribution, and dummies for tenure in integer years) as well as unit fixed effects.

B.5 Mediation Analysis

We follow Heckman et al. (2013) and Huber (2014) in conducting a mediation analysis to calculate the contribution of the estimated changes in personality characteristics presented in Panel C of Table B4 to the productivity impacts estimated in Table 2. We do so by combining treatment effects on mediators and productivity with estimated heterogeneity in productivity impacts by these mediators. We employ inverse probability weighting to account for endogeneity in mediators, as in Huber (2014).

First, we estimate the effect of the treatment on the productivity outcome, controlling for the full set of mediating variables:

$$Y_i = \beta_0 + \beta_1(PACEtreat)_i + \beta_2\mathbf{Z} + \varepsilon_i$$

where i indexes the individual, Y_i is the outcome variable for individual i, $(PACEtreat)_i$ the binary variable indicating whether individual i received the P.A.C.E. training, and **Z** the vector of mediating variables. All mediating variables are standardized to facilitate comparison.

Second, we estimate the effect of the treatment on the mediating variables, by regressing each mediating variable Z on the treatment, controlling for a vector of all other mediating factors $\mathbf{Z'}_i$:

$$Z_i = \phi_0 + \phi_1(PACEtreat)_i + \phi_2\mathbf{Z'}_i + \varepsilon_i.$$

Third, we estimate the incremental contribution of the mediating factors to the productivity treatment effect with the following specification:

$$Y_i = \alpha_0 + \alpha_1(PACEtreat)_i + \alpha_2\mathbf{Z} + \alpha_3\mathbf{Z}\cdot(PACEtreat) + \varepsilon_i$$

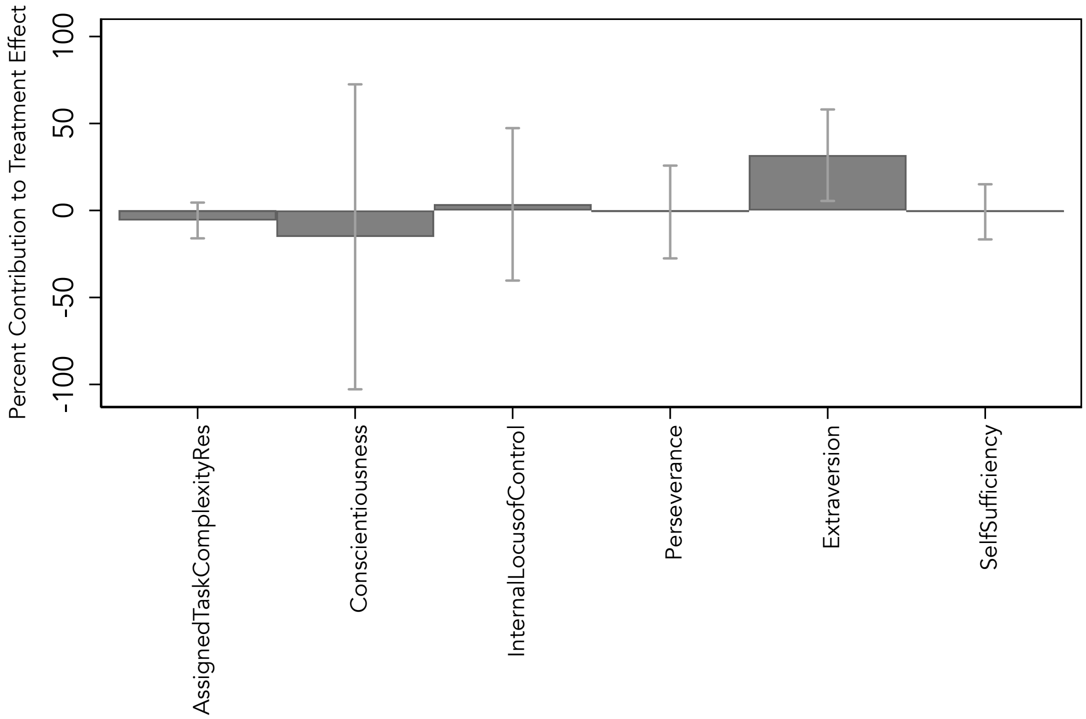
in which all mediators are included along with treatment dummy and the interactions of treatment dummy with each of the mediators.

