Training, Communications Patterns, and Spillovers Inside Organizations*

Miguel Espinosa† and Christopher Stanton‡

July 11, 2025

Abstract

This paper examines how training affects productivity across hierarchical layers within organizations. After a randomized training program for frontline employees at a government agency, trained workers' output increased while their requests for managerial assistance fell. This freed managers to focus on strategic tasks – particularly managers with the strongest connections to trained employees. A structural model of organizational hierarchies shows that spillovers to managers account for approximately 45% of the program's total benefits, indicating that evaluations focused solely on individual trainees may substantially understate the full value of training investments.

*We thank Jerome Adda, Oriana Bandiera, Morten Bennedsen, Jordi Blanes-i-vidal, Ronald Burt, Lorenzo Caliendo, Alexia Delfino, Christian Dustmann, Sarah Eichmeyer, Florian Englmaier, Guido Friebel, Luis Garicano, Robert Gibbons, Mitchell Hoffman, Jonas Hjort, Scott Kominers, Hideo Owan, Tom Nicholas, Andrea Prat, Canice Prendergast, Eric Quintane, Imran Rasul, Ray Reagans, Esteban Rossi-Hansberg, Raffaella Sadun, Antoinette Schoar, Josh Schwartzstein, Kathryn Shaw, Pablo Spiller, Chad Syverson, Steve Tadelis, John van Reenen, Timothy Van Zandt and participants of the 14th UNSW workshop in Organizational Economics, 2021 NBER Organizational Economics, 2021 NBER Personnel Economics, 2021 CEPR IMO, ICESI, HBS, LMU, Michigan, Paris-Berkeley Workshop, EIEF, LSE, SIOE, Padova, Quantil, SMS, UTAH Winter Strategy Conference, Bocconi and ESADE for helpful comments. Jairo Galvis, Juan Herrera and Daniel Garzon provided superb research assistance. We thank the anonymous organization for giving us access to the data used in this study. Under the data use agreement, the organization had the right to review the draft for confidential information or details that would allow a reader to identify the organization.

†Bocconi University, CEPR & CESifo, miguel.espinosa@unibocconi.it

‡Harvard Business School and NBER, cstanton@hbs.edu

1 Introduction

Organizations frequently invest in employee training. Compared to research on direct productivity effects and firms' ability to capture gains (Bartel, 1995; Black and Lynch, 1996; Acemoglu and Pischke, 1998, 1999; Autor, 2001), there has been less focus on the organization-wide impacts of training. Models of organizational hierarchies emphasize that increasing lower-level workers' skills frees higher-level workers' time (Garicano, 2000; Garicano and Rossi-Hansberg, 2004; Caliendo and Rossi-Hansberg, 2012), suggesting spillovers across organizational layers could be valuable. Understanding these spillovers is essential when evaluating training investments in light of their importance for growth, agglomeration, and careers (Jones and Summers, 2020; Glaeser et al., 1992; Jarosch et al., 2021).[1]

This paper examines how a randomized training program for frontline workers affects individual productivity and generates spillovers to coworkers and managers. Our data come from a Colombian federal government agency with investigative and oversight responsibilities. Twelve percent of the agency's frontline workers were randomly selected to receive 120 hours of training in late 2018. The program was implemented in anticipation of increasing external demands requiring enhanced workforce capabilities. The curriculum mainly covered legal analysis, written communication, and specific topics for each participant's work, with smaller modules addressing computer skills, goal setting, and time management.

Our analysis draws on 12 weeks of productivity data from early 2018 (several months pre-training), and 12 weeks of data from early 2019 (several months post-training). Our

[1] One prominent view is that firms under-provide training (Cappelli, 2012). According to Training Industry Magazine, organizations' attempts to quantify training gains largely focus on individual outcomes, suggesting that firms may under-invest if spillovers are significant and difficult to quantify. See: https://trainingindustry.com/articles/measurement-and-analytics/how-to-identify-the-right-training-kpis-for-your-learning-and-development-programs-spon-eidesign/.

productivity measure is based on workers' weekly goal achievement, a common performance metric in the public sector (Rasul and Rogger, 2018; Rasul et al., 2018). An independent unit sets and evaluates goal achievement for all employees, including managers, and training did not formally affect goal setting or evaluation. To estimate spillovers, we leverage variation in connection strength between trained workers and both managers and untrained coworkers, which we measure from email counts between all worker pairs.

Our empirical approach begins with reduced form estimation, which yields three main findings. First, the program directly increased trained frontline workers' goal achievement by approximately 10% in the medium-term (4-6 months post-training), which is noteworthy given the mixed evidence on training effectiveness (Card et al., 2018). Untrained frontline workers' goal achievement remained relatively stable over the same period.

Despite challenges in capturing value from training (Becker, 1994), the study organization likely realized benefits from the individual gains. In particular, there is a fixed wage schedule, eliminating upward labor cost pressure, and trained workers were more likely than others to remain with the organization over the next 3 years. However, we note that the organization's unique positioning as a prestigious, high-paying government employer with fixed wages may limit the generalizability of the direct gains estimates in other settings.

Second, we find significant spillovers to managers, which we term *vertical spillovers* because they occur across the organizational hierarchy.[2] Importantly, managers are assigned their own "strategic" tasks, and their goal achievement and evaluation are not contingent on frontline workers' performance. In the raw data, managers' average goal achievement

[2] Goal achievement manipulation is also not likely to fully explain our direct returns results, as we would not expect to find spillovers to managers if trained workers' gains solely came from manipulation.

increased by 2.2 percentage points (3%) between the pre- and post-periods.

Managers with the strongest pre-period connections to trained workers experienced the largest productivity gains. Using an exposure design based on email counts with training participants, we can fully explain managers' 2.2 percentage point increase in goal achievement. We attribute about 1 percentage point of managers' gains to training spillovers when we use a more conservative measure of exposure based on the share of emails with participants.

Email data and surveys indicate that managers benefited because trained workers required less assistance. In pre-period data, managers' goal achievement is negatively correlated with emails from frontline workers, suggesting that responding to workers diverts time from their own tasks. Following training, emails from trained workers to managers declined substantially, allowing managers to focus on their strategic work.

Third, we find negligible spillover benefits for untrained frontline workers. Although we document that trained workers began providing help to untrained workers, the measured impact of these *horizontal spillovers* for untrained workers' individual goal achievement is small. However, help from trained workers functions as an informal layer of management that further freed existing managers' time.

After establishing these results, we develop and estimate an organizational hierarchy model that incorporates spillovers. We extend the canonical knowledge hierarchies models of Garicano (2000), Caliendo and Rossi-Hansberg (2012) Caliendo et al. (2015, 2020), and Gumpert et al. (2022) with two key modifications. First, we allow managers to perform their own strategic tasks rather than devoting their entire time to helping lower-level workers. Second, both lower-level workers who need assistance and higher-level workers who provide it must pay helping costs. In our framework, increasing frontline workers' skills reduces their time cost of seeking help, allowing them to take on additional tasks. Reduced help requests also create vertical spillovers by freeing managers' time for strategic tasks.

We structurally estimate the model to value spillovers relative to the direct returns. We use a GMM estimator to recover individual-level production parameters, including knowledge levels for each layer, helping costs, and wages. Given these input choices, we compute the organization's cost minimization problem to infer the relative importance of frontline workers' and managers' tasks in the production function.

Using the production function estimates, vertical spillovers to managers represent approximately 45% of the total program benefits when we assume that 1 percentage point of managers' goal achievement increase comes from spillovers (our conservative reduced form estimate). Given that managers are paid more than twice as much as frontline workers, it is not surprising that their tasks have about 75% more weight in the aggregate production function, yielding a substantial spillover value. Because of the importance of managers' tasks, the organization would have needed to train 1.9 times as many frontline workers to achieve the same output level in the absence of vertical spillovers.

These estimates suggest that measuring training gains from individual-level outcomes will not fully capture organization-level returns (Bartel, 2000).[3] In fact, we demonstrate theoretically that if trained workers transition from producing individually to helping others, their individual productivity metrics may even decline while the organization as a whole

[3] Our vertical spillover estimates complement studies using saturation designs to estimate peer effects at the same level (De Grip and Sauermann, 2012; Adhvaryu et al., 2023b). Other relevant papers are Levitt et al. (2013), who examine learning by doing cascades across workers, Sandvik et al. (2020), who run an experiment showing firms can generate knowledge spillovers by increasing contacts between coworkers, Sandvik et al. (ming) who examine selection into training, and Bandiera et al. (2007, 2009) who examine social incentives between coworkers and managers. Kugler et al. (2022) estimate training spillovers to workers' relatives, providing a wedge between social and private training returns.

benefits. Despite our unique setting, the spillover estimates are likely more generalizable than estimates of direct returns, suggesting that training can improve organization-wide performance and government quality (Rasul and Rogger, 2018; Bandiera et al., 2021).

Our findings are also relevant for several streams of literature in organizational economics. First, we extend the empirical literature on hierarchies and organizational structure, which has typically leveraged across-firm variation in market size/demand or changes in communications infrastructure (Garicano and Hubbard, 2003). For example, Caliendo et al. (2015) study firm growth, showing that most French manufacturing firms exhibit headcount, wages, and hours patterns consistent with hierarchical structures. Growing firms add layers, while wages decline in lower layers – consistent with task substitution as firms scale. These relationships are causal – with distinct implications for quantity-based and revenue-based TFP (Caliendo et al., 2020), can generate intra-firm inequality across layers (Friedrich, 2022), and also hold for the services sector (Garicano and Hubbard, 2007, 2009). Changes in technology (Bloom et al., 2014) or infrastructure Gumpert et al. (2022) alter communication costs and organization, leading to equilibrium adjustment even absent demand changes.

Our insider approach with one organization extends the literature in several ways.[4] First, we can distinguish between different theoretical models of hierarchy, which is difficult in most settings because managerial spans of control change with demand or communications technology shifts (Chen and Suen, 2019). In monitoring hierarchies, spans of control determine the likelihood of detecting shirking, while in knowledge hierarchies spans of control respond to question volume from lower-level workers. Because the agency didn't change headcount

[4] The closest work looking inside one firm is from Adhvaryu et al. (2023a), who study how training automobile line workers to accommodate product shifts results in delayering – the rationale being that higher-level workers need to be closer to the line to provide rapid help when the firm is undergoing change.

or formal structures, our evidence supports the knowledge hierarchy interpretation: training increased workers' goal achievement and reduced their questions to managers without altering monitoring. In addition, we show that slight modifications to canonical hierarchies models can fit relatively complex organizational structures where workers in higher layers simultaneously handle support functions and their own directly assigned tasks. Finally, in contrast to the common focus on top-down effects of bosses on subordinates and organizations (Lazear et al., 2015; Bloom et al., 2015, 2020; Hoffman and Tadelis, 2021; Bianchi and Giorcelli, 2022), we show how bottom-up skills changes can benefit managers.

2 The Setting and Empirical Strategy

We study a randomized training intervention at a Colombian federal government agency. A confidentiality agreement prevents us from naming it, but it is a prestigious oversight, inspection, or investigative institution within the federal government. Federal positions in Colombia generally offer competitive compensation, and this agency pays well relative to other public entities. Frontline workers earn wages at the 78th percentile of service-sector workers nationally.[5]

The "core" part of the organization has 5 divisions, and workers perform different functions in each division. For example, the "Execution Division" (36.9% of employees) answers citizen requests, conducts investigations, and issues findings for disciplinary proceedings.[6] Workers joining the agency are assigned to a division and placed within one of five wage bands based on their educational qualifications and prior government sector experience. We categorize employees in wage bands 1 and 2 as "frontline workers" – typically high school

[5]Based on analysis of Colombian Great Integrated Household Survey data.

[6]Other divisions are Administration (19.3% of employees), Finance (13.7%), Human Talent (14.9%), and Planning (14.9%).

graduates or those with bachelor's degrees. Those in wage bands 3 through 5 are "managers" in our analysis and generally hold bachelor's, master's, or doctoral degrees. Compensation follows a strict schedule based on wage band, occupation title, and public sector experience. As in many public sector settings, earnings are not linked to short-term, individual performance metrics (Finan et al., 2017; Miller, 2018) and attrition rates are very low (Grissom et al., 2016; Hur and Abner, 2022), but employees can be terminated for long-term poor performance.[7]

2.1 Goal Achievement and Evaluation

Our primary productivity measure is goal achievement, which refers to weekly performance evaluation on assigned tasks. An independent oversight group with limited interactions with employees sets and assesses goals.[8] Due to this separation, the organization's leadership believes that goal setting and evaluation did not consider workers' training status.

Managers are assigned their own tasks, which are evaluated by the oversight group. These tasks, which we refer to as strategic, are distinct from frontline workers'. Although managers may assist frontline workers and authorize actions beyond workers' responsibilities, managers are evaluated solely based on their strategic responsibilities, not on the success of employees they support.

While our dataset does not include the content of individual goals, we have access to workers' weekly goal achievement scores and were given example goals in interviews. For example, frontline workers in the Execution Division typically focus on advancing specific

[7]Only two workers left during our main sample, one untrained frontline worker and one manager. Limited hiring and turnover outside of government changes is common in Colombian government agencies.

[8]The oversight body is familiar with workers' tasks and contains specialists to set and measure task completion. There are period meetings between the oversight body and division-level managers to discuss overall objectives, after which work is assigned to individuals.

cases or investigations, while managers are evaluated on higher-level tasks such as filing case audit reports, planning and coordinating investigations, and developing contingency plans when case progress deviates from expectations. Our primary productivity measure is the weekly aggregate goal achievement score, which is reported on a 0–100 scale and which we re-scale to range from 0 to 1, representing each worker's achievement as a fraction of the maximum evaluation score.[9]

2.2 Training Program

Organizational leaders anticipated that the future case mix would become less routine and case volume would increase. In July 2018, the organization announced that a 16-week frontline employee training program would run from August through December 2018. Budget considerations limited participation to 63 employees. The selection was determined by lottery among frontline workers in wage bands 1 and 2. All employees were informed of the selection process and were made aware that no other similar training programs were planned for the future.

The training was delivered through three full-day sessions each month, totaling 120 instruction hours. The curriculum covered four general skills components and one division-specific module. The general skills components included: (i) goal setting, scheduling, and time management; (ii) Microsoft Excel proficiency; (iii) Colombian legal analysis; and (iv) effective written communication for legal documentation. The division-specific modules were tailored to employees' work. For example, Finance division workers focused on banking and public finance principles, while the Execution division received training in national and inter-

[9] Goal evaluation is based on a weighted average of four components, with the two most important being how well targets are met and how efficiently resources are used.

national law to fit the anticipated future case mix. More time was devoted to legal analysis, written communication, and division-specific modules than to the time management and Excel components.

