Returns to On-the-job Soft Skills Training*

Achyuta Adhvaryu†
Namrata Kala‡
Anant Nyshadham§

December 2019

Abstract

We evaluate the causal impacts of on-the-job soft skills training on the productivity of Indian garment workers. Treated workers were 20 percent more productive than controls after the program. Productivity gains were larger among workers with lower baseline leadership skills, lower educational attainment, and greater technical skills. Wages rose only modestly with treatment (by 0.5 percent), with no differential turnover, suggesting that although soft skills raised workers' marginal products, there was a substantial wedge between productivity and wages, consistent with frictions in the low-wage labor market. The net return to the firm was 258 percent eight months after program completion.

Keywords: soft skills, non-cognitive skills, worker training, productivity, India
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, BREAD, NBER, JPAL; kala@mit.edu
§University of Michigan & NBER; nyshadha@umich.edu

1 Introduction

Soft skills – e.g., teamwork, leadership, relationship management, effective time allocation, and the ability to assimilate information – 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). Surveys of employers from around the world corroborate that soft skills are in great demand, and that firms often struggle to find workers with high levels of these skills (Cunningham and Villaseñor, 2016).

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). 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). It is not obvious that inculcating these skills in a meaningful way is possible: 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).

Moreover, when general training is delivered within the firm (as it often is[1]), 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. In perfectly competitive markets, workers' wages would need to increase commensurate to their marginal products; any firm that paid below marginal product would lose the newly trained workers as they received higher wage offers at other 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). 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 fourfold. First, is it possible to improve soft skills meaningfully for workers with low stocks of these skills? Second, if skills do improve, what are the causal impacts on workplace outcomes, including productivity, wages, and retention? Third, how does training in soft skills interact with baseline stocks of skills (both soft and technical) when generating improvements in workplace outcomes? That is, for which types of workers (e.g., those with lower soft skills at baseline or those with some degree of technical skill) is the value of training in soft skills largest? Finally, does it pay for firms to provide on-the-job soft skills training to workers, and what does this rate of return tell us about the nature of labor market frictions as pertains to soft skills?

To answer these questions, we partnered with the largest ready-made garment export firm in India

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

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 successful task execution. These skills are important inputs into production in the ready-made garments context. Workers produce in teams, and need effective communication 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., machine malfunction, requesting breaks or help to complete tasks, etc.). They need skills for personal goal-setting, planning, and prioritization to maintain motivation to achieve hourly and daily production targets. And they need problem-solving frameworks to effectively deal with daily shocks to production.

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 untrained workers on treatment lines to control workers on control lines. We assessed the impacts of soft skills training on 1) direct and indirect measures of the stocks of these skills; and 2) retention, productivity, wages, task complexity, and other workplace outcomes. Finally, we computed the firm's returns, combining our point estimates with data on the program's costs and the firm's accounting profits.

Endline survey results for treated and control workers and pre/post-module testing of treated workers indicate that stocks of soft skills improved in several important dimensions. Specifically, treated workers showed a pronounced increase in extraversion, which may impact productivity via improvements in the ability to communicate and solve issues collaboratively with peers and supervisors.[2] Survey results indicate greater self-assessment of workplace quality (relative to peers of the same technical skill grade), consistent with an increase in self-regard. Pre/post data from assessment tools designed to measure learning in each of the program's modules show that initial stocks of knowledge in each of the program's target areas were low, and that treated workers substantially improved these stocks through the program (most markedly for communication skills). Finally, a broader set of soft skills were measured for a supplemental non-experimental sample of trained and untrained workers matched via propensity score. These data reveal higher grit, openness, autonomy, and motivation among trained workers, in addition to the same gains in extraversion found in the experimental sample. Each of these dimensions is consistent with themes and topics emphasized throughout the core

[2]Mediation analysis indicates that roughly 30% of the impact on productivity is due to increases in extraversion. These workers were also more likely to request and complete technical skill development training, generating complementary improvements in "hard" skills. Mediation analysis indicates that the assignment to more complex tasks as result of this technical skilling actually contributes negatively to impacts on productivity, though only modestly, with increases in soft skills helping to overcome any losses in productivity due to these more complex task assignments.

modules of the training.[3]

Direct impacts on workplace outcomes, measured using the firm's administrative data, are consistent with the acquisition of soft skills by workers. Our measure of productivity, efficiency, is the number of garments produced divided by the target quantity of garments at the daily level. Treated workers are more productive by about 11 percentage points (20% higher than the control mean) and more likely to be assigned to complex tasks. Impacts were largest among those with a lower stock of leadership skills at baseline, consistent with improvements in soft skills driving gains in productivity. Productivity gains also persist up to 8 months after program completion (when we ceased data collection), suggesting that learned skills translated into lasting improvements in workplace outcomes. Workers on treatment lines who did not receive the program are also (weakly) more productive and are assigned to more complex operations, generating line-level impacts on productivity post-program completion. Wages went up only slightly as a result of treatment – an increase of about 0.5 percent. The program had no sustained impact on turnover or attendance.[4]

Taken in sum, we interpret the results to indicate that the program increased workers stocks of soft skills, which in turn led to productivity improvements. Combined with the fact that there was essentially no impact on wage or long-run turnover, our results are consistent with the presence of labor market frictions that prevent workers from capturing more of the productivity rents that ensue from training (Acemoglu, 1997; Acemoglu and Pischke, 1999). Soft skills are largely unobserved in the hiring and wage-setting process in this setting 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 caused by training) among both current and future potential employers likely generates the observed difference in impacts of soft skills training on marginal 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 73% by the end of the program period. Eight months after program completion, fueled by post-program increases in productivity, the return climbed to over 250%. 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 garment factories in South Asia (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.,

[3]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 observed 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. 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. 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 the grit questionnaire.

[4]Retention was actually higher in the treatment group relative to control during the program period; this effect diminished after program completion. We use a dynamic inverse probability weighting procedure, described in detail in section 4, throughout our analysis to correct for potential changes in the size and composition of the treatment and control groups over time.

2017; Ashraf et al., 2017; Campos et al., 2017; Groh et al., 2012; Schoar, 2014). These previous studies are mostly 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, tying 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.[5] This individual-level analysis allows us to document for whom soft skills training produces the largest impacts on productivity. We find that training is most impactful for those workers identified by factory HR representatives to be lacking in leadership skills at baseline, indicating that training in soft skills was a substitute rather than a complement for baseline stocks of skill. We also find that productivity gains were significantly larger for workers with less traditional education, consistent with the notion that these productive soft skills are imparted to a certain extent via traditional schooling. Additionally, productivity gains are concentrated among machine operators with some technical skilling, suggesting that in factory settings like ours, soft skills likely complement technical skills in the determination of productivity.

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). 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). 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 less educated machine operators with low baseline stocks of these skills.

2 Context, Program Details, and Experiment Design

2.1 Context

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. Women comprise the majority of the workforce in garment factories, and new labor force entrants tend to be disproportionately female (Staritz, 2010).