To calculate the percentage contribution of each mediating factor to the total treatment effect of the P.A.C.E. training, for each mediating factor Z, we divide the product of its $\phi_1$and α_3 by the total effect of treatment on the outcome, β_1. This methodology needs three assumptions to hold: conditional independence of treatment, conditional independence of the mediator, and common support (Huber et al., 2016). The randomization of treatment guarantees the first assumption. There may be concerns with the second assumption, if there are baseline imbalances in the mediators or if they are endogenous. We address these concerns by constructing inverse probability weights and including them in all regressions described above (Huber, 2014). The third assumption can be verified by ensuring none of the mediators perfectly predicts treatment.

Finally, we draw 10,000 bootstrap samples of the original data with replacement and apply the three step procedure described above to each pseudo-sample to compute standard errors. Specifically, we run each of the above inverse probability weighted regressions and calculate the percentage contribution of each mediating factor to the total treatment effect for each pseudo-sample and use the standard deviation of the percent contribution across the 10,000 bootstrap replications as the standard error. We construct 95% confidence intervals using this standard deviation for each mediator.

The results presented in Figure B1 suggest that the large effect on extraversion may contribute

Figure B1: Percentage Contribution of Mediators to Treatment Impact on Productivity



meaningfully to the productivity gains, but the other contributions are unfortunately imprecisely estimated given the small sample in the survey. Accordingly we do not rely on the survey outcomes nor this mediation analysis to interpret the mechanisms by which the training produced productivity gains. Rather we base our interpretations largely on the heterogeneity and spillover analyses discussed in section 5. We note that improvement in extraversion as a result of the training as a mediator for productivity gains is also consistent with the preferred teamwork and collaboration skills interpretation discussed in section 5.

B.6 Survey Outcomes for Supplemental Propensity Score-Matched Sample

We also estimate treatment effects on a broader set of survey measures of soft skills and personality traits using a supplemental (non-experimental) sample of propensity score-matched trained and untrained workers. Though this analysis provides additional evidence for whether the estimated productivity impacts of the training are delivered by way of gains in soft skills, it has two key drawbacks: 1) workers in the original sample from the randomized experiment had mostly left the firm by the time we fielded this subsequent survey, so we are unable to leverage the randomized treatment assignment; 2) we do not observe productivity for the new propensity score matched sample in this survey, as factories discontinued the collection of worker-level productivity.

From the five factories that were part of the study, we surveyed all the female sewing machine operators in January 2019-February 2019. We first asked if they had already completed P.A.C.E. training, or if not whether were interested in participating in the training in the future. From each unit,

Table B5: Supplemental Propensity Score Matched Sample Balance Checks

	(1) Non-P.A.C.E.		(2) P.A.C.E.		Difference
Variable	N	Mean/(SE)	N	Mean/(SE)	(1)-(2)
Age (Years)	331	34.866 (0.399)	344	34.733 (0.387)	0.133
Kannada	331	0.810 (0.022)	344	0.797 (0.022)	0.013
Tenure (Years)	331	3.023 (0.109)	344	3.220 (0.110)	-0.197
1(Skilled)	331	0.360 (0.026)	344	0.311 (0.025)	0.048
1(High School Completion)	331	0.305 (0.025)	344	0.346 (0.026)	-0.041

Notes: None of the differences are statistically significant at convention levels. P-values from t-tests of differences are all above 0.10.