2.3 Summary Statistics and Treatment Assignment

Our dataset covers 655 employees over two distinct twelve-week periods from April through June in both 2018 (the pre-period) and 2019 (the post-period). The dataset includes weekly goal achievement metrics for individual workers and managers linked to employee characteristics including gender, education, wage band, and division. We also have email metadata documenting daily bilateral communication patterns between all employees during the same periods.

Table 1 presents descriptive statistics and examines balance in treatment assignment. Columns 1-3 display statistics split between frontline workers (n=526) and managers (n=129). To understand differences among frontline workers, Columns 1 and 2 provide separate statistics for wage band 1 and 2 workers. Frontline workers across wage bands are relatively similar to one another, though some differences exist. For example, 50% of wage band 2 workers hold a bachelor's degree or higher, compared to 27% of wage band 1. By contrast, all managers have at least a bachelor's degree, with 36% holding a master's or PhD. Managers earn, on average, 2 times as much as workers in wage band 2. We group wage band 1 and 2 workers in subsequent analyses due to their demographic similarity, their eligibility for the training program, and their assignment to similar tasks.[10]

Columns 4-5 split the data by frontline workers' training status. Column 6 presents a

[10] Later we will show that workers in both frontline wage bands turn directly to managers for help, suggesting they are in the same layer of the organizational hierarchy.

balancing test of random assignment to training based on worker characteristics. Although two characteristics show some imbalance, the F-statistic from a joint test (F=0.885, p=0.547) fails to reject the null of no systematic differences between treatment and control groups.

In evaluating the organization's returns, Becker (1964) suggests examining whether workers capture value through higher wages or exits to external opportunities. There are no differential wage increases for training participants during the post-period (Columns 4 and 5). Organizational leadership confirmed that this pattern persisted, with no systematic wage premiums for trained workers except in cases where workers received temporary promotions.[11] To assess effects on retention, we tracked whether frontline workers remained with the organization through December 2022. Fourteen of the 526 frontline workers left, none of whom participated in the training program. Retention data for managers was not available.

The final rows of Table 1 present preliminary evidence of the training program's effects. Among trained workers, average goal achievement increased from 71.9% in the pre-training period to 78.5% in the post-training period—a 6.6 percentage point (or approximately 10%) improvement. In contrast, goal achievement for untrained workers remained essentially stable, averaging 72.6% in the pre-period and 72.1% in the post-period. Of particular interest, managers (Column 3) experienced a 2.2 percentage point (3%) improvement in goal achievement, moving from 70.8% to 73.0%. This motivates our analyses of vertical spillovers.

[11] Temporary promotions backfill turnover in higher layers or cover absences. These positions come with higher wages but do not represent a net cost of the program, as the organization fills these slots internally by rule. As in many other public entities, permanent job positions are filled through a national-level call for applications. Candidates must meet specific requirements, including excelling in a public sector knowledge test. These hiring waves are costly and occur infrequently. Temporary promotions may last only a short time or remain in place until a new hire fills the position.

2.4 Identifying Spillovers from Email Data on Connections to Trained Workers

We use email data to capture connections between workers that may drive same-level (horizontal) and across-level (vertical) spillovers from training and to understand how communications patterns change after training.[12]

Our reduced-form identification strategy uses two connection measures, calculated from pre-period email data: (1) the level or count of emails with program participants, and (2) the share of emails with program participants relative to all emails with frontline workers.[13] Because our spillover mechanism involves help requests, we focus on emails that managers receive from eventually trained workers and emails that untrained workers send to trainees.

There is substantial variation in exposure. On average, managers received 1,670 emails from eventually trained workers during the pre-period (SD=893, or 54% of the mean). Their average email share with eventually trained workers was 12.1% (SD=2%, or 16.5% of the mean). Although correlated, the two measures capture different aspects of interaction. Some managers work extensively with frontline workers in general, resulting in high exposure levels. The share measure accounts for overall communication volume, so variation comes from idiosyncratic connection strength with trained workers. Moving to our horizontal connections measures, untrained workers send an average of 668 emails to eventually trained colleagues during the pre-period (SD=376), and approximately 12% of untrained workers' emails were with eventually trained workers (SD=3.4%).

[12] We do not observe email content and we cannot distinguish message threads (initial contacts versus replies). We also cannot observe whether messages were sent to individuals or multiple recipients. We use surveys to confirm that emails are a reliable proxy for total communication, as they complement other forms of interaction. In addition, surveyed frontline workers reported that 74% of their email requests to workers in higher layers involve asking for help (see Figure A1).

[13] Table A1 provides details about our email-based connection measures and correlations between alternative metrics.

Because all managers and untrained workers have at least some connection with program participants, we focus on a transformed variable that captures relative differences in connection strength. We define exposure as:

$$Exposure = \frac{\chi - \min \chi}{\max(\chi - \min(\chi))}, \tag{1}$$

where χ is a measure of pre-period emails (levels or shares) with eventually trained workers. This yields a 0-1 measure, where the most exposed manager (or untrained co-worker) takes a value 1 and the least exposed takes a value of 0. Figure A2 displays the distribution of exposure to trained workers in levels (Panel A) and shares (Panel B).

This identification approach assumes that connections between workers would have remained stable in the absence of the training program. Although direct tests are challenging because communications changes are a mechanism for our results, email patterns in the pre-period are highly persistent and suggest stable relationships (see Appendix Figure A3). If our measures reflected only temporary connections, our estimation approach would likely fail to detect spillovers through the organization.

However, there are two potential biases that the exposure design might introduce. First, the email levels measure is potentially vulnerable to transitory fluctuations in workloads. If a manager experienced unusually high communication volumes during the pre-period, the levels measure could erroneously attribute mean reversion in help requests to training effects, overstating vertical spillovers. Second, although share-based measures normalize away the total communications workload, they may suffer from mean reversion in relative connection strength, biasing spillover estimates to zero from measurement error. Thus, the most likely sources of potential bias go in opposite directions for our different exposure measures, making their combination useful for triangulating around the true parameter.

3 Stylized Facts and Reduced Form Estimates

We now turn to analyzing the program's effects. We organize the empirical findings around four key stylized facts that motivate and inform our structural valuation analysis.

3.1 Fact 1: Training Increased Workers' Goal Achievement

Figure 1 shows how goal achievement changes vary by training status. The figure shows a clear upward shift for trained workers, representing an approximate 10 percent increase.[14] The positive shift occurs across the support of the pre-period productivity distribution. Moreover, the program appears to be randomly assigned, as the pre-period density of goal achievement is similar for trained and untrained workers (as each data point represents an equally sized bin).

We quantify uncertainty and probe robustness to controls using a difference-in-differences regression of the form:

$$\log(y_{it}) = \beta_i + \beta_t \times \beta_d + \delta_1 Trained_i \times Post_t + \delta_2 Trained_i \times Post_t \times X_i + \varepsilon_{it}. \qquad (2)$$

The main coefficient of interest is δ_1. In some specifications, we include δ_2 to capture treatment effect heterogeneity based on characteristics, X_i. The model with the most controls includes individual fixed effects (β_i) and time-by-division fixed effects ($\beta_t \times \beta_d$). We cluster standard errors by worker.

Table 2 presents the results. Because the dependent variable is log goal achievement, the 0.105 coefficient on *Trained x Post* indicates that trained workers' goal achievement increased

[14] Figure 1 contains distinct clusters of goal achievement scores, which arise from rounding of the underlying measures that are aggregated into the final score. The plot is similar when we include division fixed effects, suggesting the lumpiness is not due to unobserved differences in functions or tasks. Changes in trained workers' goal achievement are similar for the Execution division and others, suggesting the gains are unlikely to be driven by the changing nature of legal cases.

by about 11 percent. Columns 3 and 4 add interactions with wage band, education, gender, and prior performance, which are reported in Appendix Table A2. Differences by wage band are most relevant, as we later examine how trained wage band 2 workers assist untrained workers. There is no differential effect of training on wage band 2 workers in Column 3, but the point estimate is negative with division-by-time fixed effects in Column 4. However, as Columns 5 and 6 show, no characteristics that determine treatment effect heterogeneity survive a LASSO variable selection procedure. Only the main treatment effect remains.[15]

We do not have data on longer-term goal achievement, but we can track temporary promotions for 3 years after the program. Over this time, trained workers were more likely to receive short-term promotions into temporarily open positions (57% compared to 28%), suggesting the program lifted long-term performance (see regression results in Table A3).

Robustness: The potential outcomes framework underlying equation (2) stipulates that post-period log goal achievement for untreated workers equals $\beta_i + \beta_t \times \beta_d$. This imposes the stable unit treatment value assumption (SUTVA) that there are no spillovers to untreated workers. We have also estimated models, following De Grip and Sauermann (2012), that allow untrained workers' goal achievement to deviate from $\beta_i + \beta_t \times \beta_d$. SUTVA violations do not alter our estimates of direct returns. Appendix Table A4 details the approach and shows that our estimates of direct returns are larger than those in Table 2. No variables that account for spillovers to untrained workers survive LASSO variable selection procedures.

[15] The LASSO (Least Absolute Shrinkage and Selection Operator) (Tibshirani, 1996) is a regularization procedure that entails variable selection. It penalizes the regression model, potentially setting some coefficients to zero. The LASSO is often used to discriminate between which variables enter a final model, and Zhao and Yu (2006) provide conditions for when the procedure selects the true model. We use the plugin penalty from Belloni et al. (2016) that accounts for clustering, as detailed in the table notes.

3.2 Fact 2: Managers Most Connected to Trained Workers Had the Largest Goal Achievement Increases

Figure 2 Panel (a) shows the evolution of managers' goal achievement for the top and bottom quartile of connections to trained frontline workers (based on email levels). Managers in the top quartile have a noticeable improvement in post-period goal achievement, averaging 5.5 percentage points or roughly 8% of the mean. Post-period goal achievement for bottom quartile managers is nearly identical to the pre-period. Notably, there are both high- and low-exposure managers across the range of pre-period goal achievement scores. This pattern supports our identification strategy, which treats exposure to trained workers as effectively random.

We estimate the following regression for the full sample of managers:

$$\log(y_{it}) = \beta_i + \beta_t \times \beta_d + \delta_m Exposure_i \times Post_t + \varepsilon_{it}. \tag{3}$$

The variable $Exposure_i$ is manager i's exposure measure, based either on the level or share of pre-period emails with eventually trained frontline workers. Because $Exposure$ ranges from 0 to 1, δ_m can be interpreted as the relative effect of moving from the least to the most exposed manager. β_i is an individual fixed effect, β_t is a time fixed effect, and β_d is a division fixed effect. The identifying assumption is that $Exposure_i$ is uncorrelated with managers' unobserved productivity trends, which is likely to hold if exposure is approximately random across managers.[16]

We lack a control group of completely unexposed managers, but we can interpret $\delta_m \times \overline{Exposure}$ as a lower bound on the average spillover effect under two conditions: i) non-

[16] Figure A4 examines the correlation of manager characteristics with their exposure to trained workers, showing that managers' pre-period goal achievement, wages, wage band, gender, and education do not predict pre-period connections with eventually trained workers.

negativity – no manager is made worse off because a connected frontline worker receives training, and ii) monotonicity – expected treatment effects are increasing in the degree of exposure to trained workers.[17]

Table 3 displays results using the level- (Columns 1-3) and share-based (Columns 4-6) exposure measures. Across all estimates, managers who have stronger pre-period connections with eventually trained workers have greater productivity gains. For the email level exposure measure, the point estimates range from 0.104 to 0.177, implying average goal achievement gains of between 3 and 5 percentage points given that mean exposure in levels is 0.4 (See Table A1). These estimates imply post-period average goal achievement for managers ranging between 0.738 and 0.76, whereas the actual average is 0.73. The estimates using the share measure are also positive. The average effects range from a 0.8 to a 1.6 percentage point increase in manager goal achievement given that the mean of the email shares exposure measure is 0.47. The difference in estimates across the two measures suggests that managers who were busiest providing help (captured in levels) benefited most. However, if transitory workloads in the pre-period mean-revert, the levels-based exposure measures will likely be biased upward. Yet if *relative* connection strength mean reverts, then the estimates based on shares are biased downward. By using connections measures with different potential biases, the true effect on managers' goal achievement likely lies between our estimates.

In what follows, we often assume the true effect on managers' overall goal achievement totals about 1 percentage point, which is relatively conservative. Still, under this conservative estimate, the importance of spillovers to managers is substantial.

[17] Figure A5 shows the monotonicity assumption appears to hold when the exposure measure is based on email levels or when it is based on email shares and division -by- time fixed effects are included in the model.

Robustness: We have probed the robustness to a variety of different controls, specifications, and variable selection techniques. Table A5 shows that the results are not sensitive to: i) controls for horizontal spillovers to untrained workers; ii) controls for connections with untrained workers and their second-order connections with trained workers and managers; iii) controls for mean-reversion, allowing initial productivity to evolve differentially across deciles of pre-period goal achievement; iv) controls for imbalance of trained workers' characteristics; v) controls for both emails sent and received from trainees;[18] vi) controls for shifts in communication networks and manager workload by including pre-period email volume and contemporaneous emails with untrained workers;[19] vii) winsorization of the exposure measure at the 90th and 10th percentiles; and viii) letting managers in different wage bands have different trends.

Managers' goal achievement is not linked to frontline worker performance: Managers' have their own goals. In Appendix Table A7, we show there is little correlation between contemporaneous manager goal achievement and the goal achievement of connected workers after we control for help requests via emails.[20] Managers' goal achievement is negatively related to weekly email volume from frontline workers, with elasticities around -0.09. However, managers' goal achievement is not responsive to connected workers' goal achieve-

[18] In Appendix Table A6 we explore alternative definitions of manager exposure to trained workers by including measures based on emails sent from managers to workers. A LASSO procedure selects only Exposure based on emails that managers receive.

[19] Figure A6 explores whether managers who are more connected to program participants receive a different number of emails from untrained workers in the post-period, which would violate the SUTVA. Although managers with greater connections to trained workers receive slightly more emails from untrained workers, the effect size is small and the regression coefficient is imprecisely estimated. We find no evidence of changes in communication patterns among managers in the post-period (see Figure A7).

[20] Establishing that managers' goal achievement is not linked to workers' is difficult because managers are expected to become more productive if their connected workers need little help, motivating why controls for help requests are required for this test. Our test regresses manager log goal achievement in the pre-period on emails received and connected workers' goal achievement, where connections are calculated based on emails in weeks other than the focal one.

ment. The lack of correlation helps to validate that workers and managers are evaluated on separate, distinct goals and that our estimates of vertical spillovers are not mechanical.