Garments are usually sewn in production lines consisting of around 50-70 workers arranged in sequence. Most of the workers on the line are assigned to machines completing sewing tasks (one

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

person to a machine). The remaining workers perform complementary tasks to sewing, such as folding or aligning the garment to feed it into a machine. Each line produces a single style of garment at a time. The line is 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). 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.

2.2 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, problem-solving, and financial literacy for its trainees. The program began with an introductory ceremony for participants, trainers, and firm management. The core modules were: Communication (9.5 hours); Problem Solving and Decision-Making (13 hours), and Time and Stress Management (12 hours). Additional modules included Execution Excellence (5 hours); Financial Literacy (4.5 hours); and Legal Literacy and Social Entitlements (8.5 hours).[6] Table A1 provides an overview of the topics covered in each module. 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 personal and professional situations. At the close of the program there was a graduation ceremony.

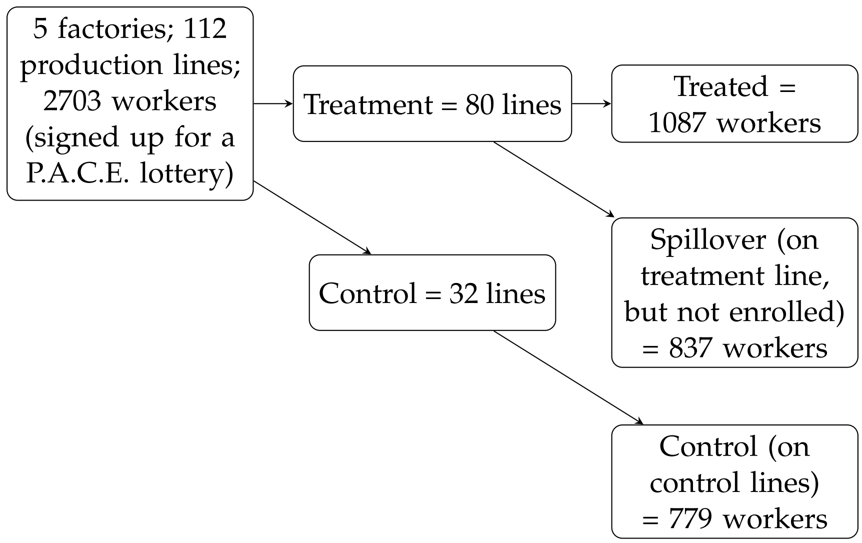
Workers participated in two hours of training per week. One hour of workers' production time a week to the program was allocated to the training, and workers contributed one hour of their own time. Training sessions were conducted at the beginning of the production day in designated classroom spaces 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. Accordingly, the training groups were balanced in size with no more than 3-4 employees from a given line in each group.

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

[6] Additional modules on Water, Sanitation and Hygiene (6 hours) and General and Reproductive Health (10 hours) were also included, but were not considered core modules. Pre/post assessments were not conducted for ancillary modules such as sanitation.

in July 2014.[7]

Figure 1: Experimental Design



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 workers 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.[8] 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.[9]

Figure 1 presents a schematic diagram of the experimental design.[10] 2703 workers signed up for the treatment lottery, from which 1087 were chosen for treatment. Out of the 1616 untrained workers,

[7] See appendix A for more details regarding the program.

[8] 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.

[9] 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.

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

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 our primary control. The latter group (untrained workers in treatment lines) is used to estimate treatment spillovers. Summary statistics and balance checks are discussed in Section 3.4.[11]

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. These data then automatically wirelessly synced to the server. 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. Note that though productivity was being recorded prior to the program implementation, the worker-hourly level data was not kept prior to the introduction of the tablet computers for production writing but rather discarded after line-daily level aggregate measures were input into the data server. Accordingly, line-daily level 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.

3.1.1 Productivity

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).[12] At the worker-hour level, we define "pieces produced" as the number of garments that passed a worker's 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 worker's station that had completed placket attachment by the end of a given production hour would be recorded as that worker's 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.

[11]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).

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

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 task specific increments. A line with 60 machine operators then would have an average worker-hourly SAM of 0.5.[13] As the name suggests, this measure is standardized across the global garment industry and is drawn from an industrial engineering database.[14] 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 $60/.5 = 120$. As mentioned in the previous section, 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.[15]

3.2 Human Resources Data: Attendance and Salary

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 indicates current grade level. 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) (Adhvaryu et al., 2019). 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). Workers can request to have their grade reassessed 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. The salary data are available until six months post-program completion, unlike the productivity and attendance data, which are available for eight months after program completion. 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.

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.[16] The survey covered,

[13]Mean SAM across worker hourly observations is 0.61 with a standard deviation of 0.20.

[14]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).

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

[16]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

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	0.616	0.843	0.642	0.688	-0.026	0.473
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
Spillover Treatment	**Control Workers in Control Lines**		**Control Workers in Treatment Lines**			
Number of workers	779		837			

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

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.[17] Finally, the survey covered worker's 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.

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 task complexity in the announcement month (the first month of observation for these outcomes).

We fail to reject that the difference between treated and control workers for any of these outcome means at baseline is statistically significantly different from zero. Average attendance rates are about

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.

[17] 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 worker's level of agreement with the statement. Survey questions are available on request.

90%, and average tenure with the firm is about 1.4 years. The average worker is about 27-28 years old. Over 60% of both samples are high school educated and speak Kannada.

The summary statistics and differences presented in Table 1 apply to the direct treatment comparison. Analogous balance checks for spillover comparisons were performed as well. We find no significant differences, and do not present them here for the sake of brevity.

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 and pay, and then for differences in survey measures of self-reported personal and professional outcomes. We follow this section with a discussion of potential mechanisms, including an investigation of heterogeneity in treatment impacts on productivity by baseline leadership and technical skills as well as traditional schooling, a mediation analysis on the experimental sample, and additional survey results from a supplemental non-experimental matched sample of trained and untrained workers.

4.1 Retention

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

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

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, $1[T_w]$ is a dummy variable that takes the value 1 if the worker is a trained worker on a treatment line and 0 if she is a control worker on a control line, and it is interacted 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 factory x year x month fixed effects ψ_{uym} (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 - 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.[18]

We conduct a variety of tests and conclude that there is little discernible effect on the size or composition of retained workers over the observation period, and discuss implications of these results for the analysis of conditionally observed outcomes used later in the analysis.

[18] We designed the experiment to allow measurement of spillover effects, and find some evidence suggesting such spillover effects occurred.

4.2 Dealing with Potential Bias from Selective Attrition

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 (February 2015). Results are presented in Table C1 and demonstrate that all baseline characteristics are balanced on means at both points in time.[19]

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. For example, if there exists 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 can weight all observations prior to that time by the probability of being able to measure the outcome at each point in time. Accordingly, 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, which 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), on the treatment indicator interacted with month by year fixed effects and baseline

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

characteristics (attendance, education, tenure, age, skill grade, productivity and task complexity).[20] 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.[21]

4.3 Attendance

Along with retention, the attendance roster allows for estimation of treatment impacts on additional outcomes of interest such as attendance (a binary variable that is 1 if the worker is at work today and 0 if not). 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 little evidence of impacts on these outcomes.