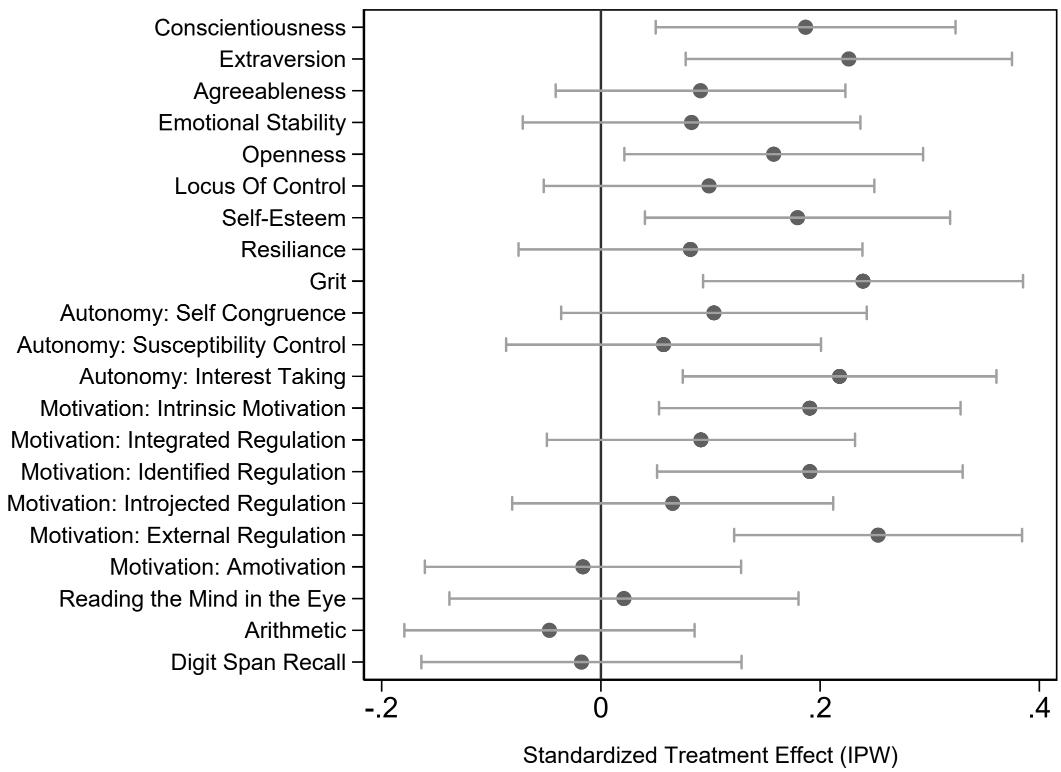
we matched the sample of workers who have completed the PACE training to the workers who have expressed interest in PACE using propensity score matching without replacement. We used for following variables to calculate the propensity score on which we matched: age of the worker, if the worker spoke the native language of the state (Kannada), tenure in the firm, if the worker is classified as skilled, and the workers' completed level of education. We then administered a comprehensive psychometric survey on this matched sample of workers between April 2019 to July 2019. Table B5 reports balance tests.

Figure B2 presents standardized treatment effects on 18 non-cognitive skills and personality traits as well as a measure of social cognition (i.e., reading the mind in the eye) and two other measures of cognitive skills (arithmetic and digit span recall) as placebos.[51] We find large and significant (roughly .2 standard deviations, significant at the 5% level) impacts on many of these survey measures, including openness and grit and several measures of autonomous functioning and motivation. We remeasured a few of the dimensions collected in the original experimental survey as a consistency check. We find a large and significant impact on extraversion of similar magnitude to that presented in Panel C of