3.3 Fact 3: Horizontal Spillovers to Untrained Frontline Workers Are Limited

Returning to Figure 2, Panel (b) plots changes in goal achievement for untrained workers who are in the top and bottom quartiles of connections with trained workers. Unlike the gap observed among managers, the difference in goal achievement between more and less connected untrained workers is small. Regressions confirm that horizontal spillovers to untrained frontline workers are relatively limited. Using a regression like the one for vertical spillovers, we find point estimates ranging from 0.029 to 0.039, none of which are statistically significantly different from zero (see Appendix Table A8).

3.4 Fact 4: Email Patterns Are Consistent With Knowledge Hierarchies Models

After the program, trained workers' emails to managers fell, while untrained workers' began to email trained workers more frequently. Figure 3 shows changes in log emails between the pre- and post-periods, broken down by recipient type and sender (purple for untrained workers and light green for trained workers). While overall email volumes declined year-over-year, the reduction was particularly pronounced for emails from trained workers to managers, which decreased by 74.4% overall and by 61.6% after netting out the expected email reduction that would have occurred in the absence of training.[21] This pattern is consistent with models of knowledge hierarchies, as skill improvements reduce workers' need to get help from upper-layers.

Trained frontline workers also became more central within the organization's communi-

[21] We estimate the expected reduction in emails using changes in emails sent to managers by untrained workers in wage band 2. These workers' emails with trained workers do not increase, suggesting they do not get help from program participants. Appendix Table A9 presents the full set of email changes between different sender and receiver pairs.

cation network, suggesting they emerged as an informal helping layer for untrained workers. In particular, trained wage band 2 workers received substantially more emails from untrained wage band 1 workers, while untrained wage band 2 workers did not (see Appendix Figure B3). Even though trained workers' help had little effect on untrained workers' productivity, this informal support role likely benefited managers by further reducing their assistance burden. We explicitly account for this possibility in the theoretical framework and structural estimation of the model.

Email patterns also inform our modeling choices, supporting the decision to combine wage bands 1 and 2 and to abstract from pre-training peer communication (for analysis that treats wage bands 1 and 2 separately, see results in Appendix B.) What is particularly important for models of hierarchy is that both wage band 1 and 2 workers ask managers for help directly, which we confirm in Appendix Figure A8. In fact, relative to wage band 2, wage band 1 workers send a greater share of their emails to Managers. If wage band 2 were the first source of help for wage band 1, we would expect the opposite pattern.

Although frontline workers did frequently email one another in the pre-period (the most common email destination was another frontline worker), these emails were likely about coordination or project-related administration rather than help. There is very little correlation between emails with other frontline workers and goal achievement, which we would expect under a helping role (Figure A9 Panel (a)). Given the lack of correlation between frontline workers' emails and their productivity in the pre-period, we abstract away from other rationales for horizontal communication. By contrast, there is a negative relationship between frontline workers' pre-period goal achievement and email volume sent to managers, which will motivate how we model productivity (see Figure A9 Panel (b)).

4 A Model of Hierarchies and Spillovers

To quantify the relative importance of direct gains versus spillovers, we develop a structural model that builds on the knowledge hierarchies literature but adapts it to organizations where managers have their own strategic tasks. Our approach builds on Garicano (2000) and Caliendo and Rossi-Hansberg (2012) while adopting the cumulative knowledge setup in Gumpert et al. (2022). Appendix C.1 provides additional details about the baseline model and explains why the modifications we describe below are needed to fit the goal achievement patterns from the reduced form results. We estimate the model using generalized method of moments (GMM) and then infer the value of frontline workers' tasks relative managers' from the first-order conditions of the organization's cost minimization problem.

4.1 Model Setup

We study an organization that is initially structured as a 2-layer hierarchy. The organization has n_1 first-layer frontline workers and n_m managers. We denote workers' and managers' knowledge levels as z_1 and z_m, respectively. Workers draw problems with difficulty z from an exponential distribution with parameter $\lambda > 0$ and work on them sequentially. Workers can solve a problem autonomously using 1 unit of time if $z \leq z_1$, which occurs with probability $1 - e^{-\lambda z_1}$. If $z > z_1$, the worker asks a manager for help. Problems that require managerial help take $1 + H_m$ units of worker time and h_m units of manager time. Frontline workers' goal achievement, $\phi = \frac{1-e^{-\lambda z_m}}{1+H_m e^{-\lambda z_1}}$, is a renewal process where the denominator is the expected time a worker needs to process a task, accounting for the probability of needing managerial help. The numerator is the fraction of problems that can be resolved, $1 - e^{-\lambda z_m}$, which is determined by managers' knowledge.[22]

Each manager has one unit of time to allocate between their own strategic tasks (denoted S) and assisting workers who request help. We define a manager's goal achievement as $S = 1 - \frac{n_1}{n_m} h_m \frac{e^{-\lambda z_1}}{1+H_m e^{-\lambda z_1}}$, the time left over for strategy after allocating $\frac{n_1}{n_m} h_m \frac{e^{-\lambda z_1}}{1+H_m e^{-\lambda z_1}}$ units of time helping (assuming help requests are spread across managers equally, $e^{-\lambda z_1}$ is the probability a worker with skill z_1 needs help on a problem, and $1/(1 + H_m e^{-\lambda z_1})$ is the number of problems workers address in a time period). Final output, Q, depends on both the production tasks completed by frontline workers and the strategic tasks done by managers.

4.1.1 Effects of Training a Subset of Frontline Workers

To illustrate how the model can fit the reduced form stylized facts, we examine how output and help requests change when some workers receive training. We consider team-level production with 1 manager, n_1 frontline workers (here n_1 is used as workers/manager), and output for the team $Q = (n_1 \phi)^a (S)^b$. The experiment can be thought of as inducing some manager-worker groups to be exposed to trainees while others are not.

We analyze the organization's cost minimization problem in the next section. For now, we assume that the optimal inputs for a team are $\{n_1^*, z_1^*, z_m^*\}$, and we assume that 1 frontline worker gets trained and has knowledge $z_1^t > z_1^*$. We examine two cases. In the first, the trained worker only works on his own tasks. In the second, the trained worker both works on his own tasks and provides help to untrained workers. The share of untrained workers' help requests handled by the trained worker is given by an exogenous parameter, ρ. The two propositions below illustrate when it is better for the organization (and for untrained

[22] Using the standard approach in this literature, we treat workers' output as non-stochastic (which can be justified by assuming that workers draw a large number of tasks or that the organization can balance work across a large number of workers and managers).

workers) for the trained worker to provide help.

Proposition 1 (Only Managers Help)**.** *Consider training one frontline worker and hold fixed n_1^*, z_1^*, and z_m^*. Trained and untrained workers only ask questions to managers, and helping costs are symmetric, with $h_m = H_m$. The following results hold: 1) Total manager time spent on strategic tasks increases. 2) The trained worker's goal achievement increases (solving more production tasks). 3) Untrained workers' production does not change.*[23] *4) Total output increases.*

Proof: See Appendix C.2. The intuition is straightforward and shows how our model captures the main reduced form results. As frontline workers' skills improve, so too does their goal achievement as $\frac{\partial \phi}{\partial z_1} > 0$. While we do not model incentives, to the extent that asking for help is costly, workers' total effort may change little, while goal achievement increases. In addition, help requests to managers fall, increasing their time on strategic tasks and leading to a vertical spillover.

The next result suggests that if training leads to a sufficiently large knowledge increase, the organization may benefit when trained workers form a new helping layer between untrained workers and managers. To analyze this case, we assume that untrained workers pay time cost H_t when getting help from a trained worker and the responding trained worker pays h_t. If the problem then gets escalated to the manager, additional time costs of H_m and h_m are paid by the untrained worker and manager, respectively.

Proposition 2 (Trained Workers Help)**.** *Consider training one frontline worker and hold*

[23]This result for untrained workers only holds in a version of the model without congestion in queuing for help. The model with congestion in getting help (which we take to the data) can generate spillovers to untrained workers even if trained workers do not provide help because training makes them less likely to add to the queue to get managerial assistance.

fixed n_1^*, z_1^*, and z_m^*. The manager helps the trained frontline worker. The trained worker helps untrained workers with exogenous probability $\rho \geq 0$, and the untrained worker will, in turn, ask the manager for help if the problem is beyond the trained worker's capabilities.[24] *Assume helping costs are symmetric ($h_m = H_m$ and $h_t = H_t$). The following results hold: 1) Relative to the case where trained workers do not provide help (Proposition 1), managers spend more time on strategic tasks for any $\rho > 0$. 2) The trained worker solves weakly fewer production tasks relative to the case where he does not provide help. For ρ sufficiently large, the trained worker may solve fewer production tasks than untrained workers, implying individual production measures do not capture workers' skills. 3) Moving from $\rho = 0$ to $\rho > 0$ only increases untrained workers' production if the time required to get help from the trained worker is sufficiently small relative to requesting help from a manager, $H_t < H_m$, and the probability the trained worker can solve the problem is sufficiently large. 4) Relative to the case where the trained worker does not provide help, changes in total output are ambiguous.*

Proof: See Appendix C.2.

Figure 4 presents three scenarios illustrating when it is beneficial to use trained workers as a helping layer. In the first scenario, the trained worker requires less managerial help and uses the additional time to help untrained workers, subject to the constraint that his own production does not fall. In the second scenario, the trained worker's helping can reduce his own production; for illustrative purposes, we cap the trained worker's goal achievement reduction at −2%. In the third scenario, the trained worker does not provide help. Each panel displays different output changes as a function of the trained worker's knowledge.

[24]This assumes that the trained worker's knowledge does not overtake the manager's knowledge, i.e. $z_1^t < z_m^*$. The trained worker's helping time must respect their own time constraint, but if $h_t \leq h_m$ and the original allocation is feasible, then the trained worker's time constraint will always be satisfied even if $\rho = 1$.

Panel (a) traces out the change in output for the trained worker. Panel (b) illustrates how the model can capture horizontal spillovers to untrained workers, while showing that the sign of these spillovers is ambiguous for untrained workers' individual productivity. The figure has a negative spillover when the trained worker helps, but it can be positive.[25] Panel (c) illustrates vertical spillovers to managers, which are positive and increase when trained workers provide help. Panel (d) considers the combined organizational impact under a Cobb-Douglas production function. The figure shows that for small increases in z_1^t, output is highest when the trained worker concentrates solely on production. For larger increases in z_1^t, there is a threshold after which providing help raises total output even as individual production falls.

Remark 1 (Calculating Benefits from Training): Trainees' individual productivity is not a sufficient statistic for the returns to an organization. Focusing on individual output ignores vertical spillovers to managers. In addition, when trainees help untrained workers, their own output can stay flat or even fall due to helping time commitments while overall organizational performance rises.

In our empirical implementation, we introduce congestion effects. With congestion, if multiple workers queue for a manager's help, they experience waiting times proportional to their position in the queue. Congestion potentially explains why an organization may split a team manager's time between helping and other tasks, as filling a manager's entire time

[25] The pattern in the figure is U-shaped in the trained worker's knowledge. For a small increase in z_1^t relative to z_1, the trained worker fields help requests but is unlikely to solve them, causing untrained workers to spend more aggregate time seeking help rather than working on problems. Untrained workers' goal achievement initially falls as z_1^t increases because trained workers free up capacity for help under the the constraint that their own solutions cannot decline. The slope eventually turns positive as the trained worker's knowledge increases, but at all points untrained workers' goal achievement falls. Note that the horizontal spillover can be positive under alternate parameter values, which Figure A11 illustrates.

allocation with help requests can lead to delays for workers.

Remark 2 (Congestion): Although the main intuition remains, differences in results with congestion are: 1) Even when the trained worker does not provide help, output weakly increases for untrained workers due to congestion reductions. 2) Help from trained workers can further alleviate managerial congestion and requires numerical fixed point methods to characterize the full effect for untrained workers.

4.1.2 The Organization's Problem

We now characterize the organization's problem, allowing us to value the direct training gains and spillovers. The organization's goal is to produce at minimum cost subject to an output target, where costs are wages times headcount. The expression for wages is from Caliendo and Rossi-Hansberg (2012), where w is a base wage and c is the marginal cost of knowledge acquisition. There are n_1 frontline workers and n_m managers who each handle n_1/n_m help requests. The production function is $Q = (n_1\phi)^a(n_mS)^b$, where a and b are the output elasticities for frontline and strategic tasks, respectively, and $a+b$ determines returns to scale. The organization's problem in the pre-period is:

$$\min_{\{n_1,n_m,z_1,z_m\}} n_1 w(1+cz_1) + n_m w(1+cz_m) \text{ subject to} \tag{4}$$

$$Q = (n_1 \frac{1-e^{-\lambda z_m}}{1+H_m e^{-\lambda z_1}})^a (n_m S)^b \geq Y. \tag{5}$$

When μ is the Lagrange multiplier on the output constraint, there are 5 first order conditions that characterize a solution (see equation 9 in Appendix C.2).[26] However, even though we observe individual goal-achievement in the data, we do not know a and b, so overall

[26]We have suppressed a time constraint that $S \geq 0$ because managers' strategy time is valuable directly, meaning that for any $b > 0$ and $Q > 0$, we will have $S > 0$. At an optimum, non-negativity constraints are also slack, and we have suppressed them to simplify the derivation.

output in the pre- and post-periods is not directly known. By assuming that the pre-training organization solves the first order conditions, we can infer the values of a and b that rationalize the choices of wages and headcount if we have estimates of ϕ, S, w, and c that go into the first order conditions. The conditions for this approach to work are given in Proposition 3.

Proposition 3 (Identification)**.** *Take the first order conditions with respect to z_1, n_1, z_m, and n_m. If the 4×3 matrix of derivatives of these first order conditions with respect to μ, a, and b has rank 3, then the output elasticities a and b are identified given estimates of w, c, z_1, z_m, h_m, and λ and input data n_1 and n_m.*

Proof: See Appendix C.2.

The motivation for the training program can be thought of as a need to shift the output level from $Y = Q_{Pre}$ to $Y = Q_{Post}$ where $Q_{Post} > Q_{Pre}$. Intuitively, we infer a and b from the organization's pre-period choices, which will allow us to recover the realized output Q_{Post} after the training program occurred, valuing the direct gains relative to spillovers.