4.4 Productivity and Task Complexity

Next, we investigate treatment impacts on two key outcomes of interest from the productivity data: efficiency and SAM. As discussed above, SAM measures task complexity, and efficiency is actual pieces produced divided by target pieces (the latter being calculated from SAM). All of these variables are only measured if a worker is retained by the factory, and present in the factory that day. Accordingly, these conditionally observed outcomes are weighted in the analysis as discussed above. The weights are obtained as discussed in section 3 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.

In the SAM regressions, we follow the above specification in equation 1 exactly. However, in the efficiency regression, we replace the worker fixed effects with worker by garment style fixed effects. These are to account for any treatment impacts on the task complexity as identified in the SAM regression. We also include as additional controls days that the style has been running on the production line and total order size to account for learning dynamics at the line level that might impact worker productivity across the life of the order.

The results from these regressions are presented in Table 2. Treated workers are more efficient after the program (relative to the month of treatment assignment announcement) by nearly 11 percentage points, about 20% relative the control group mean (column 1). Impacts on productivity are stronger after program completion, the during treatment coefficient less than one third the size of the after treatment coefficient.[22]

[20]Since workers salaries are homogenous within skill grade level, grade proxies for skill level as well as salary.

[21]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 generally quite 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.

[22]The fact that workers are absent from the production line for one hour per week raises the concern that productivity gains from the program may arise because workers may be happy with the reduction in working time, or more efficient in the remainder of the time given a shorter work day. However, the productivity gains only 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 data collection ended). There are no significant productivity gains during the training period.

Table 2: Impacts of P.A.C.E. Treatment on Productivity

	(1)	(2)
	Efficiency Produced/Target	SAM (Operation Complexity) Standard Allowable Minute
After X P.A.C.E. Treatment	0.108**	0.0384**
	(0.0510)	(0.0180)
During X P.A.C.E. Treatment	0.0300	0.0334**
	(0.0274)	(0.0147)
Additional Controls	Days on Same Line-Garment, Total Order Size	None
Fixed Effects	Unit X Month X Year, Worker X Garment	Unit X Month X Year, Worker
Weights	Inverse Predicted Probability from Probit of Working on Treatments X Mo-Yr X Baseline Characteristics	
Observations	290,763	290,763
Control Mean of Dependent Variable	0.542	0.565

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Standard errors are clustered at the line level. Observations in columns 1 and 2 are weighted in regressions 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 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.

Additionally, we find that treated workers are assigned to more complex tasks both during and after treatment (tasks to which they are assigned are expected to take about 2.3 seconds (0.038 minutes) more, roughly 7% of the control group mean, as presented in column 2). Thus, not only are workers in the treatment group assigned to more complex tasks during and after the program, they are more productive even at these harder tasks once treatment ends as a result of the training in soft skills they received. We explore the contributions of both the assignment to more complex task assignment and gains in some measured soft skills to the large and significant increase in productivity below in section 5.4.

4.4.1 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 15 percentage points higher efficiency after the treatment, supporting the notion that productivity impacts are not driven by changing composition of the sample over time.[23] We also

[23] 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.

present in the appendix results for productivity and task 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 roughly 8% of the mean.

4.5 Salary, Career Advancement, and Career Expectations

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 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. Except for the salary data which is 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, we first estimate the retention probability weights as detailed in section 3, and then estimate equation 1 using those inverse probability weights, with the log of gross salary as the outcome.[24] 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 on average less than half a percent more wages in the period after the program completion, which translates to roughly 30 INR or less than 0.5 USD a month. Thus, despite being assigned to more complex tasks and being more productive, treated workers are not paid meaningfully higher wages.

Columns 2-6 of Table 3 presents the results from analyses of related survey outcomes. Treatment workers are about 8.7 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

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

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

	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Gross Salary)	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	0.00492*					
	(0.00270)					
During X P.A.C.E. Treatment	0.00137					
	(0.000906)					
Announced X P.A.C.E. Treatment	0.000221					
	(0.000647)					
P.A.C.E. Treatment		0.0871**	0.158***	0.0293	0.122*	0.0645
		(0.0414)	(0.0467)	(0.0185)	(0.0648)	(0.0667)
Fixed Effects	Unit X Month X Year, Worker			Unit, Education, Age, Tenure		
Weights	Inverse Predicted Probability from Probit of Retention on Treatments X Mo-Yr X Baseline Characteristics					
Observations	28,692	621	621	621	621	621
Control Mean of Dependent Variable	8.909	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. Controls in regressions for survey outcomes 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).

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 5).

5 Mechanisms

Our interpretation of the productivity and task complexity results is that skills and learned traits like communication and effective teamwork; problem solving and decision-making; time and stress management; and motivation and aspirations are "soft" inputs into production. Reinforcing these skills through the P.A.C.E. program should thus directly affect workplace outcomes. Across the categories of results presented below, impacts are consistent with a direct treatment effect on the stock of soft skills. In particular, the narrative that emerges is one that is consistent with the P.A.C.E. program increasing the stock of soft skills. This is indicated in part by the fact that treated workers exhibit more extraversion, are more likely to seek out and avail themselves of government and employer benefits to which they are entitled, and are also more likely to exhibit forward-looking behavior via savings and aspirations for their children's futures. They are also more likely to proactively increase their stock of hard skills by requesting technical training and to be assigned to more complex tasks as a result, but exhibit greater productivity despite this assignment to more complex tasks.[25]

Below, we provide support for this interpretation using evidence from a survey of treatment and control workers; from assessments of the treatment group's knowledge before and after the completion of the program's core modules; from additional survey data on soft skills and personality traits from a supplemental non-experimental sample of matched trained and untrained workers; and from estimates of treatment spillovers. We present evidence of heterogeneity in productivity impacts by stocks of baseline skills and conduct a mediation analysis to decompose treatment effects on productivity into components that can be attributed to measured changes in soft skills and assigned task complexity. We also review several alternative interpretations and discuss the plausibility of each in more detail in the appendix.

5.1 Survey Results

The first piece of evidence supporting a change in the stock of soft skills as a result of the training comes from a survey we administered to treatment and control workers in the month after program completion. We consider the impact of the program on survey outcomes that might plausibly reflect the skills taught by P.A.C.E. For instance, since the program targets the stock of non-cognitive skills such as the ability to acquire and use information more effectively, we consider outcome variables regarding whether workers avail themselves of government and firm welfare programs like pension schemes and subsidized health-care. Similarly, we test whether there is an increase in workers' savings, especially for important future considerations like education (their own or their children's), and risk and time preferences. Furthermore, we test whether the program impacted personality characteristics (conscientiousness, locus of control, perseverance, extraversion and self-sufficiency) and mental health (self-esteem, hope/optimism, and mental distress.).

As mentioned previously, the survey measures are cross-sectional. The regression specification is thus the same as for the survey outcomes in the previous section: we regress the outcome on the binary treatment variable and include factory fixed effects and retention weights from the attendance data matched by survey date. Table 4 presents estimates of the impact of P.A.C.E. treatment on four categories of survey outcomes. We discuss results within each category in turn to lay out our reasoning.