[51]From top to bottom, the first 5 measures correspond to the elements of the five-factor model in psychology known as the Big Five. Locus of control is oriented such that a more positive score reflects a more internal locus of control, identical to the measure presented in Table B4 above. Self-esteem is also measured identically to that from the experimental sample survey. Resilience reflects the standardized score from the 6 question Brief Resilience Scale. Grit is meant to measure a combination of passion for perseverance in the pursuit of a goal and reflects the standardized score from the 10 question scale. The three autonomy measures are subscales obtained from the 15 question Index of Autonomous Function module. The six motivation measures are subscales of an 18 question motivation module meant to capture the different types of motivation emphasized in self-determination theory. The Reading the Mind in the Eye measure is the standardized number correct from the test of the same name meant to measure social cognition. The arithmetic measure is the standardized number correct from a timed arithmetic test we designed. The digit span recall measure is the maximum number of digits recalled correctly in a sequence of increasing length.

Table B4 (.2 as compared to .164 of a standard deviation). Given that we obtained imprecise estimates on conscientiousness and self-esteem from the original experimental evaluation despite finding significant effects on related measures like aspirations and self-assessment of skill, we exercised more care in the translation and training of surveyors on these modules of the survey. The results in Figure B2 indicate that these efforts indeed improved the measurement of these dimensions.

Figure B2: Propensity Score Matched Treatment Impacts on Survey Outcomes



Each of these dimensions is consistent with themes and topics emphasized throughout the core modules of the training.[52] For example, the Problem Solving and Decision-Making module, the longest of all modules, emphasized the importance of self-reliance in problem solving consistent with improvements in measures of autonomous functioning. The second longest module, Time and Stress Management, emphasized and practiced personal goal-setting and organizing and prioritizing tasks and activities in service of those personal goals, both crucial elements of external and identified regulation in motivation.

As discussed above, the final core module, Communication, introduced different types of communication (e.g., submissive vs. assertive) and had participants role-play to both assess and practice the most effective forms of communication in different scenarios. We interpret impacts on extraversion (and possibly self-esteem) to be reflective of these exercises. In addition, impacts on openness might reflect the emphasis on role-playing throughout several of the modules. Beyond these three

[52]We present a detailed description of the topics covered in and time devoted to each module in Section A.2.

core modules, additional sessions also addressed topics that map to measured skills and traits. For example, execution excellence explicitly focused on motivation and teamwork and linked planning, conscientiousness, and attention to detail in work to career goals. Additionally, the themes and topics emphasized across these modules, when taken together and reviewed and consolidated, as was done in the final two sessions of the program, map well to the combination of skills measured in grit, which reflected one of the largest standardized treatment effects in Figure 3.

We interpret this supplemental evidence of impacts on additional dimensions of non-cognitive skills and potentially productive traits as likely contributing to the portion of the productivity impacts left unexplained by the mediation analysis above. Unfortunately, given that we do not observe the same productivity data for this non-experimental sample, we are unable to confirm this interpretation with an analogous mediation analysis. We do, however, present additional evidence to support the validity of this supplemental evidence.

Note that social intelligence and cognitive measures, interpreted here as placebos, show small and insignificant differences between trained and untrained workers, supporting the validity of the comparison in this non-experimental exercise. We also demonstrate the robustness of these results to alternate estimation specifications (i.e., nearest neighbor fixed effects and no correction) as well as to corrections for multiple hypothesis testing in section C.2 and C.3, respectively. The pattern of results is nearly identical across specifications. Taken together, no detectable differences in the placebo measures and the consistency across estimates from different specifications strengthens our confidence in this supplemental analysis, despite the obvious caveat that variation in treatment is not randomized.