4.2 GMM Estimation

Our estimator uses the moment conditions implied by the model to recover the underlying parameters that govern knowledge levels, helping costs, and production technologies in the organization. Table A10 displays the moment conditions we describe below. Five moment conditions come from pre-period data. We match frontline workers' and managers' average pre-period goal achievement to $\phi = \frac{1-e^{-\lambda z_m}}{1+H_{m,Pre}e^{-\lambda z_1}}$ and $S = 1-\frac{n_1}{n_m}\frac{h_m e^{-\lambda z_1}}{1+H_{m,Pre}e^{-\lambda z_1}}$. These expressions have 4 unknown parameters: λ, z_1, z_m, and h_m.[27] We match average wage levels and wage

[27] The term $H_{m,Pre}$ is the expected time it takes a worker to get help from a manager and reflects congestion or queuing for help. Let $\tilde{h}_{Pre}$ be total help requests fielded by a manager so that $h_m\tilde{h}_{Pre}$ is the total time each

ratios between managers and workers to functions of the form $w(1 + cz)$ where $z \in \{z_1, z_m\}$. These 3 moment conditions based on wage data add 2 unknown parameters, w and c.

We match four theoretical moments to post-period data. For these moments, we substitute λ_{Post} for λ, allowing the problem distribution to shift over time. When computing the model where trained workers provide help, we assume they address a share, ρ, of connected, untrained workers' help requests. For this exercise, untrained wage band 1 workers with above-median exposure to trainees are defined as connected and all other workers are unconnected. We match unconnected workers' changes in goal achievement over time to pin down λ_{Post}. We also match post-period goal achievement for trained workers, adding two new parameters: z_1^t, trained workers' knowledge, and h_t, trained workers' helping time cost (when $\rho > 0$). We match connected workers' goal achievement when ρ help requests go first to trained workers and $1 - \rho$ go directly to managers. Finally, we match managers' post-period goal achievement, identifying managerial helping costs separately from knowledge levels.

We estimate the model via one-step GMM and verify that the derivative matrix of the moment conditions with respect to the parameters has full rank (an identification condition). We bootstrap the standard errors using 150 random samples drawn with replacement.

Parameter Estimates: Table 4 Panel A displays the parameter estimates. The first column corresponds to the model when trained workers do not provide help (akin to Proposition 1, with $\rho = 0$). The second column (akin to Proposition 2) assumes that trained workers help connected untrained workers, with a value of $\rho = 0.12$.[28] In practice, the estimates are

manager spends helping. This yields an expected wait time to receive help of $H_{m,Pre} = h_m(1 + \tilde{h}_{Pre})/2$. $h_m\tilde{h}_{Pre}$ is the solution a fixed point problem, which we compute numerically in an inner loop in the estimation algorithm. We solve for a different value, $H_{m,Post}$, when we match post-period moments.

[28] We fix ρ because of challenges separately identify it from z_1^t and h_t. The value of ρ that we choose provides a plausible estimate of trained workers' help provision based on the email data. For example, untrained workers in wage band 1 increase emails to trained workers by about 4.5 - 6 percent of the baseline level

very similar, although the trained workers' helping cost, h_t, is not estimated when $\rho = 0$.

Combinations of parameters are useful to understand the environment. For example, in the model with help, $z_1 = 0.196$, $z_m = 1.632$ and $\lambda = 1.123$ imply that frontline workers handled just 20% of their pre-period tasks autonomously, while managers could solve 84 percent of their problems. Managers' helping cost, h_m, implies that each request takes about 10% of a manager's unit of time. For workers, congestion raises the burden of asking for help, with an expected waiting time of 0.21 time units to get help in the pre-period.

Trained workers' have substantial knowledge gains. $z_1^t = 1.08$ in the model with help means they handle 70% of their problems autonomously (untrained workers handle 19% of their own problems under the distribution with parameter λ_{Post}). The model fits the reduction in emails from trained workers to managers reasonably well. The last row in Panel B shows that help requests from trained workers to managers fall by about 58% compared to a 74% decline in the raw data and a 62% decline in relative terms after accounting for the year-over-year reduction in overall email volume.[29]

The most important findings are the estimates of a and b, the output elasticities of the production aggregator. These estimates come from the solutions to the first order conditions for the organization, and we verify that the rank condition in Proposition 3 holds conditional on the GMM estimates of the input parameters. Frontline workers' production tasks and

of emails sent to managers. Because we define highly connected untrained workers based on those with above-median connections, it suggests that trained workers likely handle a share of approximately $\rho = 0.12$ help requests from connected workers. Tables A11 and A12 show estimates while varying ρ.

[29]The model also suggests that trained workers' costs of helping others are substantially smaller than the cost of turning to a manager for help, as h_t is under 0.01. The low helping costs are partly mechanical due to the choice of ρ, as trained workers' increased goal achievement is difficult to rationalize if they spend a significant share of their time helping others. As Table A11 shows, smaller values of ρ raise the estimate of h_t. Despite this identification challenge, the importance of vertical spillovers is similar across the values of ρ.

managers' strategic tasks have output elasticities of 0.30 and 0.53, respectively. Manager's strategic tasks have implicit valuations that are relatively high compared to workers', and changes to managers' time spent on strategy are likely valuable.

Decomposing Program Returns: Panel B provides estimates of output changes (using the post-period distribution of problems) under different assumptions about vertical spillovers. When managers' full goal achievement gains are attributed to spillovers, the total output increase from the training program is 1.94 and 2.46 percent in the two models. These calculations include the direct training gains, managers' gains, and any goal achievement increase for untrained workers. The difference across columns reflects horizontal spillovers and congestion alleviation. When only 1 percentage point of managers' gains come from spillovers, output gains are 1.05% and 1.18%, respectively. The next two rows show the share of gains from vertical spillovers under different assumptions. Approximately 74% of the productivity benefits come from vertical spillovers when we attribute managers' full goal achievement change to the program. Between 45 and 52% of the benefits come from vertical spillovers when managers' goal achievement gains from spillovers total 1 percentage point.[30]

To approximate the worth of the vertical spillovers, we ask how much more the organization would have needed to spend on training to reach the same post-period isoquant in the absence of spillovers (Figure A12 illustrates the approach). In the absence of vertical spillovers, the number of trained workers would need to roughly double to stay on the isoquant where managers' goal achievement increased by 1 percentage point.

[30]Appendix C.3 details a sensitivity analysis when we extend the model to have 3 layers rather than 2. In the 3 layer model (Table A13), we find that about 50% of the program gains are due to vertical spillovers to managers when 1 percentage point of managers' goal achievement increase is due to spillovers from subordinates' training.

5 Evidence on Mechanisms, Alternative Explanations, and Discussion

To provide evidence on mechanisms and to address potential alternative explanations, we conducted a follow-up survey in August of 2020. The agency distributed the survey to 63 trained and 105 untrained workers who were present in the pre-period.[31]

To assess how emails proxy for overall communications, the survey asked about the frequency of face-to-face interaction with email contacts.[32] The majority of workers interact either several times per week or at least weekly with those they frequently email, suggesting that emails and face-to-face communication are complements (see Figure A1). The survey also probed why frontline workers email managers. Three of four workers reported that asking for help is the main reason to contact superiors, with the remaining responses split evenly between asking for authorizations and reporting on task progress.[33] When presented with data on reduced emails to managers in the post-period, trained workers had no abnormal perceptions of directives to change help requests going to managers (see Table A14).

When fielding the survey, we sought to understand the relatively large direct gains from training. We asked respondents about changes between the pre- and post-periods that would explain trainees' goal achievement gains. Trained workers reported much greater improvements in their knowledge of task requirements, their understanding of appropriate

[31] The survey contained 9 questions and had an estimated completion time of less than 10 minutes. The survey was described as part of research on the organization's working environment conducted by independent researchers. Participation was voluntary and was not incentivized. Fifty-two percent of the trained workers (N=33 workers) and 54% of the untrained workers (N=57) took the survey. Appendix D contains the English version of the survey.

[32] During the sample period, the organization prohibited the work-related use of other communication technologies such as WhatsApp and Skype.

[33] We conducted an additional survey with a smaller, more targeted group in 2024 to understand how help requests propagate through the organization. Figure A10 shows that 51% of emails entail requests for help or responses to help requests, and 30% are requests for authorizations. About 38% of workers' questions to managers are through email.

workflows and protocols, and their general skills and knowledge (Table A14). Still, we suspect the direct gains may be larger for the agency than in other settings. In most contexts, one would expect that rent sharing would result in implicit incentives for on-the-job productivity gains. In our setting, these rent sharing incentives are absent, while the wage function in the model is written to account for the need to pay trainers when increasing skill. Given the relatively high estimated price of skill acquisition, workers in the agency likely had little incentive to improve in the absence of the program. In fact, the training program did not change perceptions of career incentives or working hours, consistent with these muted incentives (see Table A14, although trained workers were more likely to receive temporary promotions).[34]

Finally, it is possible that the nature of workers' tasks changed over time. As we were concerned that changes in email patterns might reflect changes in team-related tasks, we probed whether task interdependence changed. Multi-person tasks did not change differentially by training status (Table A14). It is also unlikely that training on a new type of legal case or precedent fully explains our results, as workers in functional areas related to the organization's operations had similar training gains to those focused on cases.

6 Conclusion

Our analysis demonstrates that training programs can generate substantial value beyond their direct effects on participants. By documenting and quantifying spillovers through organizational hierarchies, we show that individual-level evaluations may understate program returns, as approximately 45% of the overall benefits from a frontline worker training program

[34]The survey also probed potential alternative explanations for our results. Changes in monitoring or supervision do not appear to explain the findings, as 85% of workers reported that supervision remained constant (the top-right panel of Figure A1).

that we study come from spillovers to managers. Managers shifted time from providing help to their own strategic tasks, confirming a core comparative static about the relationship between worker skill and support needs in knowledge hierarchies models (Garicano, 2000; Garicano and Rossi-Hansberg, 2006; Caliendo and Rossi-Hansberg, 2012). We provide some of the first empirical evidence that employees in lower levels of a hierarchy can impact those at the top, giving a rationale for positive assortative matching between workers and managers (Bloom and Van Reenen, 2007).

Future research could explore how different types of training, variations in hierarchical structures, or technologies that reduce the cost of training interact. For example, if the price of skills acquisition falls, it may flatten organizations (Rajan and Wulf, 2006; Guadalupe and Wulf, 2010). Additional work could also explore how our results vary across labor market conditions and contexts. A primary reason why firms can capture value from training comes from labor market frictions and wage compression. The low turnover rates and limited ability to hire and fire in our study organization suggest these frictions are present and necessitate training rather than other forms of workforce adjustment. While the ability for firms to capture the direct gains from training may be limited when trainees receive credentials or certifications, it is possible that spillover benefits may be relatively durable even in labor markets with fewer frictions.

References

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

Acemoglu, D. and Pischke, J.-S. (1999). The structure of wages and investment in general training. *Journal of Political Economy*, 107(3):539–572.

Adhvaryu, A., Bassi, V., Nyshadham, A., Tamayo, J. A., and Torres, N. (2023a). Organizational responses to product cycles. Technical report, National Bureau of Economic Research.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2023b). Returns to on-the-job soft skills training. *Journal of Political Economy*, 131(8):2165–2208.

Autor, D. H. (2001). Why Do Temporary Help Firms Provide Free General Skills Training? *The Quarterly Journal of Economics*, 116(4):1409–1448.

Bandiera, O., Barankay, I., and Rasul, I. (2007). Incentives for managers and inequality among workers: Evidence from a firm-level experiment. *The Quarterly Journal of Economics*, 122(2):729–773.

Bandiera, O., Barankay, I., and Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, 77(4):1047–1094.

Bandiera, O., Best, M. C., Khan, A. Q., and Prat, A. (2021). The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats. *The Quarterly Journal of Economics*, 136(4):2195–2242.

Bartel, A. P. (1995). Training, Wage Growth, and Job Performance: Evidence from a Company Database. *Journal of Labor Economics*, 13(3):401–425.

Bartel, A. P. (2000). Measuring the employer's return on investments in training: Evidence from the literature. *Industrial relations: a journal of economy and society*, 39(3):502–524.

Becker, G. S. (1964). Human capital theory. *Columbia, New York*, 1964.

Becker, G. S. (1994). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. The University of Chicago Press.

Belloni, A., Chernozhukov, V., Hansen, C., and Kozbur, D. (2016). Inference in high-dimensional panel models with an application to gun control. *Journal of Business & Economic Statistics*, 34(4):590–605.

Bianchi, N. and Giorcelli, M. (2022). The dynamics and spillovers of management interventions: Evidence from the training within industry program. *Journal of Political Economy*, 130(6):1630–1675.

Black, S. E. and Lynch, L. M. (1996). Human-capital Investments and Productivity. *American Economic Review*, 86(2):263–267.

Bloom, N., Garicano, L., Sadun, R., and Van Reenen, J. (2014). The distinct effects of information technology and communication technology on firm organization. *Management Science*, 60(12):2859–2885.

Bloom, N., Lemos, R., Sadun, R., and Van Reenen, J. (2015). Does Management Matter in Schools? *The Economic Journal*, 125(584):647–674.

Bloom, N., Lemos, R., Sadun, R., and Van Reenen, J. (2020). Healthy Business? Managerial Education and Management in Health Care. *Review of Economics and Statistics*, 102(3):506–517.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, 122(4):1351–1408.

Caliendo, L., Mion, G., Opromolla, L. D., and Rossi-Hansberg, E. (2020). Productivity and organization in portuguese firms. *Journal of Political Economy*, 128(11):4211–4257.

Caliendo, L., Monte, F., and Rossi-Hansberg, E. (2015). The Anatomy of French Production Hierarchies. *Journal of Political Economy*, 123(4):809–852.

Caliendo, L. and Rossi-Hansberg, E. (2012). The impact of trade on organization and productivity. *The Quarterly Journal of Economics*, 127(3):1393–1467.

Cappelli, P. (2012). *Why good people can't get jobs: The skills gap and what companies can do about it*. University of Pennsylvania Press.

Card, D., Kluve, J., and Weber, A. (2018). What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. *Journal of the European Economic Association*, 16(3):894–931.

Chen, C. and Suen, W. (2019). The Comparative Statics of Optimal Hierarchies. *American Economic Journal: Microeconomics*, 11(2):1–25.

De Grip, A. and Sauermann, J. (2012). The Effects of Training on Own and Co-worker Productivity: Evidence from a Field Experiment. *The Economic Journal*, 122(560):376–399.

Finan, F., Olken, B. A., and Pande, R. (2017). Chapter 6 - The Personnel Economics of the Developing State. In Banerjee, A. V. and Duflo, E., editors, *Handbook of Economic Field Experiments*, volume 2 of *Handbook of Economic Field Experiments*, pages 467–514. North-Holland.

Friedrich, B. U. (2022). Trade shocks, firm hierarchies, and wage inequality. *Review of Economics and Statistics*, 104(4):652–667.

Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. *Journal of Political Economy*, 108(5):874–904.

Garicano, L. and Hubbard, T. N. (2003). Firms' boundaries and the division of labor: Empirical strategies. *Journal of the European Economic Association*, 1(2-3):495–502.