The first category is meant to evaluate whether P.A.C.E. treatment changes women's financial behaviors and attitudes. The results from Panel A indicate that there is a positive impact on saving for own and children's education, and the impacts are quite large (about 30% of the control group mean). This result is consistent with a major theme in the training. The concept of personal goal setting, prioritizing actions in service of these goals, and mapping workplace motivation to the pursuit of these goals was at the center of one of the longest modules (Time and Stress Management, one of the three core modules); in which saving for a child's education was explored as a specific example. Savings for other purposes show no significant impacts. We construct survey-based measures of risk-aversion and patience (with higher scores corresponding to higher levels of those variables). The estimates suggest

[25] If we restrict the sample to workers to did not request technical training, we find nearly identical treatment effects for efficiency, which suggests that technical training does not drive the treatment effect on productivity. These results are available on request.

Table 4: 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	0.0804**	-0.0465	0.166*	-0.0984	0.0637*
	(0.0313)	(0.0334)	(0.0876)	(0.0935)	(0.0351)
Control Group Mean of Dependent Variable	0.265	0.272	-0.052	0.019	0.628
Control Group Standard Deviation of Dependent Variable	0.440	0.448	0.932	0.952	0.485
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	0.0248*	0.0226**	0.0119	-0.0257	-0.0270
	(0.0141)	(0.00941)	(0.0310)	(0.0352)	(0.0303)
Control Group Mean of Dependent Variable	0.039	0.006	0.120	0.142	0.152
Control Group Standard Deviation of Dependent Variable	0.192	0.080	0.322	0.347	0.357
Panel C: Personality	Conscientiousness	Locus of Control	Perserverance	Extraversion	Self-Sufficiency
P.A.C.E. Treatment	0.0210	0.0307	-0.123	0.164**	0.0445
	(0.0732)	(0.0770)	(0.0774)	(0.0702)	(0.0877)
Control Group Mean of Dependent Variable	-0.047	-0.040	0.020	-0.071	-0.063
Control Group Standard Deviation of Dependent Variable	0.929	0.951	0.982	0.969	1.021
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	-0.172	-0.0621	-0.0422	0.0456	0.0885***
	(0.106)	(0.0819)	(0.0389)	(0.165)	(0.0280)
Control Group Mean of Dependent Variable	0.048	0.015	0.094	23.427	0.117
Control Group Standard Deviation of Dependent Variable	1.013	0.966	0.294	2.077	0.319
Fixed Effects	Unit, Education, Age, Tenure				
Weighted	Inverse Predicted Probability from Probit of Retention on Treatments X Baseline Characteristics				
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 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).

that treatment increases risk aversion as well as participation in insurance or informal risk-sharing mechanisms (about 10% of the control group mean).

The second category, availing oneself of government and employer-sponsored entitlement programs, is meant to evaluate changes in the effectiveness of information acquisition, another important soft skill. The results in Panel B show that treated workers are substantially more likely to avail themselves of government pension and government subsidized healthcare programs. The magnitude of these impacts is quite large relative to control group means, which are around 0.04 for accessing government pensions, and about 0.006 for government subsidized healthcare. Impacts on other government subsidies and firm entitlements are negligible.

The third category is meant to assess differences in key personality traits often cited as productive non-cognitive skills, namely conscientiousness, internal locus of control, perseverance, extraversion, and self-sufficiency. In general, the impact estimates (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. This result is consistent with both the theme of assertive and effective communication emphasized in one of the three core modules, Communication, as well as the practice of role-playing and participation in group activities emphasized throughout the training.

The final category of survey outcomes is meant to assess impacts on psychological wellbeing, as well as the extent to which future aspirations are affected by treatment. The results reported in Panel D show 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. This is consistent with the result on saving for education presented 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.[26] While we do not find statistically significant effects on the self-esteem measure, the results on self-reported comparisons relative to co-workers indicate that P.A.C.E. training increased self-regard with respect to workplace performance in particular relative to peers.[27]

Panels C and D present an admittedly small set of personality measures and non-cognitive skills. 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. To supplement this analysis, we fielded a subsequent survey of additional soft skills and personality measures. Unfortunately, the original sample of workers had mostly left the factory by the time we fielded this second survey; accordingly, we conducted the second survey on a separate propensity score-matched sample of trained and untrained workers. Though this additional analysis does not draw on the randomized experiment, we present results from the propensity score matched treatment effects estimation as supporting evidence below in section 5.5.

[26] Note we do not find an effect of treatment on our measure of time preferences, despite finding effects on actions consistent with forward looking behavior like savings. Accordingly, we interpret the result in Panel A as consistent with this evidence of increased aspirations rather than a change in time preferences.

[27] In a subsequent survey on a supplemental non-experimental sample, we recollected several measures including self-esteem and conscientiousness and appear to have measured these dimensions more precisely.

5.2 Pre- and Post-Module Assessments

Additional evidence on the direct impacts of P.A.C.E. on stocks of soft skills comes from pre- and post-module assessments built into the program. These assessments were designed to test the specific value added from each core program module. They 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. Figure A3 in the appendix shows the pre-module assessments for each core P.A.C.E module. Figure A4 in the appendix shows the percent change between (identical) assessments taken pre- and post-module for each core P.A.C.E. module. Taken together, the results show that P.A.C.E. participants had low baseline stocks of soft skills and improved their stocks of these skills dramatically through the training. The changes shown in Figure A4 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.[28]

5.3 Heterogeneity in Productivity Impacts

We next investigate the degree to which productivity gains from training are heterogenous by baseline stock of leadership and technical skills as well as traditional schooling attainment. We do so by estimating a modified equation 1 in which we add terms for interactions between a dummy for a high level of baseline skill and the treatment by time period terms, as well as the main effect of the high baseline skill dummy. We include three different dimensions of baseline skill, alternately: leadership, traditional schooling, and technical skills.

For leadership skill, we asked factory HR representatives to rank participants in the training lottery (before treatment assignment) into 4 levels of baseline leadership skill (defined broadly as confidence and ability to effectively communicate with and motivate co-workers). We create a dummy for being above median in this ranking (i.e., receiving one of the top 2 of 4 possible ranks). For traditional schooling, we use a dummy for educational attainment above the primary level; and for technical skill, we use a dummy for being a machine operator as compared to an unskilled helper.

The results, presented in Table 5, show that productivity gains from training were strongest among those assessed by factory HR representatives as having low baseline stock of leadership skills. We interpret this evidence as further support for our assertion that increases in non-cognitive skills that contribute to leadership (e.g., communication, motivation and aspirations, and confidence and self-regard) were a primary mediating mechanism for the gains in productivity estimated above.

This result also indicates 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

[28]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. With these caveats in mind, these results are nevertheless consistent with our hypothesis that P.A.C.E. acted on workplace outcomes by increasing the stock of soft skills.

