Multiskilling in Organizations

Ratchanon Chotiputsilp, Mitchell Hoffman, Taeho Kim, and Shing-Yi Wang*

September 2026

Abstract

How should new workers be trained? A standard approach is to train workers for one job, whereas an alternative, called multiskilling, is to train workers for multiple jobs. We study multiskilling through an RCT with a large Southeast Asian garment manufacturer, which randomly assigned half of 630 new and lower-tenure workers to multiskilling. Multiskilling reduces absenteeism by 14 percent and post-training turnover by 10 percentage points. Production lines receiving treated workers produce roughly 10 percent more output and achieve 15 percent higher efficiency. These benefits do not appear to operate primarily through workers being reassigned based on newly revealed comparative advantage or through greater organizational agility. Although treated workers develop more favorable views of jobs they experience during rotation, few ultimately switch jobs, and treated workers do not substitute across jobs more frequently. Instead, the effects are largest, though imprecisely estimated, among workers who initially rate their own performance poorly. These same workers become significantly more likely to report that promotion opportunities are not limited and that rewards are sufficient, consistent with multiskilling changing workers' beliefs about their prospects within the firm. The measured benefits of multiskilling exceed its direct training costs by roughly six weeks of worker pay per treated worker. Qualitative evidence suggests that organizational frictions, including production constraints and managers' reluctance to release valuable workers, limit the firm's ability to act on newly revealed opportunities for worker reallocation. Overall, our findings show that broadening workers' skills can improve both worker and firm outcomes, even when firms make limited use of the additional flexibility and information that multiskilling creates.

Keywords: Multiskilling, management practices, garment industry

JEL Codes: M50

Preliminary — work in progress. Please do not cite without permission.

*We thank Peter Kuhn and Kathryn Shaw for helpful comments. This work was supported by the Connaught Fund. We thank Cung Hoang, Adi Jahic, Rosemary Zhang for excellent research assistance and Imran Looji and Pattawan Chanwerawong for superb field management. We thank the garment firm and its top management for their enthusiastic collaboration. The RCT is pre-registered with a pre-analysis plan on the AEA RCT Registry under AEARCTR-0012136. The IRB protocol for this project was approved by University of Toronto (44754), Erasmus University (ETH2223-0548), and University of Pennsylvania (853902). All opinions expressed in this manuscript are those of the authors.

Workers often learn new skills, not only through school or government programs, but through on-the-job training within firms. This is especially true for low-wage workers (Acemoglu and Pischke, 1998). Despite longstanding theoretical interest in understanding firm training (Pigou, 1912; Becker, 1964), relatively little is known empirically about training within firms, and the relative advantages of different forms of training.

A standard approach to training in many firms, which we call *monoskilling*, is to train workers for a single job: a firm trains a new hire to do the one job that it needs. An alternative approach, *multiskilling*, trains workers for multiple jobs before assigning them permanently to one. Several firms are known for multiskilling and for purportedly reaping benefits from it, including the Spanish grocery chain Mercadona (Ton and Harrow, 2010), Costco (Ton, 2014), and Toyota (Adler et al., 1999), but there is very little rigorous evidence on the returns to multiskilling.

There are (at least) three main potential benefits from multiskilling, and our design lets us examine each in turn. First, multiskilling may help workers and firms discover better worker-job matches. Workers are better suited to some jobs than others, but their comparative advantage and preferences over jobs may not be known until they have some training and experience in the job. Multiskilling can generate this information and allow workers to be reallocated toward jobs that increase their productivity or job satisfaction. Second, multiskilling may increase organizational agility: workers with multiple skills can substitute for one another in response to absences and quits. Third, multiskilling may benefit workers even when neither permanent reallocation nor temporary substitution occurs. Workers may value the additional human capital itself, learn more about opportunities available within the firm, or respond to the firm's willingness to invest in their development.[1] These effects may be particularly important in lower-wage environments, where workers often receive relatively little employer-sponsored training; training workers for multiple jobs, which is costly to the firm in the time of both workers and managers, may be a way of credibly communicating investment in the workers.

To shed light on these different costs and benefits, we conduct a randomized controlled trial (RCT) with a large garment manufacturer operating across multiple countries in Southeast Asia. The firm is a major producer of sports jerseys, athletic wear, and other clothing. Driven by turnover, it hires hundreds of new workers per year for different jobs, including sewing, embroidery, heat transfer, and packing. Prior to the RCT, the firm followed a traditional monoskilling model, training new workers in a single job before they began it. We randomly assigned new and recent hires to monoskilling ("Control") or multiskilling ("Treatment"). In addition to one standard job, treatment workers were trained in one or two additional jobs for approximately one week each. At the end of

[1]We will refer to the benefits as "psychological benefits" as short-hand, even though the benefits may be purely rational. For example, in many U.S. universities, students take coursework in many different fields, even those quite disconnected from their major or later career, with the idea that they will broadly benefit from exposure to other knowledge.

the rotation, workers ranked the jobs d and managers evaluated workers who had rotated through their teams. We combined these rankings using a worker-proposing deferred acceptance algorithm to generate recommended permanent job assignments.

The setting we examine is a natural one for studying the consequences of multiskilling. In many low-wage labor markets in developing countries, formal credentials and prior work experience provide firms with limited information about workers' comparative advantage across jobs. Prior research also documents high rates of turnover and absenteeism in jobs in these settings (Banerjee and Duflo, 2006; Donovan et al., 2023). These labor-market conditions create an important tension. On the one hand, turnover and absenteeism increase the value of having workers who can perform multiple tasks and substitute for one another. On the other hand, they shorten the expected period over which firms can recoup their investments in training, potentially making firms less willing to provide it. Our experiment sheds light on how firms navigate this tradeoff. It also allows us to examine the reverse relationship: whether multiskilling—by improving worker-job matches, expanding workers' skills, or signaling the firm's investment in them—reduces turnover and absenteeism. Low-wage labor markets in developing countries are also a particularly useful setting for studying potential indirect benefits of multiskilling, because workers may otherwise perceive limited investment in their development or respect from their employers.

We implement the RCT with 630 new and recent hires across several factory locations in Southeast Asia, aligned with hiring batchesthat occurred between August 2023 and May 2025. Within each batch, we randomly assigned workers to multiskilling treatment (312 workers) or to the firm's standard single-job training (318 workers). Among treatment workers, assignment to specific rotation jobs was further randomized. We combine four sources of data: baseline surveys including an aptitude test administered before the randomization, midline manager evaluations collected at the end of each job rotation cycle, a rolling endline survey of all participants fielded roughly four months after each batch began, and firm administrative records covering personnel and resignations, timesheets, team-level production, and workplace violations.

We find that multiskilling leads to sizable improvements in worker outcomes. Treated workers are absent 14% fewer days per month than control workers. This effect is concentrated in reductions in sick leave. Furthermore, treated workers are 10 percentage points less likely to resign after rotation has ended.

These effects are somewhat striking because, while the experiment clearly generates learning about alternative jobs, relatively little permanent reassignment ultimately occurs. The matching algorithm using the preferences of treated workers and the ratings of managers recommends a department switch for about one in five treated workers, but only about one third of these recommendations are ultimately implemented.[2] Treated workers do not report significantly greater

[2]In other words, only about 7% of treated workers move into an algorithm-recommended job.

preference for the jobs they actually hold at endline relative to control workers, and managers do not evaluate treated workers as performing better in their endline jobs four months after the intervention. Thus, multiskilling generates information about alternative jobs, but produces relatively little improvement in realized worker-job match quality. These results suggest that improved permanent matching based on information learned through the rotation is unlikely to account for the bulk of the average treatment effects.

At the same time, treated workers become substantially more favorable toward the jobs on their rotation menu: by endline they are 7.9 percentage points more likely to rate one of these jobs as "strongly preferred," relative to a base rate of 15.9%, roughly a 50% increase. Thus, experiencing alternative jobs meaningfully changes workers' views of them.

Turning to the hypothesis about job agility, we find little evidence that formal substitution is an important channel in our setting. At endline, treated workers do not report spending more days substituting into other jobs than control workers. While substitution may be a central benefit of multiskilling in settings such as retail, where workers routinely move between tasks, it does not appear to account for the bulk of the effects we observe.

We next examine whether these worker-level gains translate into benefits for the firm. Because productivity in the garment industry is measured at the production-line level rather than for individual workers, we exploit variation generated by the experiment in the share of workers entering a line who were assigned to treatment. At the average treated-worker inflow among lines receiving treated workers, time worked increase by 10% and efficiency as measured by output per unit time increased by 14%. These results suggest that the individual-level improvements in worker outcomes do not come at the expense of production, but generate meaningful benefits for the firm as well.

This paper contributes to several literatures, the first being the economics literature on training within firms. The modern literature on firm training draws on the seminal contributions of Acemoglu and Pischke (1998, 1999a,b), who argue that market imperfections, such as monopsony power or imperfect information may play a key role in making training valuable for firms, for whom paying for training is difficult due to poaching concerns (Becker, 1964). Recent research measures the returns to training programs within individual firms (Bartel, 1995, 2000; De Grip and Sauermann, 2012; Adhvaryu et al., 2023; Espinosa and Stanton, 2026). To our knowledge, our paper is the first to measure the returns to multiskilling. In other fields outside of economics including management and psychology, a small literature studies multiskilling using lab experiments (Volpe et al., 1996; Marks et al., 2002), theoretical modeling (Molleman and Slomp, 1999; Slomp et al., 2005), and qualitative case studies, but their focus is not on the returns of multiskilling. To our knowledge, our paper provides the first experimental analysis of the returns of multiskilling to workers and firms in a real firm.

Second, we contribute to the literature on management practices and why firms fail to adopt seemingly profitable ones. A growing literature shows that management practices matter for productivity (Ichniowski et al., 1997; Bloom and Van Reenen, 2007), including via RCTs (Bloom et al., 2013). There is less work such as ours that experimentally measures the return to a particular management practice. A related literature asks why firms do not adopt such practices despite their apparent benefits, pointing to misaligned incentives between owners and employees (Atkin et al., 2017) and to informational and learning barriers and features of the business environment (Bloom et al., 2019). Multiskilling expands workers' capabilities and generates information about alternative matches, but taking advantage of that information requires the organization to reallocate workers across teams with competing production needs — and we show that the implementation of the reallocation is itself a source of friction. Specifically, we contribute to work in organizational and personnel economics on managing conflicts over labor allocation. Recent work argues that managers face incentives to "hoard labor," for example by discouraging qualified employees from moving to other positions when doing so would cause the manager to lose a productive worker from their own team (Friebel and Raith, forthcoming; Haegele, 2026; He and Waldman, 2026).[3] We add to this literature by providing a case where conflicts between managers appear to have contributed to a firm not adopting the practice longer-run, despite sizable measured benefits.

The rest of the paper is organized as follows. Section 1 describes the RCT and conceptual framework. Section 2 describes the data. Section 3 presents the results, and Section 4 evaluates the costs and benefits of multiskilling. Section 5 discusses the implications of our findings, and Section 6 concludes.

1 Research Design

1.1 Methodological Framework and Identification Strategy

We evaluate the multi-skilling intervention using an RCT. Workers in eligible hiring batches were randomly assigned to treatment or control, allowing us to estimate the causal effect of the opportunity to train in and experience multiple production jobs. Random assignment addresses selection concerns that would arise from studying participation in an existing firm program: workers selected for multi-skilling could otherwise differ systematically in their ability, performance, or other characteristics, and managers might choose which workers to train or which alternative jobs to expose them to based on those characteristics. Within the treatment group, assignment to available rotation positions was also randomized, limiting managers' discretion over which alternative jobs individ-

[3] A related theoretical literature studies how organizations manage disagreements through authority and integration, or through communication and delegation (Van den Steen, 2010; Hart and Holmstrom, 2010; Dessein, 2002).

ual workers experienced. We describe the intervention and randomization procedure in greater detail below.

1.2 Intervention

The intervention was a multi-skilling program that exposed workers to alternative production jobs and, following this exposure, provided the possibility of reassignment based on worker and manager preferences. Most participants were newly hired workers entering the firm's production workforce. Under the firm's standard allocation process, new workers are assigned to jobs primarily based on current vacancies and their background and prior experience. The process therefore emphasizes meeting immediate staffing needs and relies on relatively limited information about workers' preferences for different jobs or their performance across alternative production tasks.

After initially working in the job to which they were assigned through this standard process, workers assigned to treatment were randomly assigned to one or two additional jobs, in which they worked for one to two weeks each.[4] Control workers continued working in the jobs to which they were initially assigned by the firm.

The firm's production process includes eight core jobs, which collectively account for approximately 90% of production workers and occupy the same level of the organizational hierarchy. Treated workers did not rotate through all eight jobs. In most factories, the set of jobs available within a rotation batch was determined by the jobs to which treated workers in that batch had initially been assigned. Treated workers were then randomly assigned to rotate through one or two of these jobs. Thus, the intervention allowed workers within a batch to experience alternative jobs and potentially exchange assignments while keeping the overall composition of positions in the batch approximately fixed.[5]

At the end of the rotation cycle, workers reported their preferences over the jobs they had experienced, while managers evaluated the performance of workers who had rotated through their teams. We used this information to generate recommended assignments through a worker-proposing deferred acceptance algorithm. Recommendations were determined jointly within each rotation batch rather than independently worker by worker. Jobs had limited capacity, so multiple workers could not necessarily be accommodated in a job even when they preferred it to their initial assign-

[4] If a treated worker was assigned one additional job rather than two, the worker performed that job for one to two weeks, depending on job availability.