Garicano, L. and Hubbard, T. N. (2007). Managerial leverage is limited by the extent of the market: Hierarchies, specialization, and the utilization of lawyers' human capital. *The Journal of Law and Economics*, 50(1):1–43.

Garicano, L. and Hubbard, T. N. (2009). Specialization, firms, and markets: The division of labor within and between law firms. *The Journal of Law, Economics, & Organization*, 25(2):339–371.

Garicano, L. and Rossi-Hansberg, E. (2004). Inequality and the organization of knowledge. *American Economic Review*, 94(2):197–202.

Garicano, L. and Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. *The Quarterly Journal of Economics*, 121(4):1383–1435.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6):1126–1152.

Grissom, J. A., Viano, S. L., and Selin, J. L. (2016). Understanding Employee Turnover in the Public Sector: Insights from Research on Teacher Mobility. *Public Administration Review*, 76(2):241–251.

Guadalupe, M. and Wulf, J. (2010). The Flattening Firm and Product Market Competition: The Effect of Trade Liberalization on Corporate Hierarchies. *American Economic Journal: Applied Economics*, 2(4):105–27.

Gumpert, A., Steimer, H., and Antoni, M. (2022). Firm organization with multiple establishments. *The Quarterly Journal of Economics*, 137(2):1091–1138.

Hoffman, M. and Tadelis, S. (2021). People management skills, employee attrition, and manager rewards: An empirical analysis. *Journal of Political Economy*, 129(1):243–285.

Hur, H. and Abner, G. (2022). What makes public employees want to leave their job? *Public Administration Review*, 82(2):218–236. Received: 21 February 2022; Revised: 30 November 2022; Accepted: 18 December 2022.

Jarosch, G., Oberfield, E., and Rossi-Hansberg, E. (2021). Learning from coworkers. *Econometrica*, 89(2):647–676.

Jones, B. F. and Summers, L. H. (2020). A calculation of the social returns to innovation. Technical report.

Kugler, A., Kugler, M., Saavedra, J. E., and Herrera-Prada, L. O. (2022). Long-Term Educational Consequences of Vocational Training in Colombia: Impacts on Young Trainees and Their Relatives. *Journal of Human Resources*, pages 0518–9528R2.

Lazear, E. P., Shaw, K. L., and Stanton, C. T. (2015). The Value of Bosses. *Journal of Labor Economics*, 33(4):823–861.

Levitt, S. D., List, J. A., and Syverson, C. (2013). Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant. *Journal of Political Economy*, 121(4):643–681.

Miller, O. (2018). *Employee turnover in the public sector*. Routledge.

Rajan, R. G. and Wulf, J. (2006). The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies. *The Review of Economics and Statistics*, 88(4):759–773.

Rasul, I. and Rogger, D. (2018). Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *The Economic Journal*, 128(608):413–446.

Rasul, I., Rogger, D., and Williams, M. J. (2018). *Management and Bureaucratic Effectiveness: Evidence from the Ghanaian Civil Service*. The World Bank.

Sandvik, J., Saouma, R., Seegert, N., and Stanton, C. T. (Forthcoming). Should Human Capital Development Programs be Voluntary or Mandatory? Evidence from a Field Experiment on Mentorship. *Management Science*.

Sandvik, J. J., Saouma, R. E., Seegert, N. T., and Stanton, C. T. (2020). Workplace knowledge flows. *The Quarterly Journal of Economics*, 135(3):1635–1680.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288.

Zhao, P. and Yu, B. (2006). On model selection consistency of lasso. *Journal of Machine learning research*, 7(Nov):2541–2563.

Figures and Tables

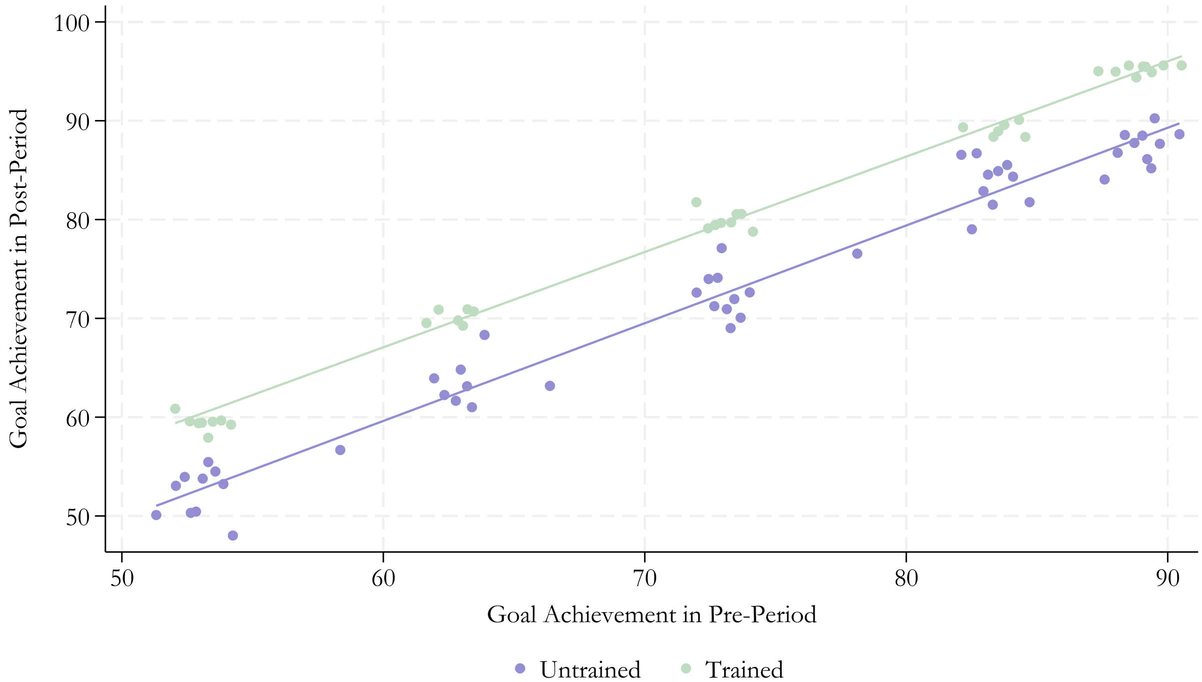
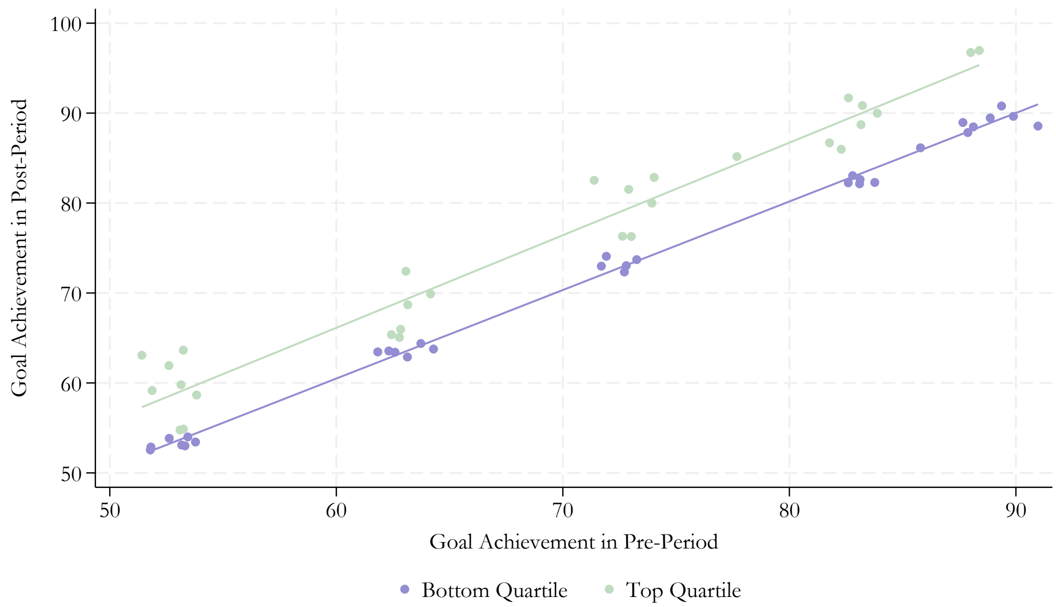
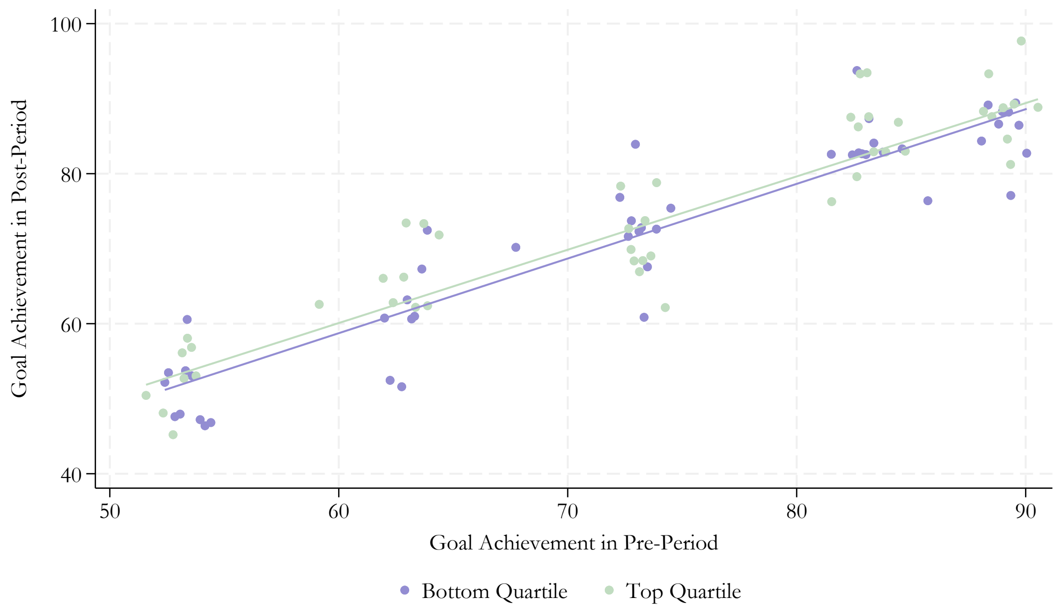


Figure 1: Goal Achievement Changes For Workers by Training Status

Note: This figure plots the relationship between individual frontline workers' average goal achievement in the pre- and post-periods. There are separate plots based on whether the worker was randomized into the training program or not.



(a) Managers



(b) Untrained Workers

Figure 2: Goal Achievement Changes by Connection Strength to Trained Workers

Note: This figure plots the relationship between goal achievement in the pre- and post-periods for managers and untrained workers. We plot the relationship separately based on the strength of a manager (untrained worker)'s connections to eventually trained workers, which we calculate from the level of pre-period emails with program participants. We split the data by quartile and display the top and bottom quartiles.

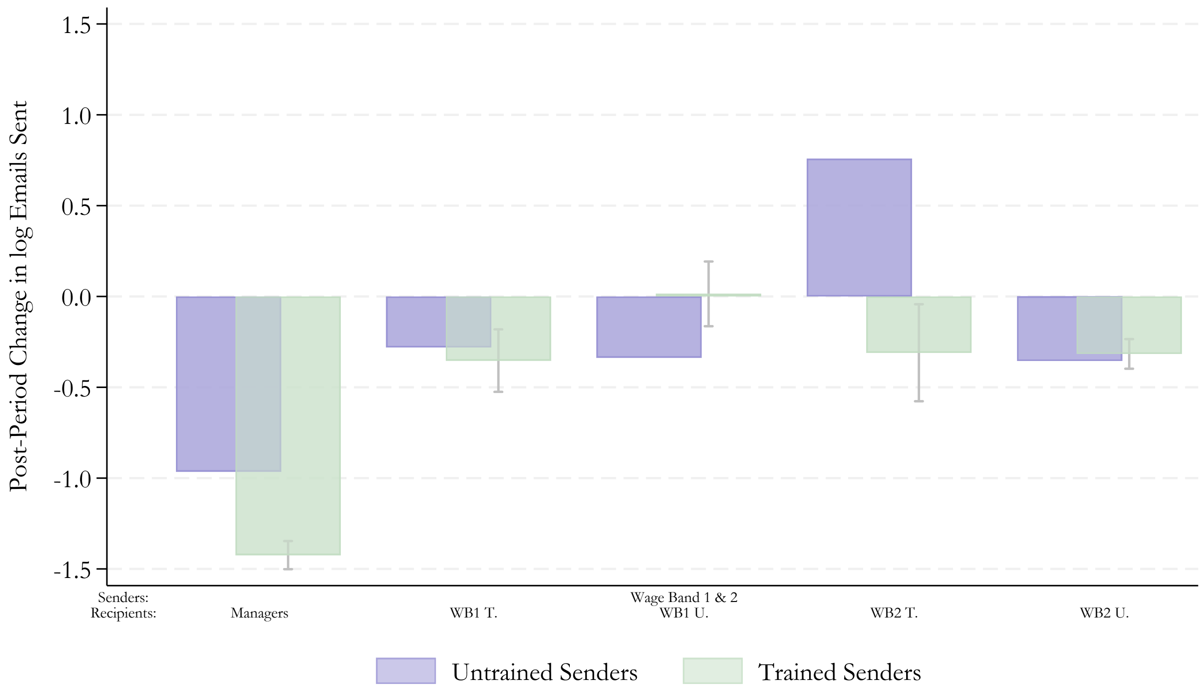
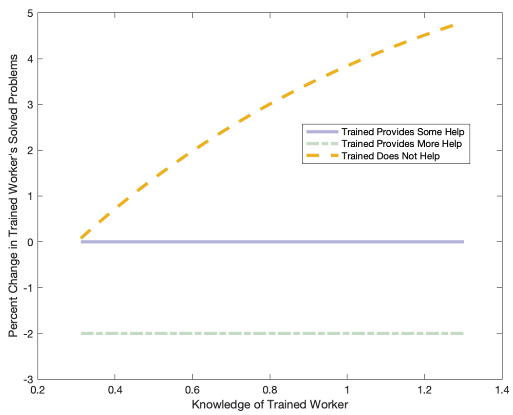
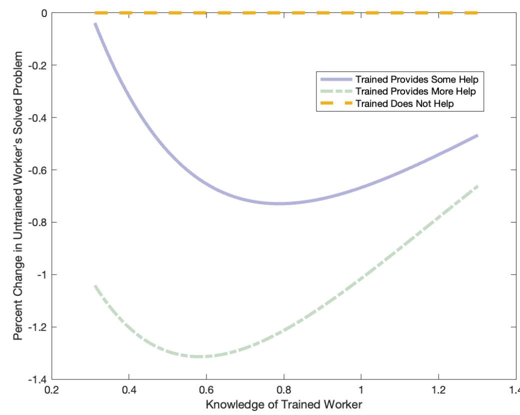


Figure 3: Changes in Log Emails Between the Pre- and Post-Period by Recipient Type

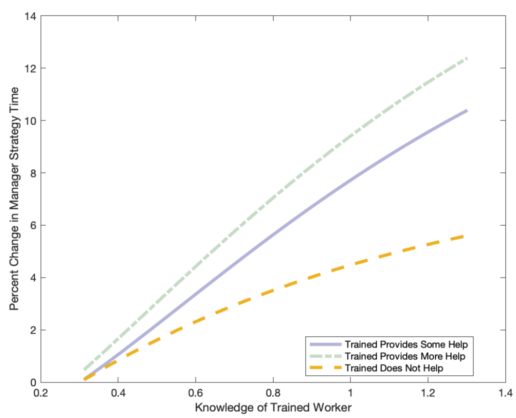
Note: This figure displays the average change in log emails sent by trained and untrained frontline workers. Each destination represents a separate recipient type by wage band and training status. Purple bars represent the average change for untrained senders and green bars are the change for trained senders. Recipient types are: Managers, Wage Band 1 Trained Workers (abbreviated as WB1 T.), Wage Band 1 Untrained Workers (WB1 U.), Wage Band 2 Trained Workers (WB2 T.), and Wage Band 2 Untrained Workers (WB2 U.). Standard errors come from a difference-in-differences regression of log emails on a post-period-by-trained dummy. The regression are run by recipient type and include fixed effects for workers and time. Standard errors are clustered by sender.