Table 5: Heterogeneous Impacts of P.A.C.E. Treatment on Productivity

	(1)	(2)	(3)
	Efficiency Produced/Target	Efficiency Produced/Target	Efficiency Produced/Target
After X P.A.C.E. Treatment	0.0622*	0.123***	0.00817
	(0.0343)	(0.0468)	(0.0359)
During X P.A.C.E. Treatment	0.0145	0.0478**	-0.00438
	(0.0165)	(0.0214)	(0.0189)
Baseline Leader X After X P.A.C.E. Treatment	-0.0430**		
	(0.0196)		
Baseline Leader X During X P.A.C.E. Treatment	-0.00708		
	(0.0116)		
Above Primary Education X After X P.A.C.E. Treatment		-0.0759*	
		(0.0402)	
Above Primary Education X During X P.A.C.E. Treatment		-0.0298*	
		(0.0179)	
Any Technical Skill X After X P.A.C.E. Treatment			0.0495*
			(0.0259)
Any Technical Skill X During X P.A.C.E. Treatment			0.0282*
			(0.0156)
Additional Controls	Days on Same Line-Garment, Total Order Size		
Fixed Effects	Unit X Month X Year, Worker X Garment		
Weights	Inverse Predicted Probability from Probit of Working on Treatments X Mo-Yr X Baseline		
Observations	258,285	290,763	290,763
Control Mean of Dependent Variable	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 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.

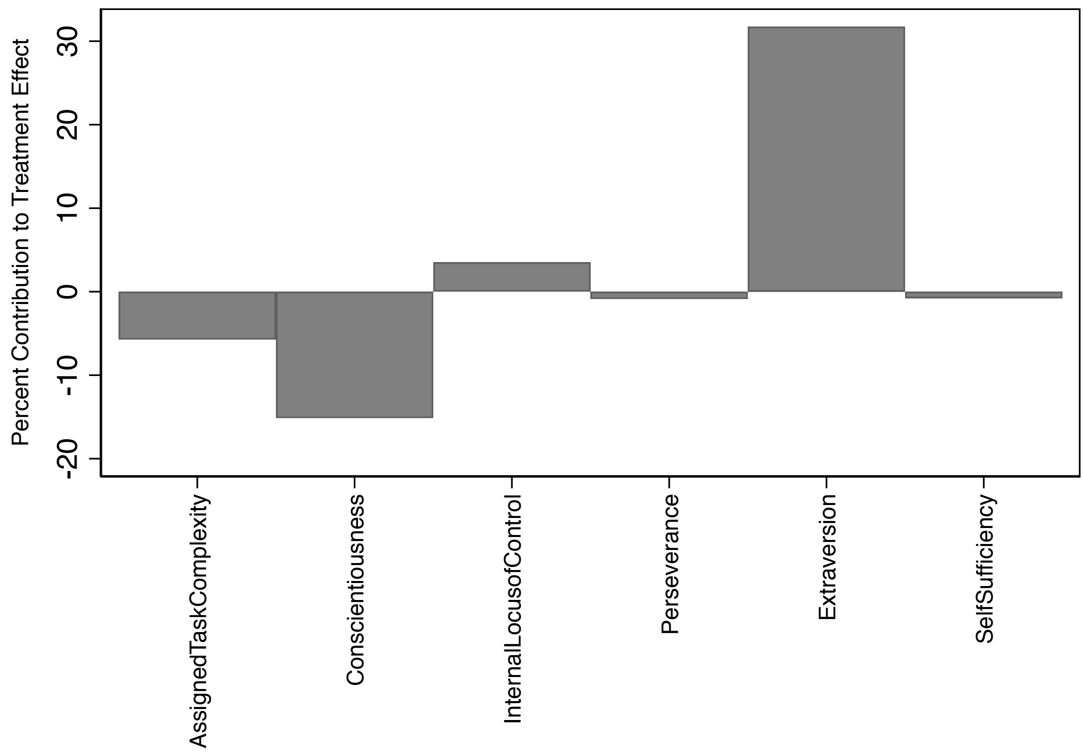
some foundational stock of skills upon which to build. Structural estimates of dynamic human capital accumulation models suggest dynamic complementarities in the productive value of non-cognitive skills, such that older children with low stocks of these skills benefit less than those who accumulated 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 result demonstrates that the productivity gains from soft skills training are actually largest among those with deficiencies at baseline. Relatedly, we find that productivity gains were also significantly larger for workers with less traditional education. This result is consistent with the notion that these productive soft skills are potentially imparted to some degree in traditional schooling (above the primary level) and so are lacking in those with less educational attainment.

Finally, we also find that productivity gains were concentrated among workers with greater technical skill. This result is consistent with the notion that machine operators require skills for communication, planning, and problem-solving to coordinate and maximize productivity more so than do non-technical helpers. We interpret this as evidence suggesting that in factory settings like ours, soft skills complement technical skills in determining productivity of workers. That is, the value of soft skills like communication, team work, planning and motivation are most productive when they are used to coordinate progress between technical operators towards a common goal.

5.4 Mediation Analysis

Figure 2: Percentage Contribution of Mediators to Treatment Impact on Productivity



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 4 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).

The results presented in Figure 2 show that 30% of the treatment effect on productivity is attributable to the large estimated gains in extraversion. The other personality dimensions contribute little to the productivity gains, as expected given the small and imprecisely estimated treatment effects on these dimensions. We also calculate the contribution of impacts on task complexity to the productivity gains. As expected, the contribution is negative, indicating that, in the absence of gains in soft skills that enhance productivity, the assignment to more complex tasks leads to reduced productivity. This negative contribution is small in magnitude, particularly relative to the large contribution of extraversion.[29] Note, however, that more than half the gains in productivity are unexplained by the measures we have in the survey on the experimental sample. Accordingly, we supplement these measures with a broader skills survey on an additional sample.

[29]Note that conscientiousness also appears to contribute negatively, but was imprecisely estimated in Panel C of Table 5 and as such we do not interpret this result strongly.

5.5 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.