[5] At one factory, accounting for approximately 37% of treated workers, implementation differed because participating workers were initially assigned to sewing and the firm had vacancies in other production jobs. These vacancies, together with sewing, constituted the positions available for rotation and subsequent matching. Workers at this factory retained their initial sewing assignment as an outside option, so the matching procedure recommended an alternative assignment only when the worker preferred it to sewing. In the remaining factories, available positions were determined primarily by the initial assignments of treated workers in the batch, so subsequent reassignment could require reshuffling positions across treated workers.

ment, and assigning a position to one worker could affect the positions available to others. The matching procedure provided a systematic way to incorporate workers' preferences and managers' evaluations while respecting these capacity constraints. The worker-proposing deferred acceptance algorithm produces a stable allocation and, among stable allocations, selects the worker-optimal assignment (Gale and Shapley, 1962). The mechanism is widely used in two-sided assignment settings; most prominently, an applicant-proposing version forms the basis of the National Resident Matching Program that assigns medical graduates to residency programs (Roth and Peranson, 1999).

On the worker side, jobs were ranked according to workers' reported preferences. In the deferred acceptance procedure, workers first proposed to their most-preferred job. Each job tentatively retained workers up to its available capacity according to managers' rankings and rejected the remaining applicants. Rejected workers then proposed to their next-most-preferred job, and jobs reconsidered their tentative matches together with new applicants. This process continued until no further proposals were made. When a worker ranked multiple jobs equally, their ordering was randomized. Managers' evaluations of workers' performance during the rotation determined jobs' rankings of workers.[6] The algorithm considered only jobs that were available within a worker's rotation batch and that the worker had experienced during the program.

Workers and managers were informed that their responses would be used to generate recommendations for treated workers' subsequent job assignments, making their responses consequential rather than purely hypothetical. The resulting assignments were presented to each factory as recommendations. Individual factories could therefore deviate from the algorithm's recommendations. Despite senior management's support for the program, as we report in Section 3, realized reassignment was limited.

1.3 Conceptual Framework

To fix ideas, consider two workers, 1 and 2, and two jobs, A (sewing) and B (embroidery), each requiring one worker. Worker i in job j produces output y_{ij} and enjoys non-wage utility a_{ij}; write $v_{ij} = y_{ij} + a_{ij}$ for the joint value of the pair. Both components are match-specific and are learned only once the worker has been trained in j and worked there (Jovanovic, 1979). Under *monoskilling*, the firm trains each worker for one job—say 1 for A and 2 for B, based on vacancies and background—at training cost τ per worker-job pair. Under *multiskilling*, the firm pays 2τ per worker to train both workers for both jobs, and can then assign either worker to either job. Multiskilling thus trains four worker-job pairs while only two are in use at any time. The question is what the extra τ per worker buys.

[6] Treated workers who did not rotate under a particular manager received a score of 3 on a scale from 1 to 5.

Reallocation (comparative advantage). With both workers trained for both jobs, the firm can keep the initial assignment, worth $S = v_{1A} + v_{2B}$, or swap, worth $S' = v_{1B} + v_{2A}$. Since it chooses the larger, the gain from breadth on this margin is $\mathbb{E}\big[(S' - S)^+\big]$. Writing $c_i \equiv v_{iB} - v_{iA}$ for worker i's comparative advantage in B, we have $S' - S = c_1 - c_2$, so the gain is

$$\mathbb{E}\big[(c_1 - c_2)^+\big] \;=\; \tfrac{1}{2}\,\mathbb{E}\big|c_1 - c_2\big| \tag{1}$$

when the c_i are independent draws from a common distribution. Two features are worth noting. First, only *comparative* advantage matters: if embroidery is simply harder for everyone, so that $c_1 = c_2$, the gain is zero, however large the absolute differences between jobs. What multiskilling buys is the ability to act on differences in the *ranking* of jobs across workers. Second, the gain is a measure of dispersion in comparative advantage, and is increasing in it; if c_i were normally distributed with standard deviation σ_c, it would equal $\sigma_c/\sqrt{\pi}$. Realizing this gain requires that the swap actually take place, so preferences and assessments can be revised without any change in assignments, in which case this channel contributes nothing.

Redeployment (agility). Even holding assignments fixed, breadth has value because staffing needs fluctuate. Suppose each worker is absent with probability p. Under monoskilling, an absence leaves that worker's job unstaffed, at cost L. Under multiskilling, the worker who is present can cover, recovering a share θ of what would otherwise be lost, so the gain on this margin is roughly $2p(1-p)\,\theta L$: increasing in absence and turnover risk, and in the cost of leaving a position unstaffed (Molleman and Slomp, 1999; Slomp et al., 2005). Because such coverage may be brief and informal, it need not register as a job change in workers' own reports, and may be more visible in line-level output.

Indirect benefits. Multiskilling may also benefit workers without changing either how they are assigned or how they are deployed. Workers may simply enjoy being trained in several jobs, whether for the variety, the skills themselves, or what those skills are worth to them later. In addition, training of this kind is costly and pays off mainly if workers can be trusted not to leave with their new skills or to misreport their preferences, so it may serve as a credible signal of the firm's respect for its workers as a group (Ellingsen and Johannesson, 2008, 2007), in contrast to practices that signal distrust (Falk and Kosfeld, 2006). Unlike the previous two channels, benefits of this kind accrue to treated workers whether or not their assignments ever change.

Adoption. Finally, the returns above accrue to the firm, while the costs fall largely on individual managers. Trainer and supervisor time is borne by a line during the rotation, and a manager asked to release an experienced worker so that a beneficial swap can occur bears that loss while capturing little of the gain (Friebel and Raith, forthcoming; Haegele, 2026). The benefits—retention, attendance, and coverage—are by contrast diffuse and delayed. A practice with this structure may

be implemented only partially, or not adopted at all, even when its benefits exceed its costs at the level of the firm.

1.4 Outcomes and Hypotheses

Our pre-analysis plan organized the study around six sets of outcomes: individual performance, turnover and absenteeism, worker satisfaction, preference discovery, temporary substitution, and workplace violations. The intervention allows us to study three implications of multi-skilling: rotation could reveal jobs that workers prefer or perform well in, while the matching procedure could use information from workers and managers to implement better assignments, workers may update their beliefs about jobs and the workplace even without changing assignments, and cross-training may give the firm greater flexibility to move workers temporarily across tasks.

For individual performance, our primary measure is the endline evaluation by each worker's immediate supervisor. The firm's production systems could not be reliably linked to individual workers across all core jobs. We therefore complement the individual analysis with production-line outcomes below.

Turnover, attendance, hours, pay, and workplace violations come from administrative records. Worker satisfaction, preferences, and temporary substitution come from the endline survey. The firm estimates that recruiting and initial general training cost approximately 300 USD per employee, more than one month of salary for an average worker, making turnover and attendance directly relevant to firm performance.

The pre-specified hypotheses were that treatment could improve performance through better matching; reduce absence and work-related quits through better fit and satisfaction, while potentially increasing quits if broader training improved outside options; increase job satisfaction; induce greater updating of job preferences; increase temporary substitution across positions; and reduce workplace violations. Because the realized rate of algorithm-induced reassignment is itself an outcome of implementation, we interpret the full pattern of outcomes jointly rather than treating improved matching as the only mechanism consistent with treatment effects.

1.5 Sample Selection and Treatment Assignment

Our sample is based on groups of new hires to several different factory locations of a large garment manufacturing firm in Asia that produces clothing for export. Our partner firm first screened new workers to identify those suitable for the program, and could choose to exclude certain new hires from the job rotation program, particularly if these workers already had extensive experience in the job they were hired for. From this list of qualified candidates, we then selected workers for the sample before randomization. While workers were hired continuously over time, we grouped

new hires into study batches defined as intervals of 3 to 5 weeks.[7] Some batches were excluded from the rotation program because our rotation schedule required sufficient variability in the jobs available in that batch, given that it was designed around fixed job slots. For example, if all new hires in one month were for a single position, rotation across jobs would be impossible.

For batches that met the qualification of having a diversity of positions open such that treated workers could be rotated into one or two other jobs in addition to their initial assignment, we randomly allocated individuals into the treatment and control groups using a computer program.[8] Individuals in the treatment group were then randomly assigned to a job rotation schedule; the timeline for each batch is shown in Figure 1.

We expected high but not perfect compliance from the firm, both with the job rotation schedule and with permanent job assignment. Our surveys, combined with the administrative data, allow us to check the degree of compliance with both; we had targeted compliance rates of 80% or higher, and Section 3 reports and discusses the compliance rates we observe. The firm has adopted a decentralized management style where each factory has a certain amount of freedom in their management, as long as the target KPIs are met. While this study was strongly endorsed by top management, the pressure to meet target output means that, at the factory level, managers sometimes had to address immediate, often manpower-related, problems by deviating from our rotation schedule and matching recommendations.

We minimized non-compliance risks in several ways. First, each factory was informed that activities related to this study would not be counted in their KPIs. For instance, one of their main KPIs is the target ratio of sewers to non-sewers; during the study period, the central planning team did not count missed targets against a factory's KPI if they resulted from the job rotation program. Second, our field team maintained close communication with each location and frequently checked monitoring data to ensure that each location respected our rotation schedule and matching recommendations.

1.6 Sample and Statistical Power

We conducted a small pilot study in May 2023 involving 16 workers, each completing a total of two to three rotations. The goal of the pilot was primarily to demonstrate proof of concept to the firm and iron out logistical issues for the firm's implementation of the program. We did power calculations based on manager evaluation scores, aiming to recruit between 500 and 1,000 new hires for this project. We used numbers from the manager evaluations obtained from the pilot work

[7] The intervals varied by factory and time, depending on the speed of hiring. Batches averaged about 10-15 workers.

[8] The initial job assignment was done by the firm. Given that batches consist of workers who joined within an interval of 3-5 weeks, some workers may have been performing their initial jobs for 1 week, while others for as long as 4 weeks.

in May 2023 and made assumptions about the expected effect size. In practice, the partner company incurs costs from this project, both explicitly (such as the wage bills of dedicated personnel) and implicitly (such as the opportunity costs of deviating from the default of job assignments), so we judged a sample size between 500 and 1,000 new hires to be a reasonable balance between statistical power and cost to the firm. The project came to a natural close as the company faced an unstable international trade environment and slowing demand in 2025 and paused hiring; the first batch was in August 2023 and the last in May 2025. The resulting analysis sample includes 630 new hires (312 assigned to treatment and 318 to control; see Table 1), somewhat below our original target given the earlier-than-planned end of hiring.

1.7 Empirical Specification

The main analysis includes regressions that take the following form:

$$Y_i = \alpha + \beta Treat_i + X_i'\gamma + e_i, \tag{2}$$

where Y_i is an outcome of worker i, $Treat_i$ is whether worker i was assigned to treatment, and X_i includes control variables such as fixed effects for each hiring batch. Where baseline characteristics are imbalanced between the treatment and control groups, we include those variables as additional controls; for outcomes collected via survey, we also include enumerator fixed effects. e_i is an error term, and standard errors are clustered at the batch level. The key individual-level outcome variables include worker performance metrics, total number of absences, employee satisfaction, worker preference changes across various jobs, substitution rates, violations, and quits.

While the treatment occurs at the individual level, we are also interested in team-level outcomes, including team performance metrics. This is especially important given that some positions do not have performance metrics at the individual level, only at the team level. At the team level, we examine the impact of the entry of a new team member who is in the treatment versus the control group, using the administrative data to compare the periods before and after the entry of the new worker, separately by the new worker's treatment status.

Because the firm's administrative data form a worker-month panel, we also exploit the time-series dimension of the data beyond the cross-sectional specification in equation (1): we estimate Kaplan-Meier survival curves and event-study specifications for attendance outcomes over time, and stacked difference-in-differences specifications for the team-level production outcomes described in Section 3.

We address concerns about multiple hypothesis testing using Westfall-Young corrected p-values, both to address the multiplicity of outcomes and when examining treatment effect heterogeneity.

2 Data

2.1 Data Collection

The data for this study combine survey and administrative records. Prior to the start of each job rotation cycle, participants completed a short baseline survey and an aptitude test, the General Aptitude Test Battery (GATB), which has been shown to predict performance among garment manufacturing workers in both developed and developing country settings (Dagenais, 1990; Dolke, 2022). The baseline survey collected demographic information, prior work experience, and workers' preferences over the various job positions within the firm.

Although new workers joined the firm continuously, they were grouped into job rotation batches based on their date of entry, with each batch generally covering workers who arrived over approximately one month. At the end of each round of rotation, we surveyed the immediate supervisors of treated workers. Managers assessed, in an incentive-compatible manner, the performance of treated workers who had rotated onto their teams. Because treated workers rotated through a maximum of three jobs, we obtained up to three manager evaluations per treated worker. These evaluations provided one of the inputs into the algorithm used to determine workers' permanent job matches.

Upon completion of the rotation cycle, all study participants—both treatment and control workers—completed a midline survey. Among other outcomes, the survey elicited workers' preferences over jobs after the rotation experience. These preferences, together with managers' performance assessments, provided the remaining inputs into the algorithm used to determine final job matches.

The intervention ran from the first batch in August 2023 through the final batch in May 2025 (Figure A.1). This was later than the summer 2024 completion date we had originally anticipated, as hiring slowed during the broader decline in international trade demand in 2024–2025. The resulting sample consists of 630 new hires, within our pre-specified target range of 500–1,000 workers but below what would have been achieved had hiring continued at its initial pace.

Rather than waiting until all batches had completed their rotation cycles to field the endline, we scheduled each batch's endline approximately four months after the beginning of its rotation. Endline surveys were therefore fielded on a rolling basis as batches reached the four-month mark, with the first conducted in April 2024 and the last in September 2025 (Figure A.1).[9] This timing allowed workers to accumulate meaningful post-rotation experience before the endline measurement while limiting the length of follow-up and the associated risk of attrition.

[9]Because treated workers spend time learning new positions during the rotation, their tenure in any particular position is initially shorter than that of control workers, potentially generating a temporary performance disadvantage. Based on discussions with our partner organization, we expected differences associated with position-specific tenure to largely converge within approximately three months.

To collect information from workers who left the firm before their scheduled endline, beginning in April 2024 we also fielded a separate survey of workers who had quit. The endline worker survey collected information on job satisfaction, substitution into other jobs, and workers' experiences with the rotation program. We also surveyed managers at endline to obtain assessments of worker performance.