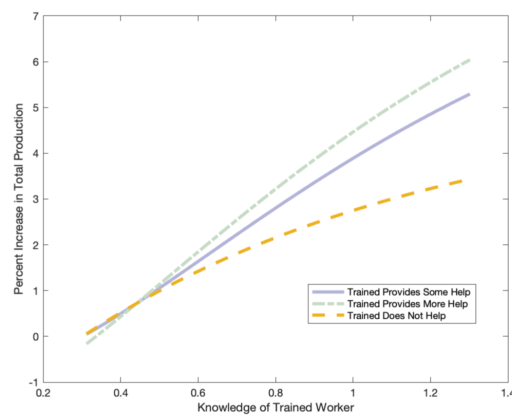
(a) Trained Solutions



(b) Untrained Solutions



(c) Manager Time for Strategy



(d) Total Output

Figure 4: Illustration of Individual and Organization-Level Output Changes as a Function of the Trained Worker's Knowledge

Note: These figures plot percent changes in output for individual workers, managers, and the organization/team as a function of one trained worker's increase in knowledge. The vertical axis displays percent changes relative to a baseline where n_1 workers report to a manager and n_1, z_1, and z_m are chosen to satisfy the production constraint $(n_1\phi)^{0.3} \times (S)^{0.53} = 1.15$ with $h_m = H_m = 0.102$, $w = .916$, and $c = 1.4$. The optimal solution in the baseline scenario sets $n_1 = 2.96$, $z_1 = 0.3$, and $z_m = 1.73$. The x-axis in each subfigure displays z_1^t for the trained worker. There are three scenarios. In the first, trained workers continue to solve the same number of problems and provide help with any additional time they have remaining (labeled "Trained Provides Some Help"). Help from a trained worker has a time cost $H_t = 0.85 \times H_m$. In the second scenario, trained workers provide more help by lowering their own solutions by 2% relative to the baseline (labeled "Trained Provides More Help"). In the third scenario, trained workers provide no help, instead concentrating on solving problems.

	(1) Wage Band 1 Workers	(2) Wage Band 2 Workers	(3) Managers	(4) Untrained Workers	(5) Trained Workers	(6) Difference of (5)-(4)
Female	0.483	0.285	0.178	0.400	0.556	0.156**
						(0.067)
Secondary Education	0.715	0.500	0.000	0.644	0.651	0.007
						(0.065)
Bachelors Degree	0.274	0.494	0.636	0.346	0.349	0.004
						(0.064)
Masters-PhD	0.011	0.006	0.364	0.011	0.000	-0.011**
						(0.005)
Execution Division	0.492	0.209	0.318	0.397	0.413	0.015
						(0.067)
Wage Band	1.000	2.000	3.341	1.333	1.286	-0.047
			(0.523)	(0.472)	(0.455)	(0.061)
Wages, Pre-Period (normalized)	1.000	1.464	2.955	1.156	1.124	-0.032
	(0.383)	(0.577)	(1.588)	(0.507)	(0.489)	(0.066)
Wages, Post-Period (normalized)	1.045	1.530	3.091	1.207	1.174	-0.033
	(0.401)	(0.603)	(1.657)	(0.530)	(0.511)	(0.069)
Goal Achievement, Pre-Period	0.720	0.735	0.708	0.726	0.719	-0.007
	(0.131)	(0.134)	(0.130)	(0.131)	(0.135)	(0.018)
Goal Achievement, Post-Period	0.723	0.740	0.730	0.721	0.785	0.065***
	(0.153)	(0.133)	(0.136)	(0.147)	(0.131)	(0.018)
Number of individuals	354	172	129	463	63	
F-statistic						0.885
						(0.547)

Table 1: Descriptive Statistics and Treatment Assignment Balance on Observable Characteristics

This table displays descriptive statistics for frontline workers (wage bands 1 and 2) and managers (wage bands 3-5). Columns 1 and 2 split the sample by wage band for frontline workers. Column 3 displays statistics for managers. Columns 4-6 provide balancing tests for random treatment assignment for trained and untrained frontline workers. Column 6 displays t-tests of differences between columns 4 and 5. Secondary Education, Bachelors Degree, and Masters-PhD are dummy variables for the worker's highest education level. Execution Division is a division dummy variables. Wages are normalized relative to the mean pre-period wage for wage band 1. Goal Achievement (GA) is the fraction of achieved goals, measured weekly and averaged over weeks. The last row computes the joint F-statistic and the associated p-value (in parenthesis) from regressing training status on frontline workers' observable characteristics. Statistical significance levels are denoted *** p<0.01, ** p<0.05, * p<0.1.

	(1)	(2)	(3)	(4)	(5)	(6)
Trained × Post	0.105***	0.104***	0.131***	0.163***	0.105***	0.104***
	(0.006)	(0.007)	(0.013)	(0.014)	(0.006)	(0.007)
Observations	12834	12834	12834	12834	12834	12834
R-squared	0.903	0.911	0.904	0.912	0.903	0.911
Worker Fixed Effects	✓	✓	✓	✓	✓	✓
Time Fixed Effects	✓	✓	✓	✓	✓	✓
Division-Time Fixed Effects	×	✓	×	✓	×	✓
Sociodemographic Controls	×	×	✓	✓	✓	✓
Post-LASSO OLS	×	×	×	×	✓	✓

Table 2: Treatment Effects of Training For Frontline Workers

Note: This table displays estimates of treatment effects from training. The dependent variable is log goal achievement and the unit of observation is a frontline worker-week. All regressions include worker fixed effects and time fixed effects. Even-numbered columns include division-by-time fixed effects. Columns 3 and 4 include interactions of sociodemographic characteristics with the trained-by-post indicator. Sociodemographic characteristics are dummies for wage band 2, a bachelors degree or more, female, and a High-Performer dummy (defined as above-median pre-period goal achievement). Online Appendix Table A2 reports the estimates on these interactions. Columns 5 and 6 report post-LASSO OLS regressions after selecting variables using the Stata rlasso package with a penalty term used in Belloni et al. (2016) that accounts for clustering. The regressors entering the LASSO are those that enter the models in Columns 3 and 4. None of the interactions with worker characteristics and training are selected by the Lasso. Standard errors are clustered at the worker level. Statistical significance levels are denoted *** p<0.01, ** p<0.05, * p<0.1.

	(1)	(2)	(3)	(4)	(5)	(6)
Exposure (Levels) × Post	0.104***	0.104***	0.177**			
	(0.009)	(0.009)	(0.074)			
Exposure (Shares) × Post				0.021**	0.021**	0.042*
				(0.009)	(0.009)	(0.023)
Post	-0.012***			0.019***		
	(0.003)			(0.006)		
Avg. Implied Change	2.98	2.98	5.05	.828	.824	1.63
N	3154	3154	3154	3154	3154	3154
R^2	.951	.952	.953	.943	.944	.953
Worker Fixed Effects	✓	✓	✓	✓	✓	✓
Time Fixed Effects	×	✓	✓	×	✓	✓
Division-Time Fixed Effects	×	×	✓	×	×	✓

Table 3: Estimates of Vertical Training Spillovers to Managers Based on Pre-Period Exposure to Eventually Trained Workers

Note: This table displays estimates of spillovers from training to managers. The dependent variable is log goal achievement and the unit of observation is a manager-week. Measures of exposure to eventually trained workers are computed from emails in the pre-training period. In the first three columns, the exposure measure is based on the level of emails received from eventually trained workers. In the last three columns, the measure is based on the share of emails with eventually trained workers relative to all emails from frontline workers. Standard errors are clustered at the manager level. All Columns include manager fixed effects. Columns 2 and 5 add time fixed effects rather than using a Post-period indicator while Column 3 and 6 include time-by-division fixed effects. The average implied change in goal achievement takes the predicted effects from the model in logs and multiplies by the individual manager's average of pre-period goal achievement to yield a percentage point change in average goal achievement due to spillovers. Statistical significance levels are denoted *** p<0.01, ** p<0.05, * p<0.1.

Help from Trained:	No	Yes		No	Yes
		Panel A: Parameter Estimates			
z_1	0.191	0.196	z_1^t	1.089	1.077
	(0.055)	(0.056)		(0.305)	(0.305)
z_m	1.618	1.632	h_t		0.004
	(0.213)	(0.175)			(0.028)
c	1.388	1.388	λ_{Post}	1.097	1.090
	(0.217)	(0.172)		(0.129)	(0.122)
w	0.912	0.907	a	0.292	0.297
	(0.065)	(0.060)		(0.033)	(0.038)
λ	1.126	1.123	b	0.538	0.531
	(0.133)	(0.131)		(0.056)	(0.060)
h_m	0.102	0.104			
	(0.012)	(0.012)			
		Panel B: Output and Help Request Changes			
Output Pct Gains with Training and Full Spillovers				1.94	2.46
				(0.47)	(0.61)
Output Pct Gains with 1 P.P. Vertical Spillovers				1.05	1.18
				(0.12)	(0.15)
Share of Gains from Vertical Spillovers (Full)				73.85	73.75
				(4.26)	(4.55)
Share of Gains from Vertical Spillovers (1 P.P.)				51.51	45.24
				(7.93)	(8.41)
Pct Change in Help Req to Mgrs from Trained				-58.87	-57.79
				(11.59)	(11.81)

Table 4: GMM Estimates and Production Function Parameters

Note: This table presents GMM estimates using the moment conditions described in the text and displayed in Online Appendix Table A10. Panel A displays parameter estimates. The model in the column without help from trained workers assumes that all untrained workers turn directly to managers for help. The model with help from trained workers fixes the share of help requests that trained workers handle for connected untrained workers. Estimating this share from the data is challenging due to the need to infer both ρ and z_1^t using only variation in post-period productivity. We set ρ to 0.12 because emails from wage band 1 untrained workers to trained workers increase by 4.5%; we only consider untrained Wage Band 1 workers with above-median exposure as connected, and we net out a base share of emails to managers (6% in survey data) related to authorizations (Online Appendix Table A11 provides sensitivity analysis around the choice of ρ). The production function takes the form $Q = (n_1\phi)^a(n_mS)^b$, where $n_1\phi$ gives the number of completed frontline tasks, S is manager time for strategy, and n_m is the number of managers. The output elasticities a and b are estimated as detailed in the text. Panel B displays model-implied changes in post-period output under problem distribution λ_{Post} after the training regime and with different assumptions about the degree of vertical spillovers. The first row displays the percent change in output when the full change in manager goal achievement is attributed to spillovers. The second row assumes manager output gains increase by 1 percentage point due to spillovers. The output gain includes horizontal spillovers in the model with help from trained workers. The third and fourth rows display the share of gains attributed to vertical spillovers. The final row displays model-implied percent changes in help requests sent to managers by trained workers. Online Appendix Tables A12 and A13 contain additional detail about output in the pre- and post-periods, output decompositions under different levels of ρ, estimates assuming constant returns to scale in the production function, and estimates assuming there are 2 layers of frontline workers. Standard errors are estimated from 150 bootstrap iterations of the entire estimation procedure.

A Appendix Tables

	(1)	(2)	(3)	(4)	(5)	(6)	Mean	SD	Obs
Panel A: Managers									
(1): Emails from Managers to Trained Workers	1.00						1,644	837	129
(2): Emails from Trained Workers to Managers	0.91	1.00					1,670	893	129
(3): Exposure, Share of Manager Emails Sent to Trained Workers	0.33	0.32	1.00				0.55	0.27	129
(4): Exposure, Share of Manager Emails Received from Trained Workers	0.53	0.25	0.62	1.00			0.47	0.18	129
(5): Exposure, Levels of Emails from Managers to Trained Workers	1.00	0.91	0.33	0.53	1.00		0.40	0.24	129
(6): Exposure, Levels of Emails from Trained Workers to Managers	0.91	1.00	0.32	0.25	0.91	1.00	0.40	0.34	129
Panel B: Untrained Workers									
(1): Emails from Untrained to Trained Workers	1.00						668	376	463
(2): Emails from Trained Workers to Untrained Workers	0.85	1.00					674	385	463
(3): Exposure, Share of Emails Sent from Untrained to Trained Workers	0.23	0.49	1.00				0.50	0.17	463
(4): Exposure, Share of Emails Sent from Trained to Untrained Workers	0.48	0.19	0.28	1.00			0.40	0.15	463
(5): Exposure, Levels of Emails from Untrained to Trained	1.00	0.85	0.23	0.48	1.00		0.39	0.25	463
(6): Exposure, Levels of Emails from Trained to Untrained	0.85	1.00	0.49	0.19	0.85	1.00	0.38	0.23	463

Table A1: Details about Email-Based Measures of Exposure to Trained Workers

Note: This table displays correlations and summary statistics for various email-based measures of exposure to trained workers. All data come from the pre-period. The unit of analysis is managers in Panel A and untrained workers in Panel B. Email measures in levels capture the total number of emails with all eventually trained workers over the 12 week pre-period. Exposure measures take raw email data and apply a transformation to the unit interval, as defined in equation (1). Measures based on email shares in rows (3) and (4) divide by the total emails sent to or received from all frontline workers. Measures based on email levels in rows (5) and (6) use counts of emails. Columns (1)-(6) display correlations between the various exposure measures.