C Additional Checks and Robustness

C.1 Balance Tests by Baseline Characteristics at Different Points in the Study Period

Table C1: Summary Statistics: Balance Checks for Baseline Characteristics at Different Points in the Study Period

	(1)		(2)		(3)	
	Full Randomized Sample (Baseline)					
	Control		Treated		Difference	
Spillover Treatment	**Control Workers in Control Lines**		**Control Workers in Treatment Lines**			
Number of workers	779		837			
	Mean	SD	Mean	SD	Mean Difference	p value
Attendance Rate (Jan-May 2013)	0.898	0.088	0.903	0.085	0.005	0.50
High School	0.602	0.490	0.604	0.489	0.003	0.47
Years of Tenure	1.432	1.316	1.353	1.231	-0.079	0.36
Age	27.712	8.947	27.420	8.089	-0.292	0.70
1(Speaks Kannada)	0.657	0.475	0.671	0.470	0.014	0.45
High Skill Grade	0.506	0.500	0.518	0.500	0.012	0.85
log(Salary) (May 2013)	8.746	0.119	8.737	0.111	-0.009	0.02**
Efficiency (Announcement Month)	0.586	0.159	0.556	0.157	-0.030	0.67
SAM (Announcement Month)	0.618	0.255	0.615	0.221	-0.003	0.46
	One Month Post Treatment (July 2014)					
	Control		Treated		Difference	
P.A.C.E. Treatment	**Control Workers in Control Lines**		**Treated Workers in Treatment Lines**			
Number of workers	344		494			
	Mean	SD	Mean	SD	Mean Difference	p value
Attendance Rate (Jan-May 2013)	0.915	0.070	0.918	0.074	-0.003	0.56
1(High Education)	0.573	0.495	0.580	0.494	-0.007	0.84
Years of Tenure	1.760	1.372	1.569	1.210	0.191	0.14
Age	30.006	9.794	28.788	8.358	1.218	0.17
1(Speaks Kannada)	0.721	0.449	0.691	0.463	0.030	0.65
High Skill Grade	0.535	0.500	0.564	0.496	-0.029	0.55
log(Salary) (May 2013)	8.770	0.128	8.756	0.112	0.014	0.19
Efficiency (Announcement Month)	0.593	0.153	0.562	0.153	0.031	0.27
SAM (Announcement Month)	0.641	0.281	0.630	0.230	0.011	0.75
	Last Month of Data Collection (December 2015)					
	Control		Treated		Difference	
P.A.C.E. Treatment	**Control Workers in Control Lines**		**Treated Workers in Treatment Lines**			
Number of workers	179		250			
	Mean	SD	Mean	SD	Mean Difference	p value
Attendance Rate (Jan-May 2013)	0.918	0.068	0.923	0.068	-0.005	0.45
1(High Education)	0.542	0.500	0.542	0.499	0.000	1.00
Years of Tenure	1.711	1.240	1.661	1.156	0.049	0.73
Age	30.525	8.334	30.096	7.957	0.430	0.63
1(Speaks Kannada)	0.737	0.441	0.721	0.449	0.016	0.81
High Skill Grade	0.564	0.497	0.554	0.498	0.010	0.86
log(Salary) (May 2013)	8.789	0.135	8.769	0.121	0.020	0.21
Efficiency (Announcement Month)	0.587	0.137	0.560	0.151	0.027	0.33
SAM (Announcement Month)	0.640	0.279	0.627	0.232	0.013	0.73

Notes: Tests of differences calculated using errors clustered at the line level according to the experimental design.