With the exception of the baseline survey and aptitude test, which were administered by the firm, trained enumerators administered all subsequent survey rounds on our behalf.

We complement the survey data with several sources of administrative records from the firm. Personnel records contain demographic characteristics and resignations; these records were extracted separately by factory between July 2025 and January 2026. We additionally obtained detailed timesheet records, which we use to measure attendance and hours worked, extracted between October 2025 and January 2026; team-level production records, extracted between October and December 2025; and workplace violation records, which report the description and date of each violation and the associated disciplinary action.

2.2 Summary Statistics

Table 1 presents baseline summary statistics for the analysis sample, separately by treatment status; treatment and control groups are balanced on every characteristic shown, consistent with successful randomization. The average new hire in the sample is in their late-20s, and about two-thirds of the workers are female. Roughly 70% of the sample have a secondary school degree or higher. While most new workers speak the local majority language, some new hires are migrants, as indicated by having a different native language. About 61% of the workers have some prior experience in the garment industry (averaging roughly 2.6 years), so close to 40% are entirely new to the industry. About 73% of the sample has a friend or family member already working at the firm, and about half of the sample has lived in the local city for over 10 years. The most common job (67%) is as a sewer, although the average self-assessed sewing skill at baseline is quite low, just above "basic" on a 1 (no skills) to 4 (advanced) scale.

At baseline, workers were also asked to rate their preferences over each of the 8 core jobs in the production process on a scale from 1 to 5, where 1 means they strongly prefer the job and 5 means they strongly dislike it; a value of 3 means neutral or no preference.Given that most of these workers have no prior experience in garment manufacturing and are answering the survey during orientation, before actually working in any specific job, it is not surprising that most respondents report near-neutral preferences with relatively little variation across positions. In addition to preferences, we are interested in whether people like the jobs they are good at, or whether there is a trade-off between jobs that are liked and jobs at which workers excel. We therefore also asked

workers how well they predict they will perform in each of the 8 core jobs, on a scale from 1 to 5 where 1 corresponds to unsatisfactory and 5 corresponds to well ahead of standard; a value of 3 corresponds to predicting they would be satisfactory in the position.

3 Results

3.1 Implementation and Assessments

Rotation compliance among treated workers was high: of treated workers with a scheduled rotation and a valid midline survey, 85.1% tried at least one of their scheduled rotation jobs and 82.2% tried all of them (Table A.1, Panel A), close to the 80% compliance rate we had targeted ex ante. On the matching side (Table A.1, Panel B), the algorithm recommended a department switch for 19.6% of treated workers, and 36.1% of workers who received a recommendation ultimately switched. Table 2 reports these same implementation outcomes as regressions on treatment status. Treated workers tried 0.91 more jobs than the control mean of 1.30 jobs (Panel A, column 1; $p < 0.01$); by construction, 20.4% of treated workers received an algorithmic switch recommendation and 6.7% had that switch accepted, compared to a control-group value of zero on both measures, since control workers were never run through the matching algorithm. The gap between the recommendation rate and the acceptance rate — roughly 20% recommended versus 7% accepted — reflects both worker and manager resistance to some recommended moves; managers, in particular, sometimes pushed back on replacing an experienced worker with a less experienced one in a given role, especially during busy production periods when doing so would depress a team's output.

Consistent with the modest rate of accepted switches, treated workers were 7 percentage points more likely than control workers to hold a job in a different department than their baseline department by endline, against a control mean of 10% ($p < 0.05$), and were 5 percentage points more likely to have experienced, relative to baseline, a switch to a job they weakly preferred to their original assignment (control mean 3%; $p < 0.05$). We find no evidence of differential survey attrition between treatment and control (Panel A; -2 percentage points against a control mean of 19%, not statistically significant).

Table 2, Panel B, reports the remaining endline survey outcomes. Managers' evaluations of workers in their endline jobs do not differ by treatment status (control mean 0.69 on a binary indicator for an above-median evaluation; treatment effect 0.01, not significant), nor do the number of days spent substituting into other jobs, the work satisfaction index, or self-reported happiness.

3.2 Preference Updating

We next examine whether preferences over jobs changed between baseline and endline. As workers experience different jobs during a rotation, we would expect them to update their views of those jobs, particularly the ones they actually tried. To measure this, for each worker we restrict attention to the jobs that were part of their batch's rotation menu, and we compile the preference rating (on the 1-5 scale, where 1 is "strongly prefer") that the worker assigned to each of those jobs.

Figure 2 plots the resulting distribution of ratings at baseline and at endline, separately by treatment status. At baseline, treated and control workers have nearly identical distributions, with most ratings clustered in the "somewhat prefer" and "neutral" categories. By endline, the control group's distribution is largely unchanged from baseline, while the treatment group's distribution has shifted visibly toward the "strongly prefer" category.

Table 3 quantifies this shift. The average endline preference rating for jobs on the rotation menu falls (i.e., preferences improve) by 0.168 points relative to the control-group mean of 2.82 ($p < 0.10$; column 2, controlling for the corresponding baseline preference). Disaggregating by preference category, this improvement is concentrated entirely in the top category: treated workers are 7.9 percentage points more likely to rate an exposed job as "strongly prefer" (column 3; $p < 0.01$), off a base rate of 15.9%, while the treatment coefficients on the remaining four categories are all small and not statistically distinguishable from zero.

We next examine whether the rotation also changed workers' preferences toward their own jobs. Table A.3 reports preferences for workers' original jobs and for the jobs they hold at endline. We do not detect a statistically significant overall change in either case. The ANCOVA coefficient on the endline preference rating is 0.106 for the original job and 0.084 for the endline job; neither is statistically distinguishable from zero, although both point toward slightly less favorable (higher) ratings among treated workers. When we disaggregate the ratings by preference category, however, treated workers are 5.1 percentage points more likely to place their original job in the "somewhat dislike" category ($p < 0.01$), relative to a control-group mean of 4.4%. We find a similar 5.1 percentage point increase for the job held at endline ($p < 0.01$), relative to a control-group mean of 4.2%.

Taken together, these results suggest that experiencing alternative jobs changes how workers evaluate the set of jobs available to them. Treated workers develop substantially more favorable views of some jobs in the rotation menu, while there is more limited evidence of a deterioration in their evaluations of their own jobs. One possible interpretation is that direct experience provides workers with information about jobs they had not previously performed, allowing them to reassess the relative attractiveness of their existing job.

3.3 Attendance and Violations

Turning to the administrative personnel data, we first examine raw monthly patterns measured after worker's batch date. Figure 3 shows that treated and control workers track each other closely in the first two months after the batch date, then diverge: treated workers have persistently fewer days absent starting around month 3, with a gap of roughly one day per month by month 12.

Table 4 summarizes these patterns with regressions using worker-month observations over the full post-treatment period (i.e., not restricted to the first 12 months). Treated workers have 0.29 fewer days absent per month than the control mean of 2.13 days (Panel A; $p < 0.01$), a 14% reduction. This decline is driven by sick leave (-0.14 days against a control mean of 0.71, $p < 0.01$), with no significant change in annual or personal leaves (Table A.5). Total hours worked rise by 3.46 hours per month relative to a control mean of 203.23 ($p < 0.05$), driven by regular hours (2.91, $p < 0.01$) rather than overtime (0.55, not significant); payroll days rise modestly (0.26, $p < 0.10$). Consistent with more hours worked, both net pay (4.10 USD, $p < 0.10$) and base pay (2.84 USD, $p < 0.05$) rise in Panel B, while incentive and overtime pay show small, insignificant increases. We show consistent patterns with alternative specifications using full pre- and post-batch panel rather than restricting to the post-period, both with and without worker fixed effects (Table A.6).

In Table A.7, we show no detectable change in workplace violations, which remain a rare outcome (control mean 0.04 violations per worker-month).

3.4 Resignation

The reduction in absenteeism is accompanied by a reduction in resignation, a slower-moving measure of worker retention (Table 5). Over the full follow-up period, treated workers are 7 percentage points less likely to have resigned than workers in the control group, relative to a control-group mean of 54% ($p < 0.10$). Treatment here spans the 21 days following the batch start date — the rotation itself together with the subsequent midline survey and matching process — so we separately examine resignations occurring during this window and those occurring afterward. The overall reduction in resignation emerges entirely after the rotation is complete: treated workers are 10 percentage points less likely to resign during the post-rotation period ($p < 0.05$). In contrast, the treatment effect on resignation during the rotation window is small, positive, and not statistically significant (3 percentage points relative to a control-group mean of 3%).

To illustrate the post-treatment result, Figure 4 plots Kaplan–Meier survival estimates for the period after the treatment window ends, restricted to workers still with the firm at that point — the same sample underlying the After Treatment estimate above. The two groups track closely for the first few months before the treated group's survival rate separates and remains persistently higher

through the end of the observed follow-up period, consistent with the regression estimate (log-rank p$= 0.035$).

3.5 Heterogeneous Effects and Mechanisms

The results so far help distinguish among several potential mechanisms. One possibility is that rotation improves workers' permanent job matches. However, widespread reassignment is difficult to reconcile with implementation: only about 7 percent of treated workers ultimately moved to an algorithm-recommended department, while the improvements in attendance and retention are evident for the treatment group as a whole. We also find no average treatment effect on managers' endline evaluations of individual performance or on overall job satisfaction. Thus, while some workers may benefit from a better permanent match, reassignment alone is unlikely to account for the average treatment effects.

A second possibility is cross-skilling: treated workers may become more valuable to the firm because they can be deployed temporarily across a broader range of jobs. At endline, however, treated workers do not report spending significantly more days substituting into other positions. This finding does not rule out other operational benefits from broader skills, but it provides little direct evidence that increased substitution is the primary individual-level channel.

A third possibility is learning or engagement. Treated workers' preferences over jobs in their rotation menu become substantially more positive even though most do not permanently switch departments. Rotation may therefore provide information not only about specific tasks but also about the range of opportunities available inside the firm, the worker's ability to perform different work, or the firm's willingness to invest in and solicit input from new hires.

We explore heterogeneity in treatment effects along several pre-specified dimensions, including prior garment-industry experience, the number and variety of jobs available in a worker's rotation batch, demographic characteristics, baseline preferences and self-assessed performance, and the worker's initial assignment. Because the sample is modest for heterogeneity analysis, we view these results as suggestive. Figure A.4 shows that no single moderator consistently predicts treatment effects across all attendance outcomes. The clearest pattern is by workers' initial self-assessment of performance: among workers who rated their own performance relatively poorly at baseline, treatment produces larger reductions in both days and hours absent.

Table A.4 provides complementary evidence from workers' qualitative assessments of the workplace. Among workers who initially rated their performance poorly, treatment increases agreement that promotion opportunities are not limited and that rewards are sufficient; both estimates are statistically significant at the 5 percent level. Treatment also produces a notably large, though imprecisely estimated, increase in workers' belief that they can perform a job on their own

initiative. These patterns are consistent with the idea that multi-skilling may be particularly valuable when workers initially have less confidence in their performance or when direct experience reveals attractive opportunities elsewhere within the firm.

3.6 Team-level Outcomes

The individual-level results above indicate that the treatment was beneficial to both workers and the firm, since absenteeism and turnover are each costly to the firm. To better understand the productivity implications of these individual-level gains, we turn to outcomes measured at the production-line level.

Because treatment was randomized at the individual level, teams vary in whether, and to what extent, they received new workers who were treated versus new workers who were control. This variation lets us ask whether the individual-level improvements documented above translate into line-level productivity gains. Lines receiving treated workers could outperform lines receiving an equivalent inflow of control workers because of more motivated workers. On the other hand, with treated workers periodically pulled out of their home team to substitute elsewhere in response to quits and absences on other lines — team-level output for treated workers' home lines could in principle be no higher, or even lower, than for control lines, even if firm-wide productivity rises.

Table 6 reports estimates from a stacked difference-in-differences design, where each rotation batch defines a stack and the unit of observation is a production line by month, over a window running from six months before to six months after the batch's start month. Because a line's exposure to a batch is not simply "treated or not" but a mix of treated and control new workers, we separately estimate the effect of treated-worker inflow and control-worker inflow, each measured as the number of batch workers of that type assigned to the line divided by the line's baseline manpower (its median staffing level over the sample period) and rescaled so that a coefficient corresponds to a 10 percentage point increase in worker inflow relative to baseline line size. Estimating both inflows separately matters because being assigned a treated worker is not only receiving a treatment effect but also receiving a new, relatively inexperienced worker; the control-worker inflow coefficient captures that bundled "new worker" effect on its own.

A 10 percentage point increase in treated-worker inflow raises output minutes (output quantity adjusted for the standard allowed minutes per garment) by 2,166 relative to a pre-period mean of 8,052, a 27% increase ($p < 0.05$), and raises efficiency (output minutes divided by input minutes) by 28.5 points relative to a pre-period mean of 71.0, a 40% increase ($p < 0.05$). The corresponding control-worker inflow coefficients are much smaller (69 output minutes and 1.2 efficiency points) and not statistically distinguishable from zero, and we can reject equality of the treated- and control-inflow coefficients at the 10% level for both output minutes ($p = 0.07$) and efficiency

($p = 0.09$). Realized treated-worker inflow across line-batch observations in our panel averages about 0.37 in these rescaled units (i.e., about 3.7% of a line's baseline manpower); applying the estimated coefficients to this realized average inflow implies that an average line receiving treated workers experienced roughly a 10% increase in output minutes and a 14.8% increase in efficiency relative to its own pre-period mean. Figure 5 illustrates the event study estimates of the team productivity outcome.

These results are robust to the sample-construction choices underlying the stacked design (Appendix Table A.8). Varying the manpower floor from 10 to 5, 15, or 20 leaves the treated-inflow coefficients essentially unchanged (2,166–2,429 output minutes; 22.8–28.5 efficiency points), while control-inflow effects remain indistinguishable from zero. Allowing overlapping rotation rounds, so that a line can appear in more than one nearby batch's window, roughly halves the point estimates (1,032 output minutes; 11.9 efficiency points), consistent with blending a line's response across adjacent batches, but both remain statistically significant. We also continue to reject equality of the treated- and control-inflow coefficients ($p = 0.02$ and $p = 0.07$).