	(1)	(2)	(3)	(4)	(5)	(6)
			Log Goal Achievement			
Trained × Post	0.105***	0.104***	0.131***	0.163***	0.105***	0.104***
	(0.006)	(0.007)	(0.013)	(0.014)	(0.006)	(0.007)
Wage Band 2 × Trained × Post			-0.010	-0.049***		
			(0.010)	(0.013)		
Wage Band 2 × Post			0.014*	0.007		
			(0.008)	(0.008)		
Bachelor + Degree × Post			-0.002	-0.002		
			(0.010)	(0.009)		
Bachelor + Degree × Trained × Post			0.007	0.015		
			(0.012)	(0.014)		
Female × Post			-0.001	-0.007		
			(0.010)	(0.010)		
Female × Trained × Post			-0.005	-0.041***		
			(0.012)	(0.014)		
High Performer × Post			0.006	0.009		
			(0.010)	(0.009)		
High Performer × Trained × Post			-0.049***	-0.060***		
			(0.011)	(0.013)		
Observations	12834	12834	12834	12834	12834	12834
R-squared	0.903	0.911	0.904	0.912	0.903	0.911
Worker Fixed Effects	✓	✓	✓	✓	✓	✓
Time Fixed Effects	✓	✓	✓	✓	✓	✓
Division-Time Fixed Effects	×	✓	×	✓	×	✓
Socio-demographic Controls	×	×	✓	✓	✓	✓
Post-LASSO OLS	×	×	×	×	✓	✓

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

Table A2: Treatment Effects of Training For Frontline Workers

Note: This table displays estimates of training treatment effects. The unit of observation is a frontline worker-week. All regressions include worker fixed effects and time fixed effects. Even-numbered columns include division-by-time fixed effects. Columns 3 and 4 include interactions of sociodemographic characteristics with the trained-by-post indicator. Sociodemographic characteristics are dummies for wage band 2, a bachelors degree or more, female, and a High-Performer dummy (defined as above-median pre-period goal achievement). Columns 5 and 6 report post-LASSO OLS regressions after selecting variables using the Stata rlasso package with a penalty term used in Belloni et al. (2016) that accounts for clustering. The regressors entering the LASSO are those that enter the models in Columns 3 and 4. Standard errors are clustered at the worker level.

	(1)	(2)	(3)	(4)	(5)
	Non-Permanent Promotions				
Trained	0.287***	0.305**	0.321**	0.361***	0.356***
	(0.049)	(0.103)	(0.112)	(0.020)	(0.026)
Wage Band 2	-0.049	0.005	-0.021	0.091	0.088
	(0.063)	(0.069)	(0.066)	(0.058)	(0.061)
Higher Education	-0.057	-0.039	-0.036	-0.091	-0.090
	(0.075)	(0.068)	(0.068)	(0.071)	(0.067)
Female	0.036	0.007	0.034	0.036	0.090**
	(0.020)	(0.023)	(0.022)	(0.025)	(0.024)
High Performer	0.012	-0.010	-0.008	0.055	0.055
	(0.051)	(0.064)	(0.066)	(0.029)	(0.029)
Wage Band 2 × Trained		-0.205	-0.214	-0.257*	-0.248*
		(0.156)	(0.172)	(0.104)	(0.115)
Higher Education × Trained		-0.216	-0.205		
		(0.142)	(0.145)		
Female × Trained		0.072	0.051		
		(0.218)	(0.214)		
High Performer × Trained		0.170	0.153		
		(0.143)	(0.149)		
Higher Education × Wage Band 2				0.044	0.035
				(0.093)	(0.090)
Female × Wage Band 2				-0.074	-0.151
				(0.110)	(0.105)
High Performer × Wage Band 2				-0.133	-0.132
				(0.110)	(0.115)
Observations	512	512	512	512	512
R-squared	.055	.061	.069	.06	.071
Division Fixed Effects	✓	×	✓	×	✓
Sociodemographic Controls	✓	✓	✓	✓	✓

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

Table A3: Estimates of Temporary Promotions in the Post-Period

Note: This table presents regressions of factors that are correlated with non-permanent promotions to higher positions in 2021 and 2022, for the Workers from wage band 1 and 2, after the end of our main sample. The Non-Permanent Promotions variable takes a value of 1 if the worker was ever promoted in these years. There were no permanent promotions for workers in the sample during this period. No data is available for 2020 due to COVID, and 14 individuals were excluded due to their departure from the organization before 2021. The estimations across different columns progressively incorporate additional controls. Standard errors clustered at the division level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Goal Achievement					
Trained × Post	0.184***	0.198***	0.189***	0.197***	0.105***	0.104***
	(0.035)	(0.039)	(0.035)	(0.040)	(0.006)	(0.007)
Exposure to Untrained Workers (Levels) × Untrained × Post	0.016	-0.047	0.020	-0.038		
	(0.184)	(0.169)	(0.185)	(0.171)		
Exposure to Trained Workers (Levels) × Untrained × Post	0.051	0.165	0.041	0.148		
	(0.427)	(0.393)	(0.428)	(0.397)		
Exposure to Untrained Workers (Share) × Untrained × Post	0.067	0.140	0.061	0.128		
	(0.143)	(0.134)	(0.144)	(0.136)		
Exposure to Trained Workers (Share) × Untrained × Post	0.020	-0.089	0.028	-0.076		
	(0.410)	(0.374)	(0.411)	(0.379)		
Wage Band 2 Worker × Untrained × Post			0.010	0.007		
			(0.009)	(0.008)		
Bachelor + Degree × Untrained × Post			-0.002	-0.002		
			(0.010)	(0.009)		
Female × Untrained × Post			0.004	-0.006		
			(0.009)	(0.010)		
Worker with high performance × Untrained × Post			0.006	0.008		
			(0.010)	(0.009)		
Observations	12834	12834	12834	12834	12834	12834
R-squared	.904	.911	.904	.912	.903	.911
Worker Fixed Effects	✓	✓	✓	✓	✓	✓
Time Fixed Effects	✓	✓	✓	✓	✓	✓
Division Fixed Effects	×	✓	×	✓	×	✓
Sociodemographic Controls	×	×	✓	✓	✓	✓
LASSO	×	×	×	×	✓	✓

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

Table A4: Regressions of Frontline Worker Log Goal Achievement on Training and Coworker Exposure Controls

Note: This table displays estimates of training treatment effects when controlling for potential Stable Unit Treatment Value Assumption (SUTVA) violations. The column structure mimics Table 2. We add several connections measures for untrained workers to the model to assess whether any of these measures would materially change the main estimate of treatment effects. The model we estimate is

$$\log(y_{it}) = \beta_i + \beta_t \times \beta_d + \delta_1 Trained_i \times Post_t$$
$$+ \sum_c \delta_{uc}(1 - Trained_i) \times Post_t \times Exposure_{ic} + \varepsilon_{it},$$

where $\sum_c (1 - Trained_i) \times Post_t \times Exposure_{ic}$ is the sum over 4 measures of connections between focal untrained worker i and other frontline workers. These measures consist of the level and share of pre-period emails with trained and untrained frontline workers, respectively. In addition, we add measures of each untrained workers' connection strength to both trained and untrained workers using the levels and shares exposure measures. None of these spillover measures survive a LASSO variable selection procedure. Columns 5 and 6 report post-LASSO OLS regressions on the variables that do survive this variable selection procedure. Standard errors are clustered at the worker level.

Control For	(1) Horizontal	(2) Mean	(3) Imbalanced	(4) Sent	(5) Contemp.	(6) All Pre-Period Emails	(7) 90/10 Normalization	(8) Wage Band* Post
Panel A: Exposure Based on the Level of Emails with Eventually Trained Workers								
Exposure × Post	0.152**	0.168**	0.183**	0.187**	0.179**	0.169**	0.146**	0.168**
	(0.075)	(0.068)	(0.073)	(0.074)	(0.074)	(0.080)	(0.066)	(0.072)
N	3154	3154	3154	3154	3154	3154	3154	3154
R^2	.954	.954	.953	.953	.953	.953	.953	.953
Panel B: Exposure Based on the Share of Emails with Eventually Trained Workers								
Exposure × Post	0.040*	0.051**	0.043*	0.044*	0.048**	0.043*	0.031*	0.039*
	(0.023)	(0.022)	(0.023)	(0.024)	(0.024)	(0.023)	(0.016)	(0.023)
N	3154	3154	3154	3154	3154	3154	3154	3154
R^2	.954	.954	.953	.953	.953	.953	.953	.953
Division-Time Fixed Effects:	Yes	Yes	Yes	Yes	Yes	Yes	Yes	

Table A5: Analysis of Robustness of Vertical Training Spillovers

Note: The dependent variable is log goal achievement. Measures of email exposure to eventually trained workers are computed in the pre-training period. In Panel A, the exposure measures are based on the level of emails received from eventually trained workers. In Panel B, these measures are the based on the share of emails with eventually trained workers relative to all emails from workers who were eligible for training. Standard errors are clustered at the manager level. Column 1 controls for how horizontal spillovers to peers might affect our interpretation of vertical spillovers by including the interaction of managers' pre-period emails from untrained workers with these untrained workers' pre-period connections to trained workers, untrained workers, and managers. We then interact each these measures with the post-period dummy. In Panel A, untrained workers' own connection measures is based on pre-period email levels, while Panel B uses email shares. Column 2 accounts for the possibility of mean reversion and controls for deciles of the managers' pre-period productivity interacted with a post-period indicator. Column 3 controls for potential imbalances in treatment by including interactions of the exposure measure with indicators for female frontline workers and those with a college degree or more education. Column 4 controls for the normalized level of emails (or share of emails) sent to trained workers in the pre-period, interacted with the post-period indicator. Column 5 controls for contemporaneous weekly emails from untrained workers, which captures changes in workload that may result from exposure to trained workers. Finally, Column 6 controls for pre-period emails from all workers x Post. All columns include manager fixed effects and division x time fixed effects.

	(1)	(2)	(3)
Panel A: Exposure Based on the Level of Emails with Eventually Trained Workers			
Exposure × Post	0.135***	0.187**	0.177**
	(0.022)	(0.074)	(0.074)
Exposure (Sent Emails) × Post	-0.048*	-0.033	
	(0.027)	(0.031)	
Post			
Avg. Implied Change	2.5	4.42	5.05
N	3154	3154	3154
R^2	.952	.953	.953
Worker Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Division-Time Fixed Effects	No	Yes	Yes
Post-LASSO OLS	No	No	Yes
Panel B: Exposure Based on the Share of Emails with Eventually Trained Workers			
Exposure × Post	0.022*	0.044*	0.042*
	(0.013)	(0.024)	(0.023)
Exposure (Sent Emails) × Post	-0.003	-0.008	
	(0.023)	(0.018)	
Post			
Avg. Implied Change	.773	1.44	1.63
N	3154	3154	3154
R^2	.944	.953	.953
Worker Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Division-Time Fixed Effects	No	Yes	Yes
Post-LASSO OLS	No	No	Yes

Table A6: Estimates of Vertical Training Spillovers to Managers Using Additional Measures of Pre-Period Exposure to Eventually Trained Workers

Note: The dependent variable is log goal achievement. Measures of email exposure to eventually trained workers are computed in the pre-training period. In Panel A, the exposure measures are based on the level of emails received from eventually trained workers. In Panel B, these measures are based on the share of emails with eventually trained workers relative to all emails from workers who were eligible for training. Each panel includes two exposure measures. One based on the emails received by each manager (exposure). The alternative measure is based on the emails sent by each manager. Standard errors are clustered at the manager level. Columns 2 includes manager and time fixed effects, while column 3 includes manager and time-by-division fixed effects. The average percentage point change in goal achievement takes the predicted effects from the model and multiplies by the individual manager's average of pre-period goal achievement. The last column shows the coefficients that survive the Post-LASSO procedure.

	(1)	(2)	(3)	(4)
Log Emails From Workers to Managers	-0.089**	-0.091**	-0.088**	-0.091**
	(0.041)	(0.042)	(0.041)	(0.042)
Leave-Out-Week Email-Weighted Worker Log GA	-0.005	-0.055		
	(0.010)	(0.067)		
Transitory Δ in Leave-Out-Week Weighted Worker Log GA			0.103	-0.997
			(0.166)	(1.218)
N	1569	1569	1569	1569
R^2	.956	.958	.956	.958
Division-Time Fixed Effects:	No	Yes	No	Yes

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

Table A7: Regressions of Managers' Pre-Period Log Goal Achievement on Emails and Connected Worker Goal Achievement

Note: This table reports regressions of managers' log goal achievement in the pre-period on measures of managers' email volume and connected workers' goal achievement. The sample only contains the pre-period. "Log Emails From Workers to Managers" is the total number of emails sent from all frontline workers to manager i in week t. "Leave-Out-Week Email-Weighted Worker Log GA" captures connected workers' goal achievement, where connection weights come from email volume between worker j and manager i in all other weeks during the pre-period. An alternative measure, "Transitory Change in Leave-Out-Week Weighted Worker Log GA", uses the same connection weights for worker j and manager i but computes the deviation in productivity in week t relative to worker j's average productivity in all other weeks. All models include time fixed effects and manager fixed effects. Standard errors are clustered by manager.

	(1)	(2)	(3)
Panel A: Exposure Based on the Level of Emails Sent to Eventually Trained Workers			
Exposure × Post	0.038	0.038	0.036
	(0.030)	(0.030)	(0.028)
Post	-0.032**		
	(0.016)		
N	11295	11295	11295
R^2	.897	.897	.907
Worker Fixed Effects	✓	✓	✓
Time Fixed Effects	×	✓	✓
Division-Time Fixed Effects:	×	×	✓
Panel B: Exposure Based on the Share of Emails with Eventually Trained Workers			
Exposure × Post	0.039	0.039	0.029
	(0.034)	(0.034)	(0.031)
Post	-0.030*		
	(0.016)		
N	11295	11295	11295
R^2	.897	.897	.907
Worker Fixed Effects	✓	✓	✓
Time Fixed Effects	×	✓	✓
Division-Time Fixed Effects:	×	×	✓

Table A8: Estimates of Horizontal Training Spillovers to Untrained Workers

Note: The dependent variable is log goal achievement and the sample is restricted to untrained frontline workers. Measures of exposure to eventually trained workers are computed from emails in the pre-period. These measures are then normalized to a unit interval, ranging from a minimum value of zero (for the least exposed worker) to a maximum value of one (for the most exposed worker). In Panel A, the exposure measures are based on the level of emails sent from untrained to trained workers. Panel B uses email shares to calculate Exposure. Column 1 includes worker fixed effects, and Column 2 includes worker and time fixed effects, while column 3 includes worker and time-by-division fixed effects. The LASSO procedure does not select any independent variable.