C.2 Alternative Estimates from Supplemental Non-experimental Sample

Figure C1: Nearest Neighbor Matched Treatment Impacts on Survey Outcomes

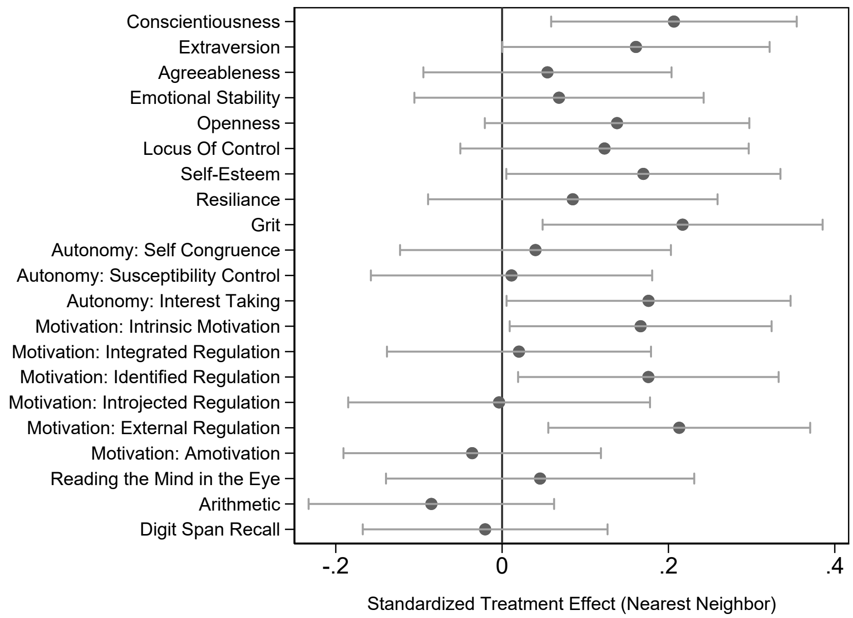
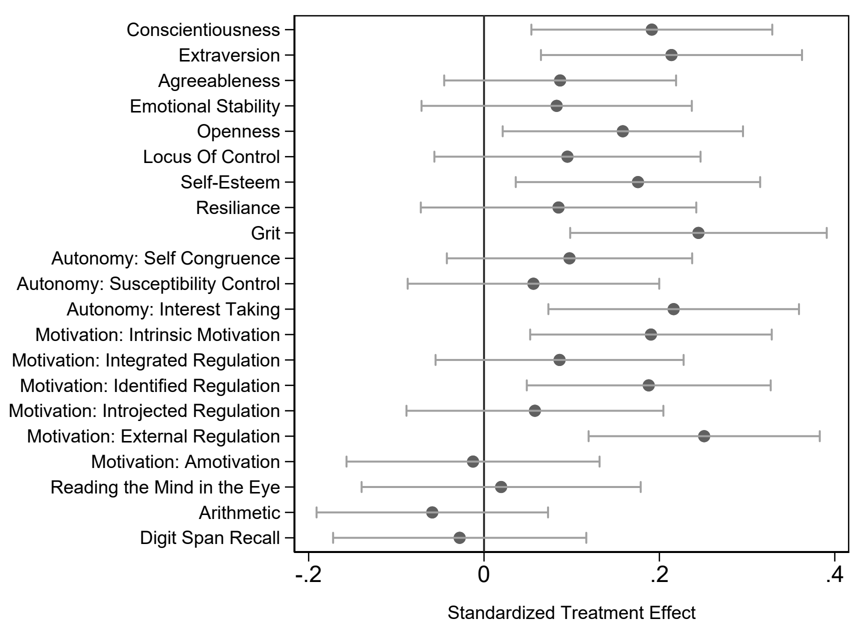


Figure C2: Unmatched Treatment Impacts on Survey Outcomes



C.3 Correction for Multiple Hypothesis Testing

In Table C2, we re-estimate the direct impacts of the P.A.C.E program on the main outcomes, correcting for multiple hypothesis testing. The regression specifications are identical to the analogous regressions in the main tables; however, in place of standard errors, we report (corrected) q-values (false discovery rates) in parentheses in this table. Each panel of the table corresponds to a set of hypothesis - for instance, we test all the productivity outcomes (efficiency and operation complexity) as one set of hypotheses, all workplace survey outcomes as another set of hypotheses, and so on. To correct the p-values for multiple hypothesis testing, we follow Anderson (2008) who recommends using the methodology of Benjamini and Hochberg (1995). This method controls the False Discovery Rate (FDR) at level q when there are M hypothesis to be tested (say $H_1, ...H_M$), by sorting the corresponding p-values in increasing order ($p_1 < ...p_M$), and rejecting c hypotheses such that c is the largest w where $pw < (qw/M)$.[53]