4 Benefit–Cost Assessment

To gauge whether the observed effects are large enough to matter for firm decision-making, we assemble a simple benefit–cost accounting around the rotation decision. We write the benefits of rotating a worker as

$$B = \underbrace{\Delta q \cdot \kappa w}_{\text{turnover}} + \underbrace{\Delta y \cdot p_m \cdot T}_{\text{productivity}} \tag{3}$$

and the costs as

$$C = \underbrace{h\, w^S}_{\text{trainer/manager time}} + \underbrace{\pi_{switch} \cdot r\,(1-\phi)\, w}_{\text{switching ramp-up}}, \tag{4}$$

where Δq is the change in the resignation probability, κ is the replacement cost of a departure measured in weeks of wages, w is the operator's weekly wage, T is the horizon in months over which benefits accrue, Δy is the monthly change in line output-minutes attributable to rotation, and p_m is the dollar value of an output-minute. On the cost side, h is the number of supervisor-equivalent weeks devoted to training and assessing fit for a rotated worker, w^S is the supervisor's weekly wage, π_{switch} is the probability of a subsequent task switch, r is the duration of the associated ramp-up period in weeks, and ϕ is the fraction of normal productivity retained during that period.

Benefits. Treatment reduces post-rotation resignation by 10 percentage points (Table 5). Fol-

lowing Blatter et al. (2012), who estimate total hiring costs of 10–17 weeks of pay using Swiss data, we assume $\kappa = 10$ weeks for our semi-skilled setting, close to the 10.5 weeks that Kuhn and Yu (2020) estimate for a retailer in China.[10] At the control-group mean pay of 249 USD per month (57 USD per week), the turnover benefit is $0.10 \times 574 \approx 57$ USD per treated worker.

Absenteeism also falls, by roughly 0.29 days per month. We do not monetize this reduction separately because its benefit may already be reflected in our measure of realized line-level output: to the extent that lower absenteeism increases production, it contributes directly to the productivity gains we estimate below.[11]

Productivity. We value an output-minute at its replacement labor cost: the wage per minute worked ($0.02, the control-group mean monthly wage divided by standard monthly hours) divided by mean efficiency (71%), yielding $p_m \approx \$0.029$ per output-minute. We base the calculation on the realized average treatment intensity in treated lines: an inflow of treated workers equal to 3.7% of a line's baseline manpower is associated with a 10% increase in output-minutes and a 14.8% increase in efficiency (output-minutes per input-minute), consistent with a productivity gain rather than simply an increase in hours worked. At the baseline mean of 8,052 output-minutes per line per month, a 3.7% inflow increases output by approximately 805 output-minutes per month, worth about $23 per month, or $280 per line-year. At the median line size of 22 workers, a 3.7% inflow corresponds to approximately 0.81 treated workers. Dividing the annual line-level benefit by this number yields a productivity benefit of approximately $345 per treated worker.

Costs. Line supervisors earn 35% more than operators, so $w^S = 1.35 \times \$57 \approx \$77$ per week. Allowing a supervisor to spend 25–50% of their time training workers and evaluating fit over two weeks implies $hw^S = (0.25\text{–}0.5) \times 2 \times \$77 \approx \$39\text{–}\$77$. We do not assign a separate cost to evaluation time because assessing fit occurs as part of training oversight, and h is defined broadly enough to include it. A further 6.7% of rotated workers switch tasks again after the initial rotation. Valuing the associated ramp-up at two weeks at 50% productivity gives $\pi_{switch} \cdot r(1-\phi)w = 0.067 \times 2 \times 0.5 \times \$57 \approx \$4$. Supervisor time and subsequent switching costs therefore sum to approximately $43–$81 per treated worker.

A rotated worker's own foregone output while learning the new task is also a cost in an accounting sense: two weeks at approximately half productivity, valued at a $57 weekly wage, amounts to roughly $57. We do not add this cost separately, however, because it is already reflected in the realized line-level productivity effect. The same applies to any output lost when experienced coworkers assist with training. Because both occur on the line during the period over which output

[10] The firm's own estimate of recruiting and initial training costs, roughly 300 USD per hire, is a lower bound covering direct costs only; Blatter et al. (2012) estimate that direct costs are about one-fifth of total hiring costs. Using the firm's figure alone, the turnover benefit is 30 USD per treated worker.

[11] We also do not separately net out the change in the wage bill from reduced absence: although wages are paid on a daily-rate basis, absence is typically covered by relief workers or overtime.

is measured, the estimated productivity effect is already net of these losses.

Summary. Combining the estimated turnover and productivity benefits yields approximately $402 per treated worker ($57 + $345), compared with $43–$81 in costs not already reflected in realized line output. This implies a net benefit of approximately $321–$359 per treated worker, equivalent to 5.6–6.3 weeks of the average worker's wages.

This calculation is necessarily approximate and relies on assumptions about the valuation and persistence of the estimated effects. Nevertheless, it suggests that the gains from multi-skilling are economically meaningful relative to its implementation costs. Importantly, the program's costs are immediate, salient, and borne locally by the workers and managers implementing rotation, whereas many of its benefits accrue over a longer horizon through higher line productivity and lower turnover. This mismatch in the timing and incidence of costs and benefits may help explain why apparently profitable management practices are not more widely implemented, a possibility we return to below.

5 Discussion

5.1 Why Did Preference Learning Lead to Limited Reallocation?

The results in Section 3 show that rotation changes workers' preferences toward jobs they experience, yet relatively little of this preference updating translates into permanent reallocation. The matching procedure recommended reassignment for 61 workers. For 57 of these workers, we can compare the recommended assignment with the worker's initial job using stated preferences;[12] 27 (47%) strictly preferred the recommended assignment and another 21 (37%) ranked the two equally. Thus, 48 of 57 workers (84%) weakly preferred the recommended assignment. Nevertheless, only 18 of these 57 recommendations (32%) were ultimately accepted.[13] Limited reallocation therefore cannot be explained simply by workers preferring to remain in their initial assignments.

To understand this gap, we use contemporaneous case notes collected by the field team documenting communication with factories during implementation. We focus on cases in which factory staffing, production, or managerial considerations prevented or delayed a recommended move. Not every recommendation that failed to proceed as evidence of an organizational friction: workers sometimes changed their minds, left the firm, or faced job-specific restrictions. Appendix Table A.9 provides a fuller accounting of these cases.

The case notes point repeatedly to constraints on releasing workers from their current teams.

[12] The remaining four workers lack midline preference data needed for this comparison.

[13] We define a recommendation as accepted if both the worker and factory management agreed to the reassignment and management subsequently followed through in implementing it.

Factories cited shortages of sewers, the need to find or train replacements, limited relief-team capacity, and existing headcount and line-balancing plans. Managers also sometimes preferred to retain workers they regarded as particularly valuable in their current roles. In one case, for example, a recommended move to cutting was delayed while HR sought a replacement for the worker's sewing position and later placed on hold because of insufficient relief-team capacity; the field notes further observed that the worker was "smart as well so he is valuable for the current team."

These cases highlight a distinction between generating information about worker-job fit and acting on that information. Reassigning a worker who appears better matched elsewhere can create a vacancy in the sending team, require a replacement to be trained, or disrupt existing production plans. Moreover, the benefits of improved allocation may accrue broadly to the firm while the incumbent team bears the immediate cost of releasing and replacing a productive worker. Although the case notes do not allow us to identify such incentive conflicts causally, they show how local staffing and production constraints can prevent information about potentially better worker-job matches from translating into actual reallocation.

6 Conclusion

We study the consequences of multiskilling for workers and firms through an RCT at a large garment manufacturer in Southeast Asia. We find sizable benefits. Multiskilling reduces worker absenteeism and turnover, while increasing production-line output and efficiency. These results suggest that training workers to perform multiple jobs, rather than immediately specializing them in a single job, can benefit both workers and firms.

Two seemingly natural benefits of multiskilling, however, largely fail to materialize. First, although workers develop more favorable views of the jobs they experience during rotation, relatively few workers ultimately move to newly preferred jobs. Thus, the firm makes limited use of the information about worker-job fit generated by the rotations. Second, multiskilled workers do not subsequently substitute across jobs more frequently, providing little evidence that the intervention creates a more agile workforce in practice. Our qualitative evidence helps explain this gap between potential and realized benefits. Production constraints, staffing needs, and managers' reluctance to release valuable workers limit the firm's ability to reallocate workers even when potentially beneficial matches are identified.

The benefits we observe instead appear to extend beyond reallocation and flexibility. Treatment effects are largest, although imprecisely estimated, among workers who initially assess their own performance poorly. These workers also become more likely to report that promotion opportunities are not limited and that rewards are sufficient. This evidence suggests that experiencing different

jobs may broaden workers' perceptions of their own capabilities and opportunities within the firm.

Our findings have implications for how firms design training, particularly in settings where workers enter with limited experience and relatively narrow job responsibilities. Firms often use pay, bonuses, promotions, and recognition to motivate and retain workers. Our results suggest that the design of work and training itself may provide another lever, even when the additional skills workers acquire are not subsequently used to reallocate them. An important question for future work is how the returns to multiskilling compare with more conventional approaches to motivating and retaining workers, such as cash incentives or recognition.

Several questions remain. Multiskilling and job rotation are common in settings well beyond garment manufacturing, including among professional and managerial workers, where the returns to learning about one's comparative advantage and developing broader organizational knowledge may be different. Understanding when and for whom multiskilling is most valuable is therefore an important direction for future research. More broadly, our findings highlight a gap between creating valuable information and flexibility and firms' ability to use them. Future work could study how organizational design and managerial incentives can reduce these barriers and unlock further gains from multiskilling.

References

Acemoglu, Daron and Jorn-Steffen Pischke, "Why Do Firms Train? Theory And Evidence," *Quarterly Journal of Economics*, 1998, *113* (1), 78–118.

_ and _, "Beyond Becker: Training in Imperfect Labour Markets," *Economic Journal*, 1999a, *109* (453), F112–42.

_ and _, "The Structure of Wages and Investment in General Training," *Journal of Political Economy*, 1999b, *107* (3), 539–572.

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham, "Returns to on-the-job soft skills training," *Journal of Political Economy*, 2023, *131* (8), 2165–2208.

Adler, Paul S., Barbara Goldoftas, and David I. Levine, "Flexibility versus efficiency? A case study of model changeovers in the Toyota production system," *Organization Science*, 1999, *10* (1), 43–68.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen, "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan," *Quarterly Journal of Economics*, 2017, *132* (3), 1101–1164.

Banerjee, Abhijit and Esther Duflo, "Addressing absence," *Journal of Economic perspectives*, 2006, *20* (1), 117–132.

Bartel, Ann P, "Training, Wage Growth, and Job Performance: Evidence from a Company Database," *Journal of Labor Economics*, 1995, *13* (3), 401–25.

Bartel, Ann P., "Measuring the Employer's Return on Investments in Training: Evidence from the Literature," *Industrial Relations: A Journal of Economy and Society*, 2000, *39* (3), 502–524.

Becker, Gary, *Human Capital: A Theoretical and Empirical Analysis*, Chicago, IL: University of Chicago Press, 1964.

Blatter, Marc, Samuel Muehlemann, and Samuel Schenker, "The Costs of Hiring Skilled Workers," *European Economic Review*, 2012, *56* (1), 20–35.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

_ , Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does Management Matter? Evidence from India," *Quarterly Journal of Economics*, 2013, *128* (1), 1–51.

_ , Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen, "What Drives Differences in Management Practices?," *American Economic Review*, 2019, *109* (5), 1648–83.

Dagenais, Fred, "General Aptitude Test Battery Factor Structures for Saudi Arabian and American Samples: A Comparison," 1990.

Dessein, Wouter, "Authority and communication in organizations," *Review of Economic Studies*, 2002, *69* (4), 811–838.

Dolke, Ashoke M., "General Aptitude Test Battery (GATB) Norms for Jobs in Textile Mills," 2022.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman, "Labor market dynamics and develop-

ment," *The Quarterly Journal of Economics*, 2023, *138* (4), 2287–2325.

Ellingsen, Tore and Magnus Johannesson, "Paying Respect," *Journal of Economic Perspectives*, 2007, *21* (4), 135–150.

_ **and** _ , "Pride and Prejudice: The Human Side of Incentive Theory," *American Economic Review*, 2008, *98* (3), 990–1008.

Espinosa, Miguel and Christopher T. Stanton, "Training, Communications Patterns, and Spillovers inside Organizations," *Journal of Political Economy*, 2026, *134* (7), 1933–1966.

Falk, Armin and Michael Kosfeld, "The Hidden Costs of Control," *American Economic Review*, 2006, *96* (5), 1611–1630.

Friebel, Guido and Michael Raith, "Talent Management: The Role of Bosses," *Journal of Labor Economics*, forthcoming.

Gale, David and Lloyd Shapley, "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 1962, *69* (1), 9–15.

Grip, Andries De and Jan Sauermann, "The Effects of Training on Own and Co-worker Productivity: Evidence from a Field Experiment," *The Economic Journal*, 2012, *122*.

Haegele, Ingrid, "Talent Hoarding in Organizations," *American Economic Review*, 2026, *116* (8), 3110–3151.

Hart, Oliver and Bengt Holmstrom, "A Theory of Firm Scope," *Quarterly Journal of Economics*, 2010, *125* (2), 483–513.

He, Qiwei and Michael Waldman, "Intra-firm Employer Learning, Talent Hoarding and Worker Visibility," *SSRN Electronic Journal*, 2026.

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi, "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines," *AER*, 1997, *87* (3), 291–313.

Jovanovic, Boyan, "Job Matching and the Theory of Turnover," *JPE*, 1979, *87* (5), 972–90.