Figure 3: Propensity Score Matched Treatment Impacts on Survey Outcomes

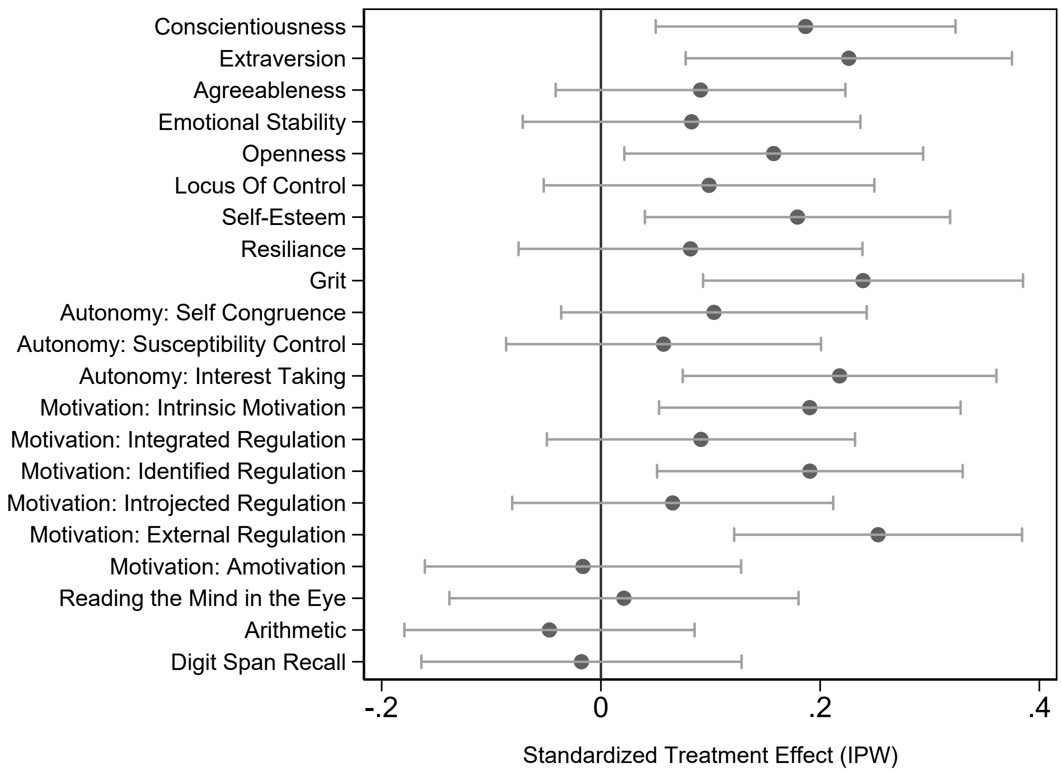


Figure 3 presents standardized treatment effects on 18 non-cognitive skills and personality traits as well as a measure of social cognition (i.e., reading the mind in the eye) and two other measures of cognitive skills (arithmetic and digit span recall) as placebos.[30] We find large and significant (roughly .2

[30]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 4 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

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 Table 4 (.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 3 indicate that these efforts indeed improved the measurement of these dimensions.

Each of these dimensions is consistent with themes and topics emphasized throughout the core modules of the training.[31] 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 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

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.

[31] We present a detailed description of the topics covered in and time devoted to each module in Section A.2, as well as evidence of the changes in knowledge of these topics as measured by pre-post testing.

corrections for multiple hypothesis testing in section C.3 of the appendix. 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.

5.6 Treatment Spillovers

Finally, we discuss treatment spillovers. Recall that the experiment was 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 all of the specifications mentioned above, 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).[32]

Results, presented in Table B3 in the appendix, show large spillover impacts on cumulative person days accrued to the firm, and imprecisely estimated but positive effects on efficiency more than two-thirds the size of the direct effects on trained workers. The existence of spillovers provides greater justification for employer investment in soft skills training, as gains compound beyond those receiving direct investment. We see some evidence for spillovers on outcomes outside the workplace, but the results are largely imprecise.

5.7 Alternative Mechanisms

Having presented evidence on the salience of direct skilling as a result of the P.A.C.E. program, we now discuss several alternative interpretations of the results and any supporting evidence of each.

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, we find spillovers in treatment for the number of days worked by workers who were signed up for the program and were on the same production line as treatment workers, but did not receive the program. These would be difficult to explain if reciprocity were the main driving force behind workplace impacts, since non-participants should not be driven by this motive. Second, the time pattern of productivity impacts (i.e., small impacts during the program, and large impacts post-program completion) does not fit well with reciprocity as a primary mechanism, since we would expect the reciprocity motive to be strongest while the program is conducted, and to dissipate over time if pay does not rise commensurately with

[32]Note that probability weights, when necessary, are calculated exactly as they are in the treatment effect estimation, using spillover treatment indicators in place of direct P.A.C.E. training.

productivity, as in this case. 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).

Second, we evaluate the possibility that the results on productivity and task 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. We reason that sheepskin effects are unlikely to explain the majority of the program's impacts given the slow onset of increased productivity over time, rather than an increase near the program's end. Additionally, once again spillover impacts are inconsistent with a sheepskin effect mechanism. 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).

Third, it is possible that workers found the classes enjoyable and they improved workers' subjective wellbeing, which in turn made workers more productive. The fact that productivity impacts are small and insignificant during the program (when such enjoyment would be most salient), and large and significant only 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 4 show that levels of moderate psychological distress, which might reflect this subjective wellbeing to some degree, are not statistically different by treatment status.[33]

Finally, we consider the idea that increased social capital drives the results on workplace impacts. The argument is that it is possible that the training sessions improved the ability of workers to create social ties, which could generate higher productivity on their production lines if it increased the extent or intensity of social connectivity on the line. 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, 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. Note that this feature does not entirely preclude social ties from being impacted by the program; it simply lowers the likelihood that this channel contributed significantly to impacts on workplace outcomes like productivity.[34]

6 Return on Investment

To quantify the profit return to the firm, we combine our treatment effect estimates on retention (person-days) and productivity 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 additional person days and incremental productivity from treated workers using estimates from the randomized evaluation. Cost involve fixed and variable programmatic costs, lost productivity due to training, and wage increases (we do not report wage as a separate category of cost in Table 6 because these impacts were essentially negligible).[35] We omit spillover impacts from the calcula-

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

[34]Indeed the observed treatment effects on extraversion may have induced better communication and thus greater social ties – something that would be part of our soft skills-based interpretation of the program's impacts.

[35]In addition, we implicitly assume in calculating lost productivity due to reduced person days that the rate of hiring or worker replacement is common across treatment and control lines such that differential attrition produces truly lost person

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)	-$57,091.68
P.A.C.E. Training Variable Cost (Lost Garments from Lost Man Hours)	-$38,314.88
Total Cost *(All numbers in present value)*	-$95,406.56
1 Year After Program Announcement	
Additional Person Days	$51,804.37
Additional Productivity	$112,785.00
Net Present Value of Subtotal	$164,589.30
Net Rate of Return	**73%**
20 Mos After Program Announcement	
Additional Person Days (End of Observation)	$68,389.79
Additional Productivity (Garments per 8 hr day)	$272,767.00
Net Present Value of Subtotal	$341,156.80
Net Rate of Return	**258%**
Assumptions	
Additional Garments per Additional Man Day	8.2
Revenue per Additional Garment	$7.00
Labor Contribution to Cost ("Cut to Make")	25%
Profit Margin on Additional Revenue from Additional Productivity	18.75%
Profit Margin on Additional Revenue from Additional Man Day	5%
Interest Rate	10%
INR per 1 USD	58

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. Additional garments per additional man day 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 480 minutes in a work day. All additional productivity and man days coefficients are taken from the montly impacts estimated in the main results 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 75% 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.

tions that follow to produce conservative estimates, given that spillover effects on productivity are not statistically significant.

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 costs are from lost production hours, and the marginal increase in wages for treated workers. For the 1087 treated workers, total program costs are approximately $95,000, about $57,000 of which are overhead costs, and the remainder variable costs.

Details on profit margins on additional revenue both from an additional person day and additional

days. This is largely true as hiring is centralized for each factory. Accordingly, firm management reported to us that it is impossible for the rate of recruitment, hiring, and training to respond to differential turnover across lines within factory.

productivity, as well as additional revenue per garment were obtained from the firm. The benefits of the program are generated by the higher number of cumulative person days accrued to the firm and by higher worker productivity. At the end of the program period, the NPV of these benefits is just over $164,000, about $52,000 of which is the result of additional person days gained during the program and the rest due to productivity gains. At the end of our tracking period (8 months after program completion), total benefits are substantially higher, more than $341,000. In the post-program period, returns via productivity gains dominate, accounting for more than 70% of the total benefits.