Overall, the significance of the main results is preserved for the set of workplace outcomes, albeit less so with the non-workplace survey outcomes. The retention and productivity impacts exhibit almost no differences in significance in Panels A and B, respectively, when the corrections for multiple hypothesis are done.[54] Workplace survey outcomes in Panel C and government and firm entitlements in Panel E also show very similar significance to the main results. Outcomes in Panels D, E and F show small increases in p-values (or q-values). For example, in the set of measures related to financial behaviors and attitudes, the positive impact on savings for children's education is significant at the 10% level in Table C2, and at the 5% level in Table 5; while, the set of personality outcomes produces a marginally insignificant positive impact of P.A.C.E. on extraversion with p-value of .108 after the correction is applied, as compared to an estimate that was significant at the 5% level in the main results. As in the uncorrected regressions, there are no statistically significant impacts on mental health, but the impact on aspirations for one's childrens' education remains positive and strongly statistically significant.

Table C3 reports treatment effects and multiple hypotheses corrections for the supplemental non-experimental sample.

[53]To implement this procedure, we use the Stata code available here: `https://are.berkeley.edu/~mlanderson/ARE_Website/Research.html`

[54]We report working and person day outcomes from the attendance dataset only for brevity, but similar equivalence is obtained when analyzing production data analogues.

Table C2: Robustness to Correction for Multiple Hypothesis Testing (Anderson, 2008)

	(1)	(2)	(3)	(4)	(5)
Panel A: Productivity and Retention	Efficiency	SAM (Operation Complexity)	Retained	Cumulative Person Days	
After X P.A.C.E. Treatment	0.108**	0.0384**	0.0062	9.25	
	(0.049)	(0.049)	(0.81)	(0.81)	
During X P.A.C.E. Treatment	0.03	0.0334**	0.0264	5.360	
	(0.27)	(0.05)	(0.22)	(0.21)	
Announced X P.A.C.E.. Treatment			0.00416	0.501	
			(0.76)	(0.76)	
Panel B: Workplace Survey Outcomes	Expect Promotion Next 6 Mos	Skill Development Training	Production Award or Incentive	Peer Self-Assessment	Line Co-Worker Self-Assessment
P.A.C.E. Treatment	0.0871*	0.158***	0.0293	0.122	0.0645
	(0.095)	(0.006)	(0.15)	(0.105)	(0.37)
Panel C: Financial Behaviors and Attitudes	Saving for Education	Saving for Other Reasons	Risk and Time Preference Index	Insurance	Informal Borrow or Lend
P.A.C.E. Treatment	0.0804*	-0.0465	0.166	-0.0984	0.0637
	(0.06)	(0.21)	(0.12)	(0.30)	(0.12)
Panel D: Government and Firm Entitlements	Gov. Pension	Gov. Subsidized Healthcare	Other Gov. Subsidy	Firm Entitlements	Community Self Help Group
P.A.C.E. Treatment	0.0248	0.0226*	0.0119	-0.0257	-0.0270
	(0.20)	(0.09)	(0.70)	(0.58)	(0.58)
Panel E: Personality	Conscientiousness	Locus of Control	Perserverance	Extraversion	Self-Sufficiency
P.A.C.E. Treatment	0.0210	0.0307	-0.123	0.164	0.0445
	(0.76)	(0.78)	(0.29)	(0.108)	(0.78)
Panel F: Mental Health and Aspirations	Self-Esteem	Hope/Optimism	Moderate Distress	Child's Expected Age at Marriage	Child Educated Beyond College
P.A.C.E. Treatment	-0.172	-0.0621	-0.0422	0.0456	0.0885**
	(0.27)	(0.56)	(0.47)	(0.78)	(0.01)

Notes: p-values adjusted for multiple hypothesis testing, q-values (false discovery rates) in parentheses (*** q<0.01, ** q<0.05, * q<0.1). Standard errors are clustered at the line level. The methodology from Anderson (2008) was used to correct for multple hypothesis testing. Specifications are otherwise identical to analogous regressions in main results tables. For conciseness, weights, fixed effects, and controls are not mentioned here, but are included in regressions where noted in analogous main tables. Similarly, observations and control means of dependent variables are omitted as well, but identical to those from main tables. For the first panel, all three outcomes (retention, working, and cumulative man days) from the attendance data is treated as one set of outcomes, and the retention information from the salary data and working and cumulative person days information from the production data together as another set of outcomes.