Kuhn, Peter J. and Lizi Yu, "How Costly is Turnover? Evidence from Retail," *Journal of Labor Economics*, 2020, *Forthcoming*.

Marks, Michelle A, Mark J Sabella, C Shawn Burke, and Stephen J Zaccaro, "The impact of cross-training on team effectiveness.," *Journal of Applied Psychology*, 2002, *87* (1), 3.

Molleman, E and J Slomp, "Functional flexibility and team performance," *International Journal of Production Research*, 1999, *37* (8), 1837–1858.

Pigou, Arthur C., *Wealth and Welfare*, London: Macmillan, 1912.

Roth, Alvin E. and Elliott Peranson, "The Redesign of the Matching Market for American Physicians : Some Engineering Aspects of Economic Design," *American Economic Review*, 1999, *89* (4).

Slomp, Jannes, Jos AC Bokhorst, and Eric Molleman, "Cross-training in a cellular manufacturing environment," *Computers & Industrial Engineering*, 2005, *48* (3), 609–624.

Ton, Zeynep, *The Good Jobs Strategy*, Houghton Mifflin Harcourt. New York, NY, 2014.

_ **and Simon Harrow**, "Mercadona," Technical Report 610-089, Harvard Business School 2010.

Van den Steen, Eric, "Interpersonal Authority in a Theory of the Firm," *American Economic Review*, 2010, *100* (1), 466–490.

Volpe, Catherine E, Janis A Cannon-Bowers, Eduardo Salas, and Paul E Spector, "The impact of cross-training on team functioning: An empirical investigation," *Human factors*, 1996, *38* (1), 87–100.

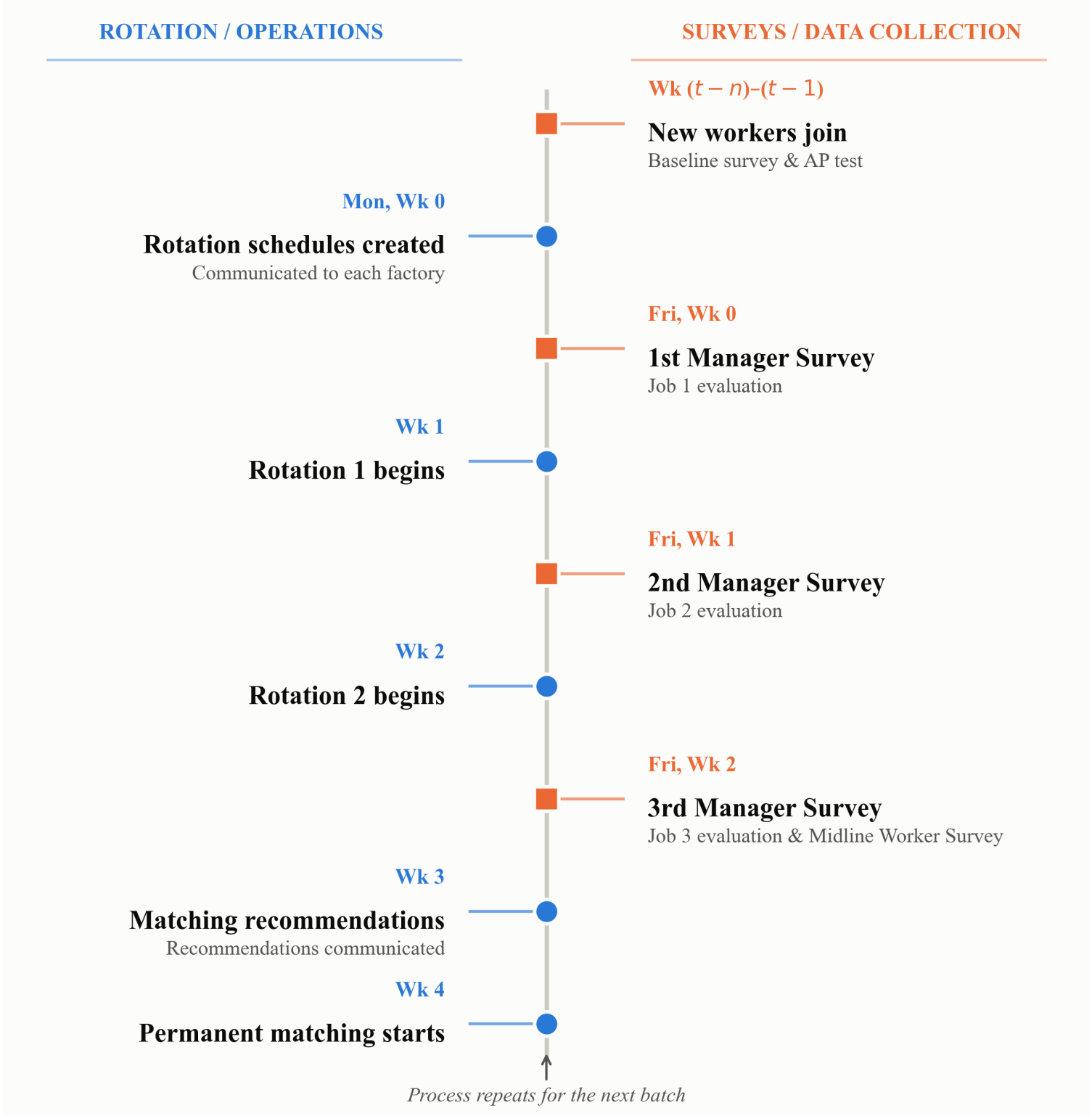
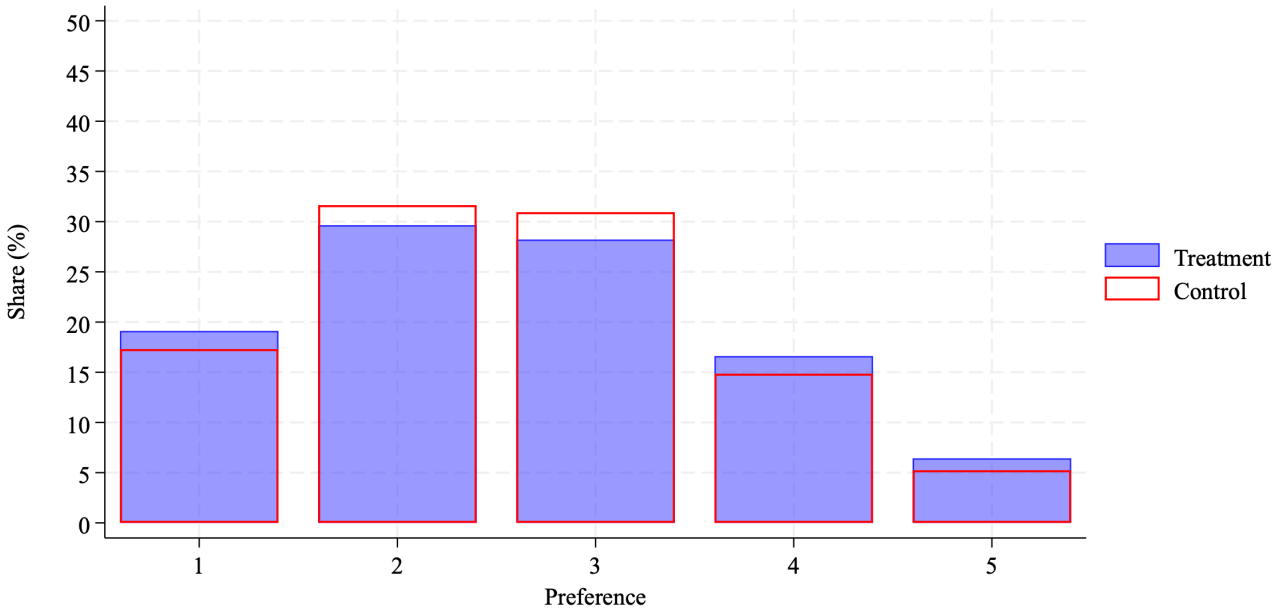
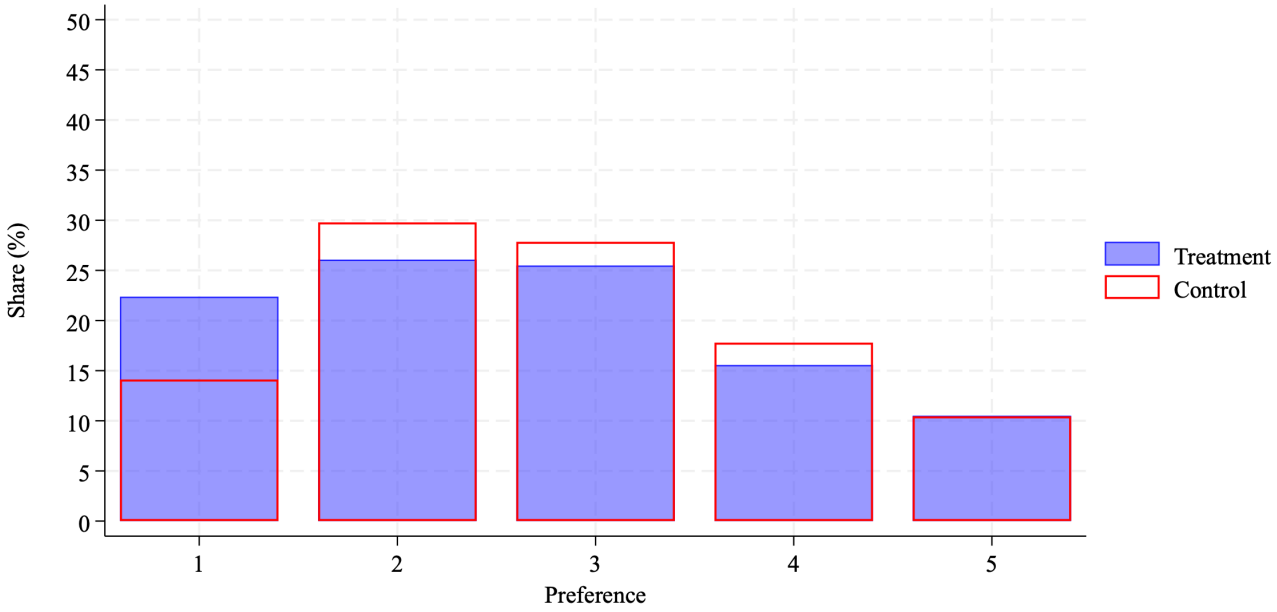


Figure 1: Timeline for Each Batch

This figure depicts the timeline for a single batch's rotation cycle. Blue circles denote rotation and scheduling milestones; orange squares denote survey and data-collection events. This process repeats for each subsequent batch, with the last batch beginning in May 2025.



(a) Baseline Preferences

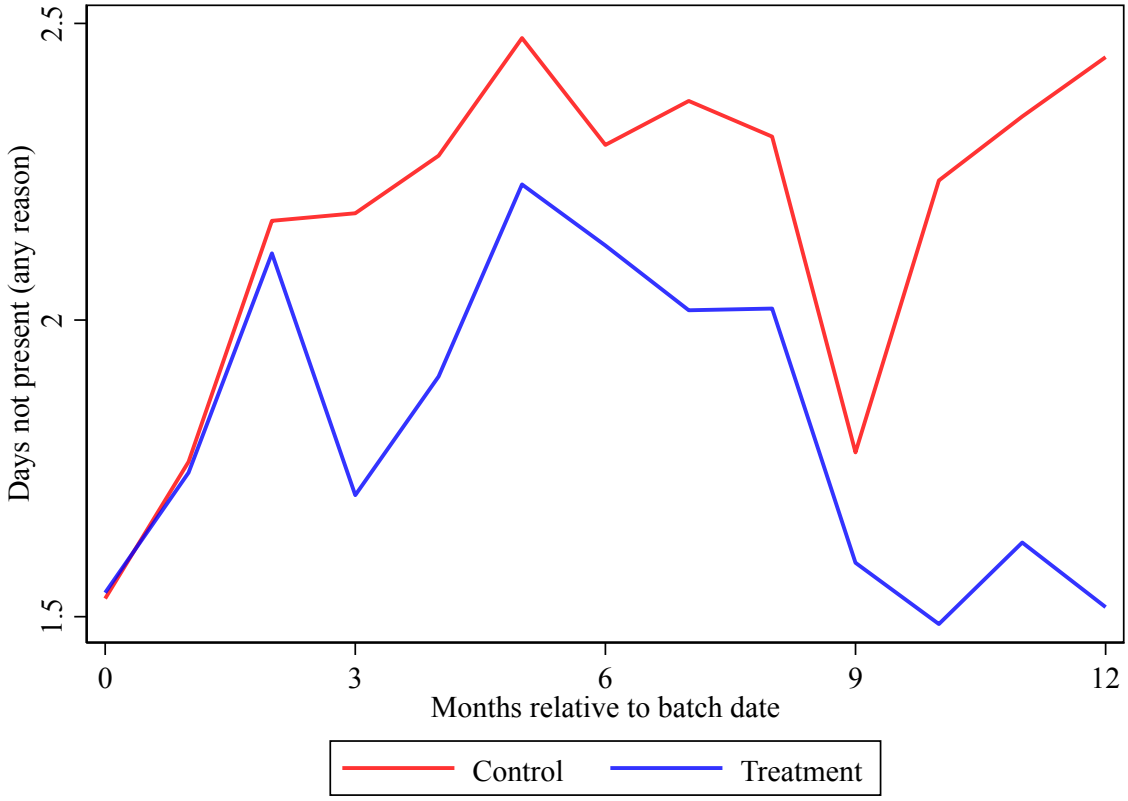


(b) Endline Preferences

Figure 2: Baseline and Endline Preferences

Notes: Distribution of stated job preferences at baseline (panel a) and endline (panel b), pooled across all six rotation jobs (cutting, sewing, embroidery, heat transfer, ironing/packing, material stock). For each job, the sample includes workers whose batch was rotated through that job; workers rotated through multiple jobs contribute one preference observation per rotated job, so observations are worker-job pairs rather than one observation per worker. Bars show the share of preference ratings (1–5 scale) by treatment status (solid blue: treatment; red outline: control).