The net rate of return at the end of the program period is thus 73% (i.e., at program end, costs had been entirely recouped by the firm, plus 73 percent additional returns). Twenty months after program completion, flow benefits mostly from post-program productivity impacts help generate a net rate of return of 258%.

7 Conclusion

In this paper we study the labor market impacts of soft skills. 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 workplace 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, 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.

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). 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 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., Molina, T., and Nyshadham, A. (2019). Expectations, wage hikes, and worker voice: Evidence from a field experiment. Technical report, National Bureau of Economic Research.

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*.

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.

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.

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*.

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.

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.

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.

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).

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.

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
December 2014	Salary Data Collection Ends
February 2015	Attendance and Production Data Collection Ends

Figure A2: Data Type and Availability

Productivity	Daily (June 2013-February 2015)
Salary	Monthly (January 2013-December 2014)
Survey Outcomes	Cross-sectional (June 2014)
Retention	Daily from Productivity Data, Monthly from Salary Data

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 roughly 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 A: 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). In separate sessions, 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 stress-management training, the first session focused on identifying stress, as well as its ubiquity. There was an exercise and 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

A.2.2 Pre-Post Testing Results

Figure A3: Avg Pre Score

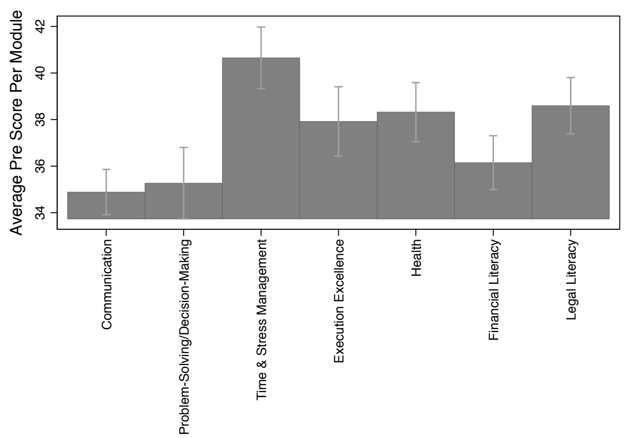
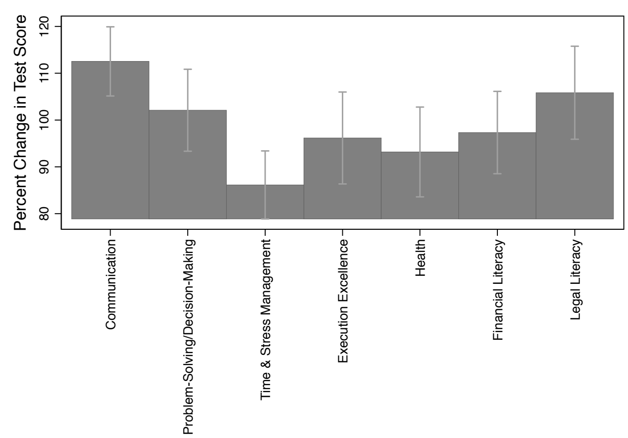


Figure A4: Post - Pre % Change in Score



Figures A3 and A4 depict average pre-training test score (A3) and normalized (percent change, A4) difference between post- and pre-training test scores administered for all core P.A.C.E. modules. Raw scores for each assessment are out of 100. These assessments were not given to control workers and accordingly cannot be analyzed in the preferred specification. Session attendance rates was high but varied slightly by training module, ranging between 94 and 99% with an average of roughly 96%.

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.

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	0.00620	9.250	0.00545	-0.00979	-0.0190
	(0.0256)	(8.683)	(0.00833)	(0.00721)	(0.0165)
During X P.A.C.E. Treatment	0.0264	5.360	0.00749	-0.00712	-0.00307
	(0.0215)	(3.258)	(0.00591)	(0.00581)	(0.0133)
Announced X P.A.C.E.. Treatment	0.00416	0.501	0.00998	-0.0109	0.00242
	(0.0136)	(1.271)	(0.0106)	(0.0106)	(0.00972)
Fixed Effects	Unit X Month X Year, Worker				
Weights	None		Inverse Predicted Probability from Probit of Retention on Treatments X Mo-Yr X Baseline Characteristics		
Observations	1,433,981	1,270,871	736,439	736,439	563,624
Control Mean of Dependent Variable	0.63	0.52	0.893	0.097	0.367

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.

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). We conduct additional tests to verify the lack of an effect on the composition of retained workers over the observation period and discuss implications of these results for the analysis of conditionally observed outcomes used later in the analysis.

B.2 Line-Level Productivity and Task Complexity Results

As a further test of robustness of our main results, we present regression results using daily productivity and task complexity at the production-line level instead of the individual-level.[36] 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 both efficiency and SAM are statistically significant at the 10% level after treatment. The magnitude of the line-level treatment effect for efficiency is about 40% of the direct treatment effect, and for SAM is about 70% 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.

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

Table B2: Impact of P.A.C.E. Treatment on Line-Level Daily Productivity and Task 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 Feb 2015)		*Line-level (including all workers on line)*	
After X P.A.C.E. Treatment	0.150**	0.0798***	0.0431*	0.0289*
	(0.0654)	(0.0255)	(0.0251)	(0.0171)
During X P.A.C.E. Treatment	0.0693*	0.0642***	0.0130	0.0174
	(0.0390)	(0.0208)	(0.0169)	(0.0134)
Additional Controls	Days on Same Line-Garment, Total Order Size	None	Days on Same Garment, Total Order Size	None
Fixed Effects	Unit X Month X Year, Worker X Garment	Unit X Month X Year, Worker	Unit X Month X Year, Line X Garment	Unit X Month X Year, Line
Weights	None			
Observations	130,187	130,187	81,258	81,258
Control Mean of Dependent Variable	0.527	0.588	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.

B.3 Treatment Spillovers

Results on treatment spillovers are presented in Table B3. Panel A presents the results for person days as well as productivity. There is a weakly statistically significant impact on the binary variable for working during the treatment announcement period, and a stronger result for cumulative person days during the treatment period - untrained workers who work with treated workers work for about 8 more days during program months relative to control workers. Productivity impacts are positive, about 70% as large as the direct treatment effects, but are not statistically significant. Panel B presents the results for career advancement variables. Similar to the effect on productivity, the spillover impacts on survey outcomes on requesting skill development training, receiving a production incentive or self-assessment relative to co-workers are not precisely measured, but again have coefficients of the same sign as the main treatment impacts. The worker self-assessment relative to co-workers is positive and statistically significant at the 10% level.