Table C3: Supplemental Sample Including Robustness to Corrections for MHT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Conscientiousness	Extraversion	Agreeableness	Emotional Stability	Openness	Locus Of Control	Self Esteem	Resiliance	Grit	Autonomous Functioning	Combined Intrinsic Motivation	Mind in the Eye	Arithmetic	Digit Span Recall
Panel A: Propensity Score Inverse Probability Weighted														
P.A.C.E. Treatment	0.187	0.226	0.0909	0.0827	0.158	0.0986	0.179	0.0817	0.239	0.147	0.0430	0.0210	-0.0469	-0.0177
	(0.008)***	(0.003)***	(0.18)	(0.29)	(0.02)**	(0.20)	(0.01)**	(0.31)	(0.001)***	(0.05)*	(0.55)	(0.80)	(0.49)	(0.81)
	[0.04]**	[0.02]**	[0.35]	[0.43]	[0.07]*	[0.35]	[0.04]**	[0.43]	[0.02]**	[0.11]	[0.65]	[0.81]	[0.62]	[0.81]
Fixed Effects	Unit, Education, Age, Tenure													
Weighted	Inverse Propensity Score													
Observations	675													
Panel B: Nearest Neighbor Matched														
P.A.C.E. Treatment	0.207	0.161	0.0545	0.0684	0.138	0.123	0.17	0.085	0.217	0.133	0.0756	0.0457	-0.0851	-0.0204
	(0.006)***	(0.05)*	(0.47)	(0.44)	(0.09)*	(0.16)	(0.04)**	(0.34)	(0.01)**	(0.11)	(0.32)	(0.63)	(0.26)	(0.79)
	[0.09]*	[0.18]	[0.55]	[0.55]	[0.25]	[0.33]	[0.18]	[0.48]	[0.09]*	[0.26]	[0.48]	[0.68]	[0.46]	[0.79]
Fixed Effects	Unit, Education, Age, Tenure, Nearest Neighbor Pair ID													
Weighted	None													
Observations	662													
Panel C: No Selection Correction														
P.A.C.E. Treatment	0.191	0.214	0.0868	0.0829	0.158	0.0951	0.176	0.0850	0.245	0.144	0.0426	0.0195	-0.0590	-0.0277
	(0.006)***	(0.005)***	(0.20)	(0.29)	(0.02)**	(0.22)	(0.01)**	(0.29)	(0.001)***	(0.05)*	(0.56)	(0.81)	(0.38)	(0.71)
	[0.03]**	[0.03]**	[0.385]	[0.41]	[0.07]*	[0.38]	[0.05]*	[0.41]	[0.02]**	[0.12]	[0.65]	[0.81]	[0.49]	[0.76]
Fixed Effects	Unit, Education, Age, Tenure													
Weighted	None													
Observations	675													

Notes: Uncorrected p-values in round parentheses, p-values corrected using Anderson (2008) methdology in square parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. Obersvations are weighted in regressions by the inverse of the predicted probability of being retained (i.e., not yet attrited with non-missing data) in the sample that day from a probit regression in the attendance roster of the retained dummy on month by year FE and their interaction with individual and line treatment dummies and baseline variables reported in Table 1. Controls include demograhpic baseline variables from Table 1 (i.e., dummies for education levels, dummies for deciles of age distribution, and dummies for tenure in integer years).