Figure 3: Monthly Absence by Treatment Status



Notes: The figure plots raw monthly means by months relative to the batch date, separately for control (red) and treatment (blue) workers. The sample is a worker–month panel restricted to the batch month and the twelve months following it.

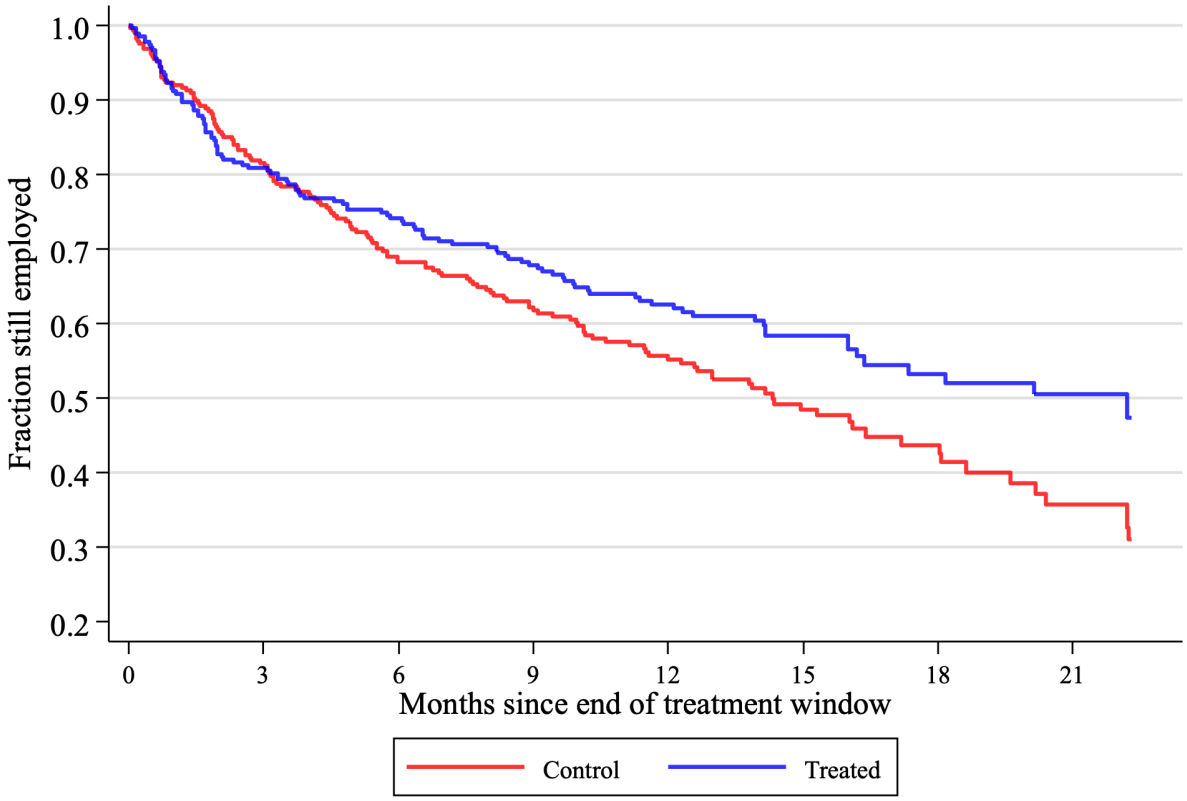
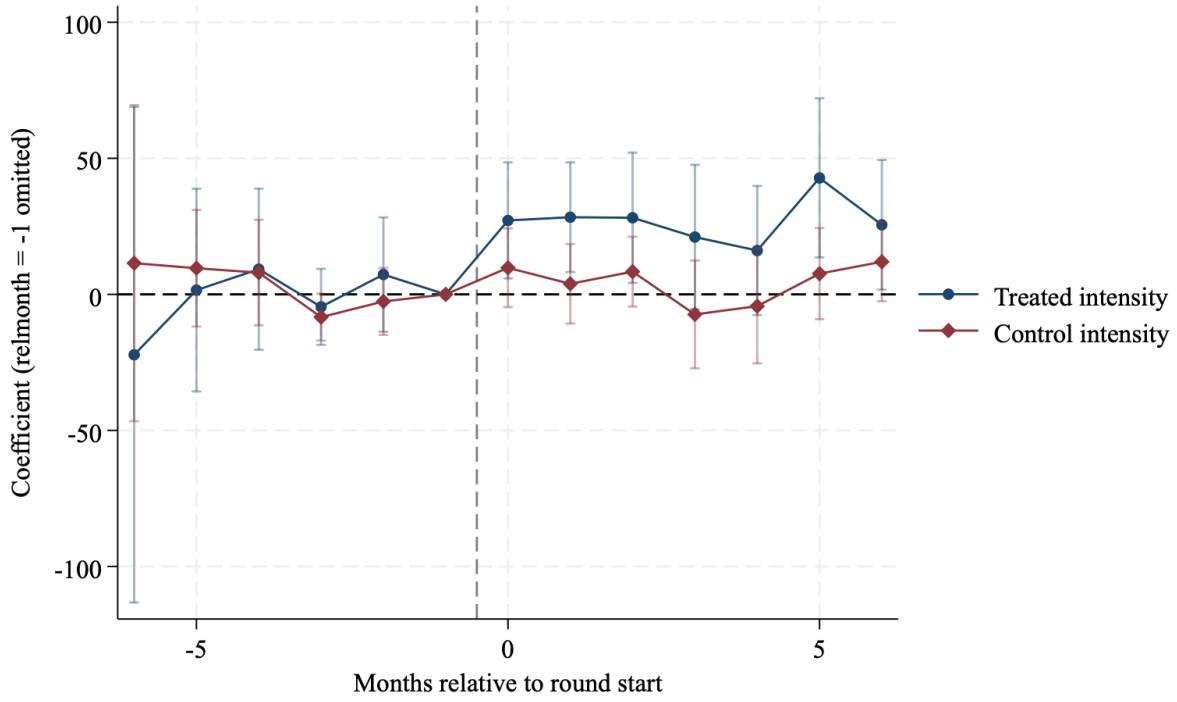


Figure 4: Kaplan–Meier Survival Estimates by Treatment Status

Note: Kaplan–Meier estimates of the share of workers still employed at the firm, by treatment status, restricted to workers who had not separated from the firm by the end of the 14-day rotation window, with analysis time measured from the end of that window. The failure event is separation from the firm; workers who had not separated are censored at the last observed separation date at their factory.

Figure 5: Event Study: Team Productivity Outcomes Around Rotation Batch Start



Notes: Estimates from a stacked difference-in-differences design. Each batch round defines a stack; within a stack the unit of observation is a production line X month, over the thirteen-month window running from six months before to six months after the round's start month. A line enters a stack if its baseline size – the median monthly manpower reading for that line – is at least ten workers and if productivity data are available for every month of the window; a round is dropped for a line when its window would overlap that of the previous retained round on the same line. Treated (control) intensity is the number of batch workers with treatment status one (zero) assigned to that line, divided by the line's baseline size, rescaled so that a one-unit increase corresponds to a 10-percentage-point increase in worker inflow relative to baseline line size. The regression interacts each intensity measure with a full set of relative-month indicators, omitting relmonth $= -1$, and absorbs stack X line and stack X relative-month fixed effects. Navy circles trace the treated-inflow path and maroon diamonds the control-inflow path, each measured relative to relmonth $= -1$; the vertical gap between the two paths at a given horizon is the dynamic treatment effect. Vertical bars are 95% confidence intervals from standard errors clustered at the line level. The dashed vertical line separates pre- and post-round months. Endpoint coefficients at $k = -6$ and $k = +6$ pool all relative months at or beyond those horizons into a single estimate.

Table 1: Baseline Summary Statistics

	Treatment			Control			
	Mean	SD	N	Mean	SD	N	p-value
Tenure $>$ 1 Month at Batch Start (0/1)	0.49	0.50	312	0.53	0.50	318	0.34
Tenure (days) at Batch Start	100.92	336.25	312	77.08	126.33	318	0.24
Age	28.94	7.16	293	28.19	6.75	303	0.19
Male (0/1)	0.35	0.48	312	0.33	0.47	318	0.61
Schooling Secondary or Above (0/1)	0.68	0.47	285	0.73	0.45	290	0.25
Native Language:							
... Language 1 (0/1)	0.14	0.34	299	0.12	0.32	306	0.47
... Language 2 (0/1)	0.38	0.49	299	0.36	0.48	306	0.52
... Language 3 (0/1)	0.02	0.13	299	0.01	0.08	306	0.24
... Language 4 (0/1)	0.40	0.49	299	0.43	0.50	306	0.46
... Language 5 (0/1)	0.08	0.27	299	0.08	0.27	306	0.83
Live in the City for Over 10 Years (0/1)	0.50	0.50	294	0.49	0.50	293	0.90
Aptitude Test: Total Score	65.56	35.91	296	66.25	35.01	306	0.81
Sewing Skills (1-4; 1=None; 4=Advanced)	2.21	0.98	287	2.20	0.97	289	0.92
Has Firm Network (0/1)	0.73	0.44	294	0.71	0.45	303	0.55
Prior Experience in Garment (0/1)	0.61	0.49	288	0.58	0.49	286	0.45
Years of Prior Experience in Garment	2.57	3.75	284	2.71	4.75	288	0.70

Notes: The table shows the baseline summary statistics. The p-value is from a regression testing the statistical difference between the treatment and the control groups.

Table 2: Endline Implementation and Survey Outcomes

	Treat	Control Mean	Obs
Panel A: Implementation			
Number of Positions	0.91***	1.30	499
	(0.08)		
Recommended Switch (0/1)	0.20***	0.00	630
	(0.03)		
Accepted Switch (0/1)	0.07***	0.00	630
	(0.02)		
Different Dept than Baseline (0/1)	0.07**	0.10	494
	(0.03)		
Preferred Switch (0/1)	0.06***	0.03	494
	(0.02)		
Survey Attrition (0/1)	-0.02	0.19	605
	(0.02)		
Panel B: Endline Survey Outcomes			
Manager Evaluation (0/1)	0.01	0.69	602
	(0.03)		
Days as Substitute	0.43	1.14	498
	(0.45)		
Work Satisfaction Index	-0.06	0.01	499
	(0.09)		
Happiness Ladder (0-10)	-0.11	7.22	499
	(0.21)		

Notes: Each row presents a regression of the outcome on treatment status, with batch fixed effects and standard errors clustered at the batch level. *Recommended Switch* and *Accepted Switch* indicate whether the algorithmic job-matching step recommended, and the worker accepted, a job switch; these are coded 0 for control workers, who were never eligible for a recommendation. *Different Dept than Baseline* equals 1 if the worker's endline department differs from their baseline department; *Preferred Switch* equals 1 if the worker's endline department is different from the baseline department and the former is preferred at endline survey. These two variables are restricted to non-attrited workers by endline survey. *Manager Evaluation* is measured on a Likert scale from 1 to 5, where [1] = Well Ahead of Standard and [5] = Unsatisfactory, recoded into a binary indicator equal to 1 if the evaluation score is greater than or equal to the median value in the full dataset. The *Work Satisfaction Index* is based on all work satisfaction items, harmonized so that a higher index means greater satisfaction. *Number of Positions* (Panel A) and *Days as Substitute*, *Work Satisfaction Index*, *Happiness Ladder* (Panel B) include worker-side endline enumerator fixed effects; *Manager Evaluation* (Panel B) includes manager-side endline enumerator fixed effects. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.10$.

Table 3: Preference Changes for Exposed Jobs

	Endline Level		By Preference Value				
	Mean	Mean	Pref=1	Pref=2	Pref=3	Pref=4	Pref=5
Treatment	-0.162*	-0.168*	0.079***	-0.018	0.003	-0.009	0.008
	(0.083)	(0.090)	(0.019)	(0.026)	(0.024)	(0.024)	(0.019)
Baseline mean pref.		0.176***	-0.027*	-0.012	-0.009	0.015	0.038***
		(0.050)	(0.015)	(0.017)	(0.015)	(0.015)	(0.011)
Control mean	2.89	2.89	.122	.252	.235	.155	.0953
Observations	499	499	499	499	499	499	499

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports OLS estimates of the effect of treatment assignment on endline preference outcomes. For each worker, outcomes are constructed by restricting preferences to jobs designated as rotation jobs in their batch and averaging across these exposed jobs. Columns (1)–(2) examine the mean endline preference level; columns (3)–(7) examine the average across exposed jobs of an indicator for whether the worker holds preference value k for that job, for $k \in \{1, 2, 3, 4, 5\}$. Where indicated, we also control for the corresponding baseline preference measure. Baseline preference is missing for workers whose exposed jobs all lack a baseline value; in those cases we impute the sample mean and include a missing-baseline indicator. All regressions include enumerator and batch fixed effects. Standard errors are clustered at the batch level.

Table 4: Effect of Treatment on Attendance and Payroll Outcomes

	Panel A: Absence and Working Hours			
	Days Not Present	Total Hours	Regular Hours	Overtime Hours
Treatment	-0.29***	3.46**	2.91***	0.55
	(0.09)	(1.46)	(1.08)	(0.55)
Control Mean (Post)	2.13	203.23	174.84	28.39
Observations	5770	5770	5770	5770
	Panel B: Payroll			
	Total Pay	Base Pay	Incentive	Overtime Pay
Treatment	4.10*	2.84**	0.65	0.61
	(2.10)	(1.24)	(0.60)	(0.95)
Control Mean (Post)	248.74	197.77	2.09	48.87
Observations	6942	6942	6942	6942

Notes: Each column reports estimates from an OLS regression using worker-month observations in the post-treatment period. All specifications include batch and calendar-month fixed effects. Standard errors are clustered at the batch level. Control means are computed over the post-period for control workers. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.10$.