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

	(1)	(2)	(3)	(4)
Panel A: Working and Production	Working	Cumulative Person Days	Efficiency	
After X Spillover	-0.0155	8.652	0.0714	
	(0.0206)	(9.332)	(0.0571)	
During X Spillover	0.0252	8.023**	0.00591	
	(0.0209)	(3.841)	(0.0319)	
Announced X Spillover	0.0317*	2.151		
	(0.0172)	(1.372)		
Fixed Effects	Unit X Month X Year, Worker		Unit X Month X Year, Worker X Garment	
Weights	None		Inverse Predicted Probability from Probit of Working on Treatments X Mo-Yr X Baseline Characteristics	
Observations	1,102,880	673,407	241,322	
Control Mean of Dependent Variable	0.519	0.382	0.548	
Panel B: Career Advancement	Skill Development Training	Production Award or Incentive	Skill Peer Self-Assessment	Co-Worker Self-Assessment
Spillover	0.0254	0.0204	0.113	0.140*
	(0.0608)	(0.0243)	(0.0687)	(0.0769)
Fixed Effects	None			
Weights	Inverse Predicted Probability from Probit of Retention on Treatments X Mo-Yr X Baseline Characteristics			
Observations	527	527	527	527
Control Mean of Dependent Variable	0.244	0.031	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 working dummies and cumulative man days are defined for every worker date observation in the data and therfore regressions do not require any weighting. Observations in attendance and advancement regressions 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. 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).

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)	
	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	0.075	0.918	0.074	-0.003	0.56
1(High Education)	0.573	0.517	0.580	0.506	-0.007	0.84
Years of Tenure	1.760	2.115	1.569	1.738	0.191	0.17
Age	30.006	12.341	28.788	10.748	1.218	0.14
1(Speaks Kannada)	0.721	1.045	0.691	0.799	0.030	0.65
High Skill Grade	0.581	0.696	0.640	0.598	-0.059	0.20
log(Salary) (May 2013)	8.770	0.160	8.756	0.140	0.014	0.19
Efficiency (Announcement Month)	0.593	0.418	0.562	0.312	0.031	0.27
SAM (Announcement Month)	0.641	0.531	0.630	0.412	0.011	0.75
	Last Month of Data Collection (February 2015)					
	Control		Treated		Difference	
P.A.C.E. Treatment	**Control Workers in Control Lines**		**Treated Workers in Treatment Lines**			
Number of workers	263		373			
	Mean	SD	Mean	SD	Mean Difference	p value
Attendance Rate (Jan-May 2013)	0.914	0.071	0.918	0.071	-0.004	0.49
1(High Education)	0.540	0.559	0.552	0.529	-0.012	0.78
Years of Tenure	1.694	1.860	1.652	1.564	0.042	0.76
Age	30.156	8.634	29.402	8.594	0.754	0.28
1(Speaks Kannada)	0.738	0.896	0.713	0.704	0.025	0.71
High Skill Grade	0.570	0.645	0.614	0.571	-0.044	0.38
log(Salary) (May 2013)	8.775	0.170	8.763	0.148	0.013	0.34
Efficiency (Announcement Month)	0.598	0.362	0.565	0.276	0.033	0.23
SAM (Announcement Month)	0.653	0.493	0.631	0.394	0.022	0.57

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

C.2 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-

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 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
	(0.0698)	(0.0759)	(0.0674)	(0.0786)	(0.0695)	(0.0769)	(0.0710)	(0.0800)	(0.0745)	(0.0742)	(0.0726)	(0.0812)	(0.0675)	(0.0745)
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	0.0850	0.217*	0.133	0.0756	0.0457	-0.0851	-0.0204
	(0.0753)	(0.0820)	(0.0761)	(0.0887)	(0.0811)	(0.0885)	(0.0841)	(0.0888)	(0.0859)	(0.0828)	(0.0765)	(0.0945)	(0.0753)	(0.0751)
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
	(0.0701)	(0.0759)	(0.0675)	(0.0786)	(0.0699)	(0.0774)	(0.0711)	(0.0801)	(0.0746)	(0.0743)	(0.0724)	(0.0812)	(0.0673)	(0.0737)
Fixed Effects							Unit, Education, Age, Tenure							
Weighted							None							
Observations							675							

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 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).

values in increasing order ($p_1 < ...p_M$), and rejecting c hypotheses such that c is the largest w where $pw < (qw/M)$.[37]

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.[38] 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.

C.3 Alternative Estimates from Supplemental Non-experimental Sample

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

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

Figure C1: Nearest Neighbor Matched Treatment Impacts on Survey Outcomes

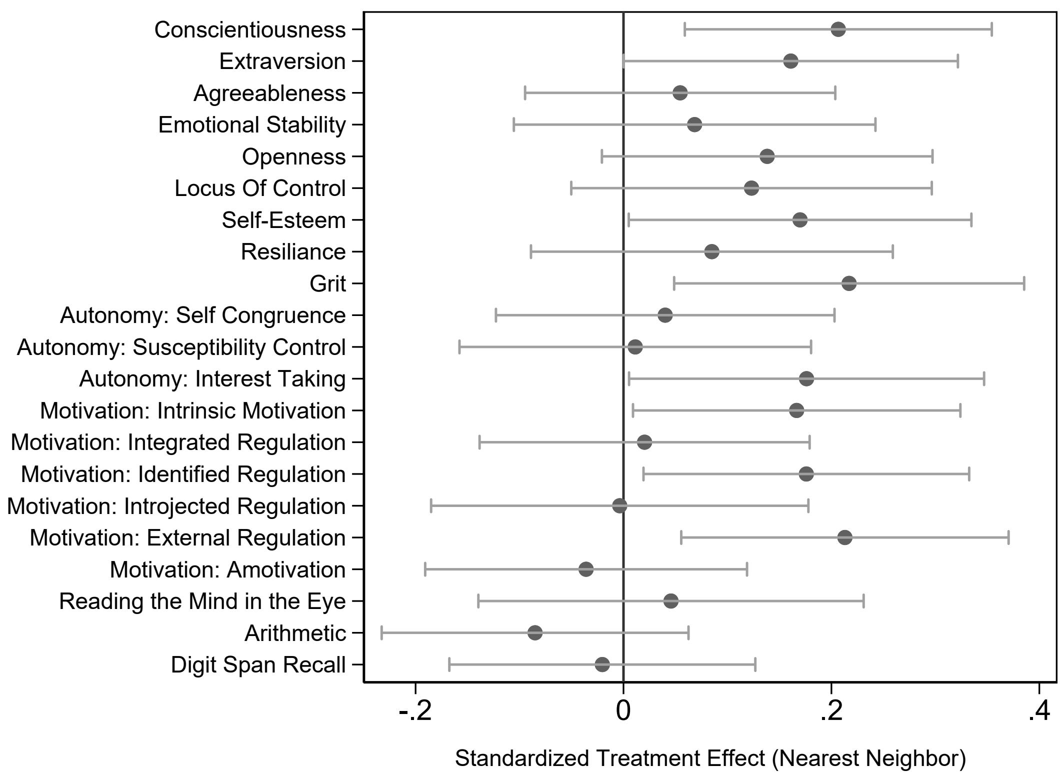


Figure C2: Unmatched Treatment Impacts on Survey Outcomes