Table 5: Effect of Job Rotation on Resignation

	(1) Overall	(2) During Treatment	(3) After Treatment
Treatment	-0.07*	0.03	-0.10**
	(0.04)	(0.02)	(0.04)
Control Mean	0.54	0.06	0.49
Observations	606	606	606

Notes: Each column reports a separate regression on a cross-section of workers, one observation per worker. The rotation window is the 14 days beginning on the batch date. "Overall" equals one if the worker resigned at any point after the batch date; "During Rotation" equals one if the resignation occurred within the rotation window; "After Rotation" equals one if it occurred afterward. Workers who resigned on or before the batch date are excluded. Workers who did not resign are censored at the last observed resignation date at their factory. All columns include batch fixed effects. Standard errors clustered at the batch level in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Stacked DiD: Team Productivity Outcomes by Treated & Control Worker Inflow

	(1) Output Minutes	(2) Efficiency
Treated inflow x Post	2166.45**	28.53**
	(816.79)	(11.77)
Control inflow x Post	68.52	1.23
	(537.00)	(6.97)
Pre-period mean	8052.02	71.04
p-value: Treat = Control	0.07	0.09
Observations	833	833

Notes: Stacked difference-in-differences at the line-round level, using a ± 6-month event window around each rotation batch's start date. Output minutes (col. 1, output quantity $\times$ Standard allowed minutes) reflect realized production; efficiency (col. 2) is output minutes divided by total input minutes worked, expressed as a percentage. Treated inflow $\times$ Post" and Control inflow $\times$ Post" measure treated/control workers assigned to the line in that round as a share of the line's baseline manpower, rescaled so that a one-unit increase corresponds to a 10-percentage-point increase in worker inflow relative to baseline line size, and interacted with a post-round indicator. Baseline manpower is the median valid manpower reading for the line. All specifications include stack-line and stack-relative-month fixed effects; standard errors (in parentheses) are clustered at the team level. In the main specification, lines with baseline manpower below 10 are dropped, each stack-line must have productivity data for every month of the event window, and a batch/round is dropped if its window overlaps the previous kept round for that line. Pre-period mean" is the outcome's mean in the estimation sample prior to the round. p-value: Treat = Control" tests equality of the treated- and control-inflow coefficients. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Web Appendix for "Multi-Skilling Workers in Organizations: Experimental Evidence from the Garment Industry"

by Ratchanon Chotiputsilp, Mitchell Hoffman, Taeho Kim, and Shing-Yi Wang

The Web Appendix is organized as follows. Appendix A presents additional figures and tables.

A Additional Figures and Tables

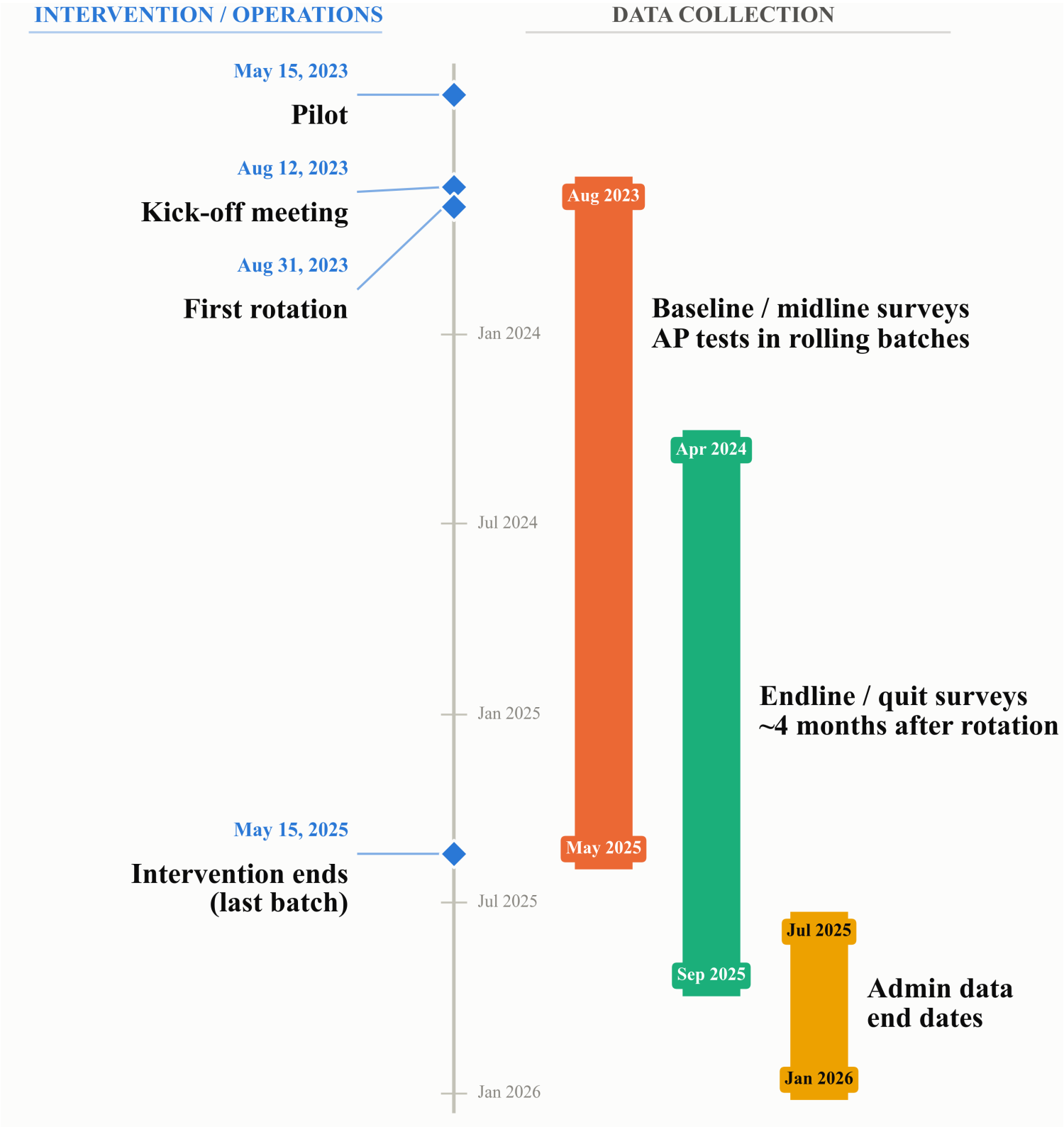
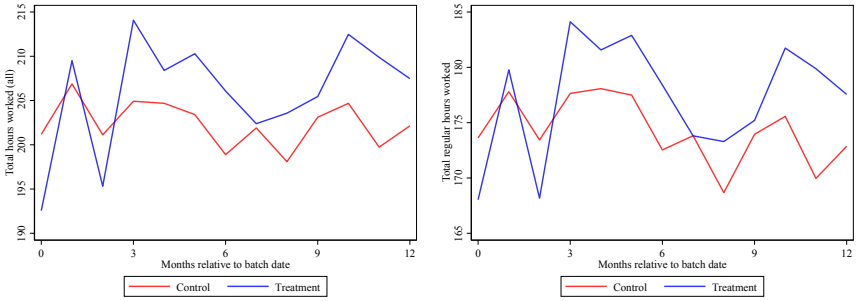


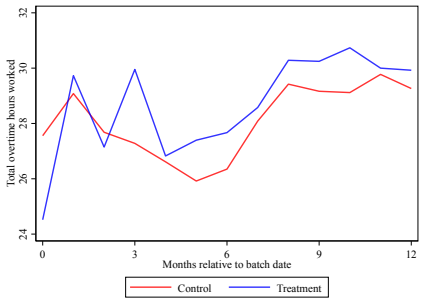
Figure A.1: Study Timeline

This figure depicts the realized timeline for the study. Blue diamonds denote point-in-time milestones. Because batches were rotated in on a rolling basis, baseline/midline data collection (orange) and endline and quit-survey fielding (teal) each occurred over an extended, per-batch or per-factory window rather than on a single date; bar endpoints show the earliest and latest date observed across batches or factories within each window.

Figure A.2: Total Work Hours



(a) Total Hours **(b)** Regular Hours Worked



(c) OT Hours Worked

Note: Each panel plots average monthly outcomes relative to the batch date, separately for treatment (blue) and control (red) workers. Timesheet hours are based on recorded clock-in and clock-out timestamps and capture actual hours worked. Regular and overtime hours are components of total timesheet hours.

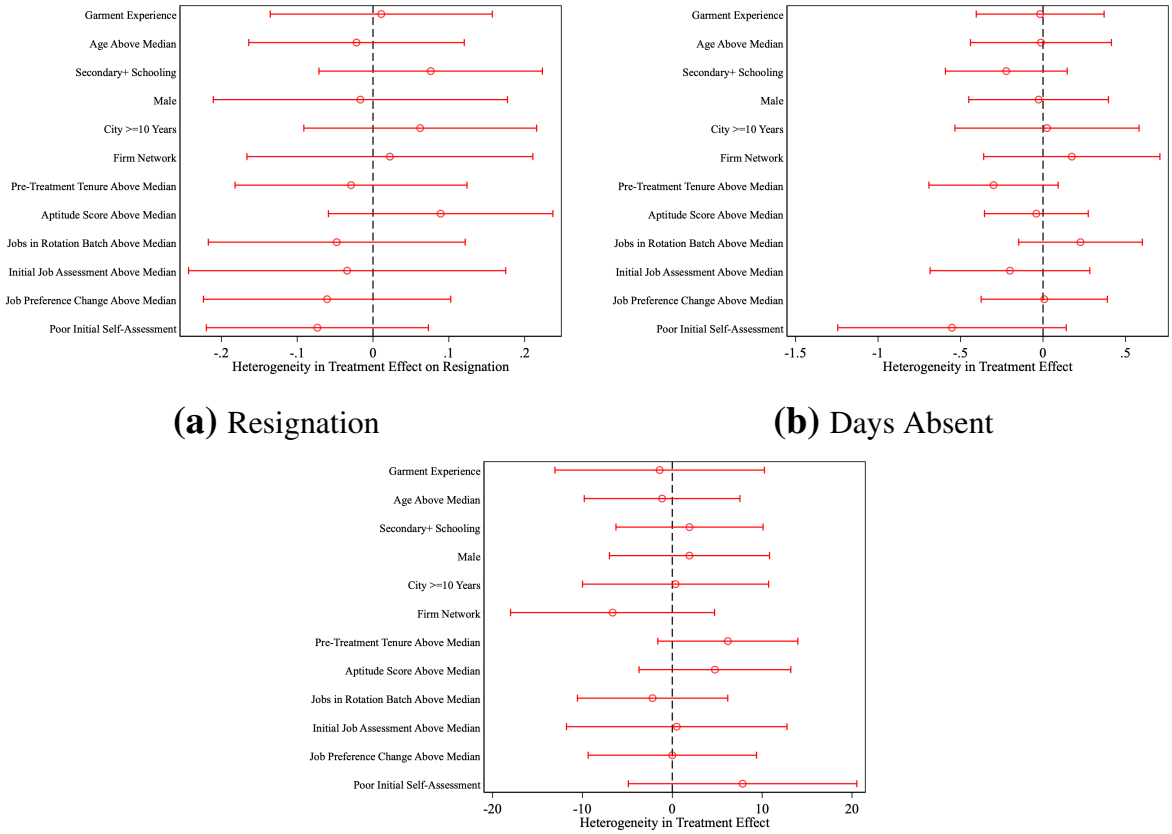
Figure A.3: Monthly Payroll Outcomes by Treatment Status



(a) Total Pay **(b)** Base Pay

(c) Incentive Pay **(d)** Overtime Pay

Note: Each panel plots average monthly outcomes relative to the batch date, separately for treatment (blue) and control (red) workers. Total Pay is total compensation; base pay reflects fixed wages; incentive pay captures performance-related compensation; and overtime pay reflects compensation for hours worked beyond regular schedules.



(a) Resignation

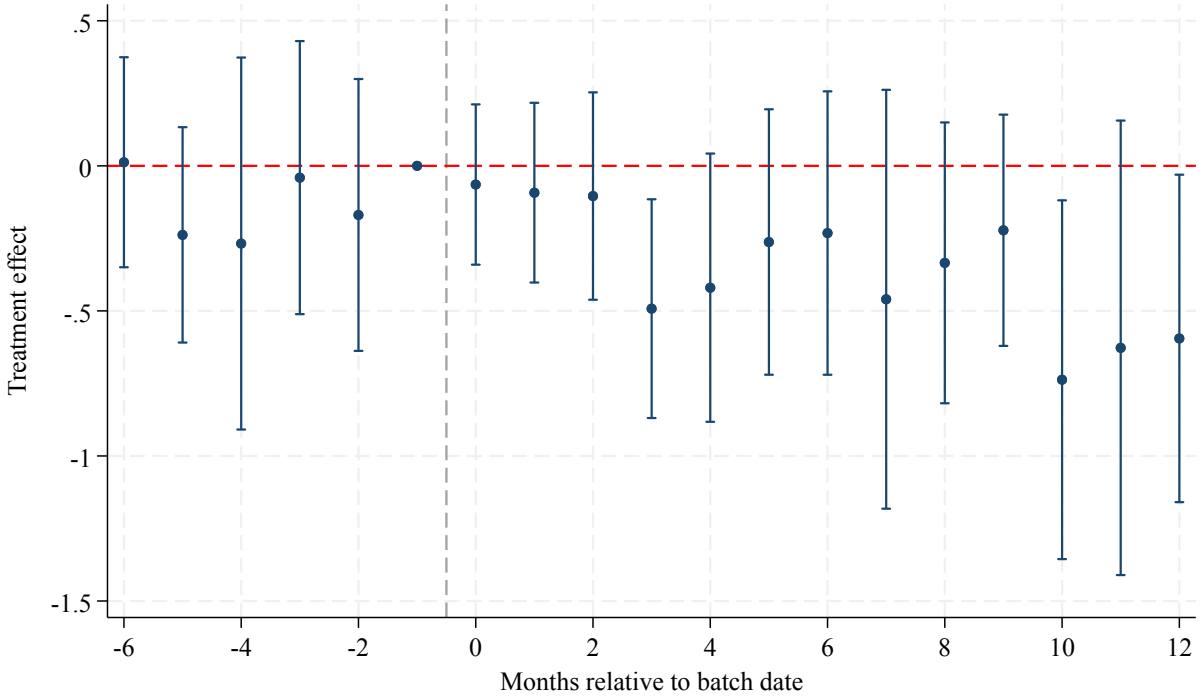
(b) Days Absent

(c) Total Pay

Figure A.4: Heterogeneity in Administrative Outcomes

Note: Each panel refers to a different outcome of interest. Each dot comes from a separate regression and represents the coefficient estimate for the interaction between each heterogeneity indicator and the intention-to-treat variable. The bands around each dot represent the 95% confidence intervals. All heterogeneity variables are binary indicators split at the full data median (greater than or equal to). Specifications include batch and year-month fixed effects. Standard errors are clustered at the batch level. Sample restricted to post-batch months only.

Figure A.5: Event Study: Absence Days by Treatment Status (Worker FE)



Note: Each panel plots the coefficient on Treat $\times$ event month, with 95% confidence intervals, from a regression with worker, batch, and calendar-month fixed effects (no separate tenure control – the worker fixed effect absorbs time-invariant tenure/demographic differences). The omitted reference period is one month before the batch date (rel_month $= -1$, dashed vertical line). Event time is capped at -6 and $+12$ months; the endpoint bins pool all worker-months beyond that window.

Table A.1: Compliance with Rotation and Matching

	N	Rate (%)
Panel A: Rotation Compliance		
Treated workers (scheduled, surveyed)	276	
Tried $\geq$ 1 scheduled job	235	85.1%
Tried all scheduled jobs	227	82.2%
Panel B: Matching Compliance		
Treated workers	312	
Recommended switch	61	19.6%
Accepted switch (of recommended)	22	36.1%

Notes: Panel A reports rotation compliance among treated workers with a valid midline survey: whether workers tried at least one of their scheduled rotation jobs, and whether they tried all of their scheduled rotation jobs. Rates are the share of treated workers (N in the first row) who tried $\geq$1 / all scheduled jobs, respectively. Panel B reports matching compliance among all treated workers: whether a job switch was recommended, and whether the recommended switch was accepted. "Recommended switch" rate is the share of treated workers; "Accepted switch" rate is the share of workers with a recommended switch who accepted it.

Table A.2: Baseline Summary Statistics (By Attrition)

	Attriters			Non-attriters			
	Mean	SD	N	Mean	SD	N	p-value
Tenure $>$ 1 Month at Batch Start (0/1)	0.34	0.48	111	0.55	0.50	494	0.00
Tenure (days) at Batch Start	43.03	92.37	111	101.99	281.07	494	0.03
Age	27.91	6.70	108	28.71	7.01	488	0.28
Male (0/1)	0.42	0.50	111	0.32	0.47	494	0.05
Schooling Secondary or Above (0/1)	0.74	0.44	102	0.70	0.46	473	0.48
Native Language:							
... Language 1	0.00	0.00	111	0.16	0.36	494	0.00
... Language 2	0.24	0.43	111	0.40	0.49	494	0.00
... Language 3	0.01	0.09	111	0.01	0.11	494	0.78
... Language 4	0.61	0.49	111	0.38	0.49	494	0.00
... Language 5	0.15	0.36	111	0.06	0.24	494	0.00
Live in the City for Over 10 Years (0/1)	0.66	0.48	106	0.46	0.50	481	0.00
Aptitude Test: Total Score	67.85	32.65	108	65.56	36.03	492	0.54
Sewing Skills (1-4; 1=None; 4=Advanced)	1.87	0.95	103	2.28	0.97	473	0.00
Has Firm Network	0.55	0.50	109	0.76	0.43	488	0.00
Prior Experience in Garment (0/1)	0.44	0.50	107	0.64	0.48	467	0.00
Years of Prior Experience in Garment	1.56	2.89	104	2.88	4.50	468	0.00

Note: The table shows the baseline summary statistics by whether the worker attrited by the endline surveys. The table is based on 605 workers who did baseline surveys. The p-value is from a regression testing for a statistical difference between the attrited and non-attrited groups.

Table A.3: Preference Change in Baseline and Endline Jobs

	Preference at Endline Survey					
	Mean	Pref=1	Pref=2	Pref=3	Pref=4	Pref=5
Panel A: Original Job						
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.106	-0.015	-0.041	0.013	0.051***	-0.008
	(0.075)	(0.045)	(0.046)	(0.032)	(0.017)	(0.009)
Baseline pref. (same job)	0.195***	-0.074***	0.009	0.034**	0.005	0.025
	(0.053)	(0.026)	(0.025)	(0.016)	(0.014)	(0.016)
Control mean	1.62	.589	.266	.105	.0161	.0242
Observations	497	497	497	497	497	497
Panel B: Endline Job						
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.084	0.009	-0.037	-0.029	0.051***	0.007
	(0.075)	(0.040)	(0.042)	(0.026)	(0.017)	(0.012)
Baseline pref. (same job)	0.277***	-0.107***	0.013	0.038**	0.038***	0.019
	(0.058)	(0.029)	(0.029)	(0.016)	(0.013)	(0.013)
Control mean	1.59	.593	.27	.105	.0161	.0161
Observations	496	496	496	496	496	496

Note: This table reports ANCOVA-style estimates of the effect of treatment on endline-survey job preferences, controlling for each worker's own baseline-survey preference for the same job, together with a missing-baseline indicator for workers whose baseline rating of that job is missing (the covariate is mean-imputed in those cases). Panel A reports results for the worker's original job; Panel B reports results for the job the worker is in at endline. In each panel, the "Mean" column regresses the raw 1–5 endline preference score (1 = most preferred) on treatment status; the remaining five columns instead use an indicator for the endline preference equaling each specific value (1 through 5) as the outcome. All columns include enumerator and batch fixed effects, with standard errors clustered by batch. $*$ p<0.1, $**$ p<0.05, $***$ p<0.01.

Table A.4: Heterogeneity Analysis: Feedback Sub-Items

	Promotion Not Limited	Supervisor Encourages Me	Workload Manageable	Supervisor Fair	Coworker Relations Good	Rules Not Burdensome	Rewards Sufficient	Job Can Be Done by Self	Low Chance of Accident	Free from Health Hazards	Encourage Friends to Apply
Treat	-0.044	-0.066	0.005	0.002	0.029	0.018	0.000	0.039	0.010	-0.093**	-0.017
	(0.043)	(0.060)	(0.034)	(0.030)	(0.030)	(0.028)	(0.043)	(0.043)	(0.054)	(0.035)	(0.039)
Treat × Poor initial performance	0.167**	0.047	0.082	-0.024	0.052	-0.025	0.199**	0.148	-0.037	0.192	-0.012
	(0.081)	(0.144)	(0.093)	(0.088)	(0.106)	(0.084)	(0.086)	(0.107)	(0.121)	(0.142)	(0.086)
Y mean	0.181	0.707	0.233	0.181	0.211	0.145	0.436	0.284	0.714	0.727	0.835
Observations	454	454	454	454	454	454	454	454	454	454	454

Notes: Each panel reports OLS estimates from a regression of the outcome on treatment status, a binary moderator, and their interaction, with batch and enumerator fixed effects. Standard errors clustered at the batch level in parentheses. All outcomes are binary indicators equal to one if the worker rates the item at or above the sample median (GEQ coding). Negatively-phrased items (Promotion Not Limited, Workload Manageable, Supervisor Fair, Coworker Relations Good, Rules Not Burdensome, Rewards Sufficient, and Job Can Be Done by Self) are reverse-coded so that, across every column, 1 corresponds to the more favorable response; see $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.5: Effect of Treatment on Various Leave Categories

	Sick	Annual	Personal
Treatment	-0.14***	-0.00	-0.05
	(0.05)	(0.02)	(0.03)
Control Mean (Post)	0.71	0.46	0.34
Observations	5770	5770	5770

Notes: Each column reports estimates from an OLS regression using worker-month observations in the post-treatment period. This table decomposes leave days by category. All specifications include batch and calendar-month fixed effects. Standard errors are clustered at the batch level. Control means are computed over the post-period for control workers. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.10$.

Table A.6: Effect of Treatment on Attendance and Payroll – Pre/Post Period, Worker Fixed Effects

	Panel A: Absence and Working Hours			
	Days Absent	Total Hours	Regular Hours	Overtime Hours
Treat x Post	-0.26**	1.79	0.72	1.07
	(0.12)	(2.83)	(2.33)	(0.75)
Control Mean (Post)	2.13	203.46	175.03	28.43
Observations	7267	7267	7267	7267
	Panel B: Payroll			
	Total Pay	Base Pay	Incentive	Overtime Pay
Treat x Post	2.94	0.95	0.66	1.38
	(3.70)	(2.76)	(0.46)	(1.41)
Control Mean (Post)	248.98	197.95	2.10	48.93
Observations	8513	8516	8513	8513

Note: Each column reports estimates from an OLS regression using worker-month observations across the full panel (pre- and post-batch, no period restriction). The coefficient shown is Treat $\times$ Post; a Post main effect is included but not shown. All specifications include worker, batch, and calendar-month fixed effects. Standard errors are clustered at the batch level. Control means are computed over control-group worker-months in the post period. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.10$.

Table A.7: Effect of Treatment on Workplace Violations

	Violations
Treatment	0.00
	(0.01)
Control Mean (Post)	0.04
Observations	6829

Notes: The table reports estimates from an OLS regression of the number of workplace violations on treatment status, using worker-month observations in the post-treatment period. The specification includes batch and calendar-month fixed effects. Standard errors are clustered at the batch level. The control mean is computed over the post-period for control workers. ***$p \leq 0.01$, **$p \leq 0.05$, *$p \leq 0.10$.

Table A.8: Stacked DiD: Team Productivity Outcomes by Treated & Control Worker Inflow — Alternate Thresholds

	(1) Output Minutes	(2) Efficiency
Panel A: Main specification (manpower ≥ 10, balanced window, non-overlapping rounds)		
Treated inflow x Post	2166.45**	28.53**
	(816.79)	(11.77)
Control inflow x Post	68.52	1.23
	(537.00)	(6.97)
Pre-period mean	8052.02	71.04
p-value: Treat = Control	0.07	0.09
Observations	833	833
Panel B: Manpower floor = 5		
Treated inflow x Post	2166.45**	28.53**
	(816.79)	(11.77)
Control inflow x Post	68.52	1.23
	(537.00)	(6.97)
Pre-period mean	8052.02	71.04
p-value: Treat = Control	0.07	0.09
Observations	833	833
Panel C: Manpower floor = 15		
Treated inflow x Post	2166.45***	28.53**
	(804.13)	(11.59)
Control inflow x Post	68.52	1.23
	(528.67)	(6.86)
Pre-period mean	8096.68	68.85
p-value: Treat = Control	0.06	0.09
Observations	807	807
Panel D: Manpower floor = 20		
Treated inflow x Post	2428.56***	22.80***
	(830.09)	(6.69)
Control inflow x Post	-103.49	2.79
	(580.06)	(5.41)
Pre-period mean	8128.64	66.66
p-value: Treat = Control	0.03	0.05
Observations	716	716
Panel E: Including overlapping rounds		
Treated inflow x Post	1031.82***	11.92**
	(379.39)	(4.72)
Control inflow x Post	-551.85	-1.94
	(439.08)	(5.03)
Pre-period mean	7762.95	67.19
p-value: Treat = Control	0.02	0.07
Observations	1397	1397

Notes: Stacked difference-in-differences at the line-round level, using a ±6-month event window around each rotation batch's start date. Output minutes (col. 1, output quantity × SAM) reflect realized production; efficiency (col. 2) is output minutes divided by total input minutes worked, expressed as a percentage. Treated inflow × Post" and Control inflow × Post" measure treated/control workers assigned to the line in that round as a share of the line's baseline manpower, rescaled so that a one-unit increase corresponds to a 10-percentage-point increase in worker inflow relative to baseline line size, and interacted with a post-round indicator. Baseline manpower is the median valid manpower reading for the line. All specifications include stack-line and stack-relative-month fixed effects; standard errors (in parentheses) are clustered at the team level. Panel A is the main specification: lines with baseline manpower below 10 are dropped, each stack-line must have productivity data for every month of the event window, and a batch/round is dropped if its window overlaps the previous kept round for that line. Panels B–D vary only the manpower floor (5, 15, 20), holding the other restrictions at their Panel A values. Panel E keeps every round regardless of overlap with an adjacent round on the same line rather than restricting to non-overlapping rounds only. Pre-period mean" is the outcome's mean in the estimation sample prior to the round. p-value: Treat = Control tests equality of the treated- and control-inflow coefficients. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Reasons Recommended Reassignments Were Not Implemented

Category	N	Share	Representative example
Factory staffing, production, or managerial constraint	13	42%	"Currently essential for Line 24's pants production . . . absence may negatively impact efficiency."
Worker declined or chose to remain	7	23%	"Workers don't want to move."
Worker- or job-specific fit or eligibility restriction	5	16%	"Worker is allergic to the chemical process in printing. Hence, cannot move."
Prior-rotation or procedural requirement	5	16%	"Never go to printing; manager wants someone who rotates to printing before."
Attrition	1	3%	"Both resign."
Total	31	100%	

Notes: Among the 57 recommended reassignments for which worker preference data are available, 18 were accepted and implemented, leaving 39 that were not. The table reports the 31 non-accepted recommendations for which contemporaneous implementation notes document a reason; the remaining 8 lack a documented reason. Categories are assigned by the authors based on the primary reason recorded in the notes. A small number of cases cite more than one reason over time and are classified according to the reason most proximate to the final non-implementation outcome. Shares are calculated among the 31 cases with a documented reason. Quotations are drawn directly from field implementation notes and anonymized to remove worker names.