Senders	Receipts:	Managers	Wage Band 1 Trained	Wage Band 1 Untrained	Wage Band 2 Trained	Wage Band 2 Untrained
Wage Band 1	Main Effect	-1.266	-0.318	-0.346	1.181	-0.372
	Trained Effect	-1.438	-0.419	-0.357	-0.305	-0.306
		(0.078)	(0.199)	(0.055)	(0.306)	(0.099)
Wage Band 2	Main Effect	-0.215	-0.222	-0.335	0.005	-0.329
	Trained Effect	-1.470	-0.200	1.054	0.001	-0.391
		(0.079)	(0.329)	(0.068)	(0.004)	(0.149)
Managers	Main Effect	-0.527	-0.166	-0.136	-0.043	-0.048

Table A9: Difference-in-differences Estimates of Changes in Log Emails

Note: This table displays coefficient estimates for the change in log emails between the pre- and post-periods by detailed sender-recipient type. Columns show the receivers while rows display the senders. Each regression is run separately for any sender type (Wage Band 1, Wage Band 2, Managers) and a receiver type (each of the 5 destinations in columns). The sample for each regression is an individual employee-level dyad. Standard errors are below the trained effect regression coefficient and are clustered at the recipient level (managers or workers). Pooling wage band 1 and 2 senders together generates the statistics that go into Figure 3.

	Model Object	Data Moment
1.	$\frac{1-e^{-\lambda z_m}}{1+H_{m,Pre}e^{-\lambda z_1}}$	Average Frontline GA (Pre-Period)
2.	$1-\frac{n_1}{n_m}\frac{h_m e^{-\lambda z_1}}{1+H_{m,Pre}e^{-\lambda z_1}}$	Average Manager GA (Pre-Period)
3.	$\frac{1+cz_m}{1+cz_1}$	Manager to Worker Ratio of Average Wages
4.	$w(1+cz_1)$	Average Frontline Worker Wages
5.	$w(1+cz_m)$	Average Manager Wages
6.	$\frac{1-e^{-\lambda_{Post}z_m}}{1+H_{m,Post}e^{-\lambda_{Post}z_1}} - \frac{1-e^{-\lambda z_m}}{1+H_{m,Pre}e^{-\lambda z_1}}$	Change in Average GA between Post- and Pre-Periods for Untrained, Unconnected Frontline Workers
7.	$\frac{(1-e^{-\lambda_{Post}z_m})(1-h_t\tilde{T})}{1+H_{m,Post}e^{-\lambda_{Post}z_1^t}}$	Average GA for Trained Workers (Post-Period)
8.	$(1-\rho)\frac{1-e^{-\lambda_{Post}z_m}}{1+H_{m,Post}e^{-\lambda_{Post}z_1}} + \rho\frac{1-e^{-\lambda_{Post}z_m}}{1+H_{m,Post}e^{-\lambda_{Post}z_1^t}+H_t e^{-\lambda_{Post}z_1}}$	Average GA (Post-Period) for Untrained, Connected Frontline Workers
9.	$1-h_m\tilde{h}_{Post}$	Average Manager GA (Post-Period)

Table A10: Moment Conditions Used in GMM Estimation

Note: This table presents theoretical moments from the model and corresponding data moments. The first two rows give average frontline worker goal achievement and manager goal achievement. $H_{m,Pre}$ is the wait time to receive help that is determined in equilibrium based on congestion. This is solved numerically with a fixed point problem, as $H_{m,Pre} = h_m(1+\tilde{h})/2$ is the expected wait time to get help when $\tilde{h}$ is the total help requests managers receive and h_m is the per-request time that managers take in responding. The Pre and $Post$ subscripts on the workers' wait time indicates that the degree of congestion will change as some workers' increasingly solve problems autonomously after training. Rows 3 - 5 provide moments on compensation. Row 6 targets changes in goal achievement for untrained, connected frontline workers. The parameter λ_{Post} allows the problem difficulty distribution to differ between periods, and $H_{m,Post}$ accounts for equilibrium congestion in getting help. Row 7 is the change in trained workers' goal achievement in the post-period, where problems are drawn from the distribution with parameter λ_{Post}. Trained workers field $\tilde{T}$ help requests at time-cost $h_t\tilde{T}$, which may reduce their own time for problem solving. The degree of help requests trained workers handle is determined by ρ, which enters row 8. With probability ρ, a connected, untrained frontline worker turns to a trained worker for help before going to a manager. If the trained worker can solve the problem, there is no need to escalate the problem further. The time cost of turning to the trained worker is $H_t = h_t(1+\tilde{T})/2$. Row 9 provides managers' post-period goal achievement, which is the residual time after all help requests are handled. In the post-period, each manager handles $\tilde{h}_{Post}$ requests which still take a per-request time cost of h_m.

Share of Help from Trained (ρ) :	.03	.06	.09	.12
	Panel A: Parameter Estimates			
z_1	0.187	0.188	0.180	0.196
	(0.043)	(0.050)	(0.041)	(0.047)
z_m	1.604	1.620	1.610	1.632
	(0.165)	(0.225)	(0.185)	(0.166)
c	1.393	1.376	1.365	1.388
	(0.218)	(0.189)	(0.173)	(0.165)
w	0.916	0.917	0.926	0.907
	(0.057)	(0.056)	(0.045)	(0.054)
λ	1.139	1.131	1.142	1.123
	(0.102)	(0.124)	(0.132)	(0.128)
h_m	0.102	0.102	0.102	0.104
	(0.012)	(0.012)	(0.012)	(0.012)
z_1^t	1.037	1.087	1.022	1.077
	(0.326)	(0.320)	(0.314)	(0.310)
h_t	0.054	0.047	0.003	0.004
	(0.032)	(0.035)	(0.028)	(0.035)
λ_{Post}	1.121	1.108	1.112	1.090
	(0.102)	(0.118)	(0.126)	(0.118)
	Panel B: Production Function Estimates and Help Request Changes			
a	0.294	0.296	0.296	0.297
	(0.034)	(0.037)	(0.037)	(0.038)
b	0.535	0.533	0.529	0.531
	(0.057)	(0.059)	(0.060)	(0.060)
Pct Change in Trained Help Req to Managers	-57.7	-59.5	-56.8	-57.8
	(12.4)	(12.5)	(12.0)	(11.7)

Table A11: GMM Estimates and Production Function Parameters Varying The Share of Help Requests Handled by Trained Workers

Note: This table presents sensitivity estimates using different values of ρ, the share of help requests handled by trained workers for connected, untrained workers. These estimates correspond to those in Panel A of Table 4 in the main text and the last row of Panel B. Output changes as implied by the various models are in

Share of Help from Trained (ρ) :	.03	.06	.09	.12
Panel A: Output in Levels Implied By First Order Conditions				
Q_{Pre} Without Training	64.08	64.24	63.27	64.21
	(12.08)	(9.31)	(10.27)	(12.43)
Q_{Post} Without Training	63.92	64.04	63.02	63.91
	(12.08)	(9.31)	(10.27)	(12.41)
Panel B: Output Change Decompositions				
Q_{Post} Pct Gains With Training and Spillovers	2.02	2.21	2.28	2.46
	(0.50)	(0.56)	(0.58)	(0.60)
Q_{Post} Pct Gains With 1 P.P. Vertical Spill	1.05	1.09	1.14	1.18
	(0.12)	(0.13)	(0.14)	(0.15)
Q_{Post} Pct Gains Without Vertical Spill	0.52	0.56	0.61	0.65
	(0.13)	(0.14)	(0.15)	(0.16)
Q_{Post} Pct Gains Without Horizontal Spill	1.90	1.96	1.87	1.92
	(0.48)	(0.51)	(0.50)	(0.50)
Panel C: Share of Gains from Vertical Spillovers				
Full Vertical Spillovers	74.48	74.84	73.35	73.75
	(4.20)	(4.48)	(4.37)	(4.62)
1 Percentage Point Vertical Spillover	51.01	48.98	46.63	45.24
	(8.58)	(8.28)	(8.32)	(8.38)
Panel D: Share of Gains from Vertical Spillovers With Constant Returns				
Full Vertical Spillovers	78.56	79.04	77.89	78.18
	(1.04)	(1.27)	(1.41)	(1.77)
1 Percentage Point Vertical Spillover	56.61	54.85	52.75	51.26
	(6.93)	(6.87)	(6.84)	(6.73)

Table A12: Changes in Output With and Without Spillovers Under Varied Levels of Help from Trained Workers

Note: This table uses the estimated production aggregator, with a and b as output elasticities estimated from the first order conditions, to display the output level, Q, in the pre-period (first row of Panel A) and changes in Q from the training program in the post-period. The first row of Panel B shows the value of Q after training with both vertical and horizontal spillovers as estimated from the post-period moments. The second row assumes vertical spillovers account for 1 percentage point of manager's gains, matching the conservative estimate from the reduced form, rather than managers' full gain. The third row shuts down vertical spillovers to managers. The fourth row shuts down horizontal spillovers to coworkers. Panel C displays the share of program gains due to vertical spillovers by comparing the change in Q with and without the spillover to managers. Panel D displays the share of program gains from vertical spillovers when the production function is assumed to have constant returns to scale, where $Q = (n_1\phi)^a(n_mS)^{1-a}$ and a is estimated from a modified set of first order conditions. Standard errors are estimated from 150 bootstrap iterations of the entire estimation procedure.

Panel A: Parameter Estimates	
z_1	0.090
	(0.008)
z_2	0.498
	(0.041)
z_m	1.793
	(0.146)
c	1.365
	(0.114)
λ_1	1.100
	(0.091)
λ_2	1.034
	(0.077)
h_m	0.124
	(0.009)
w	0.869
	(0.070)
Panel B: Production Function Estimates	
a_1	0.408
	(0.098)
a_2	0.253
	(0.092)
Panel C: Gains from Spillovers	
Share of Gains from 1 pp Vertical Spillover	0.513
	(1.996)

Table A13: GMM Estimates and Production Function Parameters with 2 Layers of Production Workers and 1 Manager Layer

Note: This table presents GMM estimates using moment conditions for an organization with 2 layers of frontline workers. Each layer of frontline workers has their own problem distribution, with parameters $_1$ and λ_2. The estimator here does not use post-period moments. Production function estimates are taken from the first-order conditions derived in the appendix. The parameter $_1$ is the output elasticity for wage band 1's tasks, $_2$ is the output elasticity for wage band 2's tasks, and is $1 - {}_1 - {}_2$ for managers' strategic tasks. Standard errors are bootstrapped using 150 iterations.

	Untrained Mean (SD)	Trained Mean (SD)	Difference (SE)
Increase Goal Understanding	0.105	0.212	0.107
	(0.310)	(0.415)	(0.083)
Directed to Reduce Help Requests to Managers	0.018	0.030	0.013
	(0.132)	(0.174)	(0.035)
Increase Promotion Probability	0.088	0.091	0.003
	(0.285)	(0.292)	(0.063)
Increased Knowledge of Task Requirements	0.053	0.879	0.826***
	(0.225)	(0.331)	(0.065)
Increased Understanding of Division-Appropriate Work	0.088	0.818	0.730***
	(0.285)	(0.392)	(0.078)
Increase Skills and Knowledge	0.035	0.909	0.874***
	(0.186)	(0.292)	(0.056)
Increase Interdependent Tasks	0.053	0.061	0.008
	(0.225)	(0.242)	(0.052)
Worked More Hours	0.053	0.061	0.008
	(0.225)	(0.242)	(0.052)
Number of individuals	57	33	

Table A14: Survey Evidence on Differences in Perceived Changes Between Trained and Untrained Frontline Workers

Note: The table shows differences and t-tests between trained and untrained workers' responses to survey questions on changes in their work environment between the pre- and post-periods. The survey was conducted in 2020 and had nine questions that each began with "Relative to 2018, in 2019 you:". These questions were then: 1) Improved your understanding of how goals are set and how they are evaluated weekly? 2) Were told explicitly that you should ask for help from colleagues and peers and rather than managers? 3) Increased your probability of promotion inside the organization? 4) Improved your ability to distinguish if tasks and projects require large or small knowledge that is specific to your division? 5) Improved your ability to recognize if the tasks and projects require the knowledge from your division or different divisions? 6) Increased the knowledge and the skills required to satisfactorily achieve goals? 7) Received a larger number of across-divisions, interdependent tasks. 8) Worked a larger number of hours a week? Respondents could choose three option answers: Yes, No, Does not apply/Do not know.

Appendix Figures

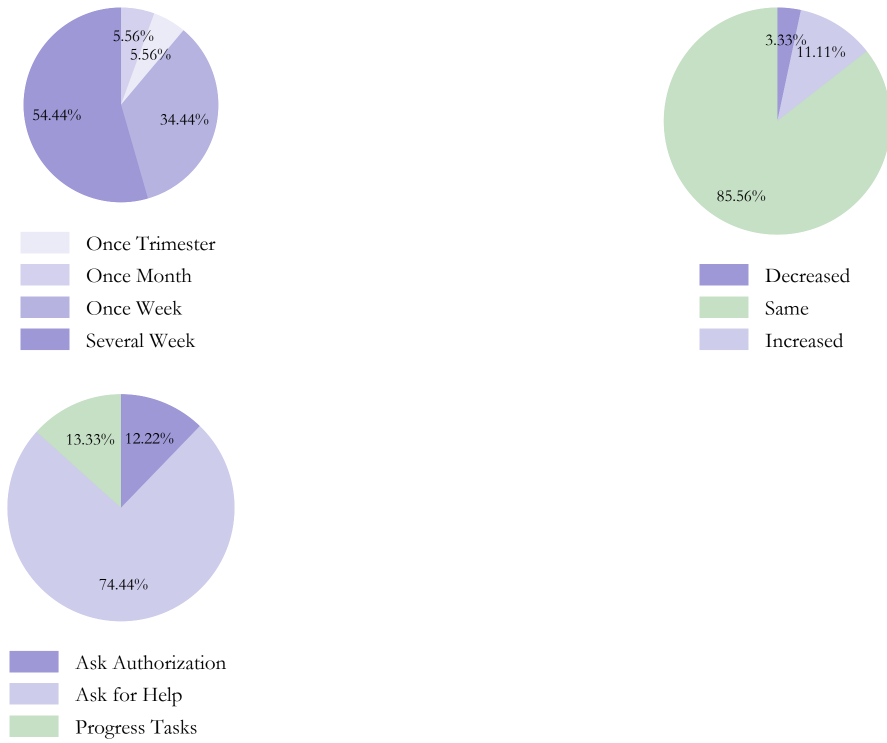
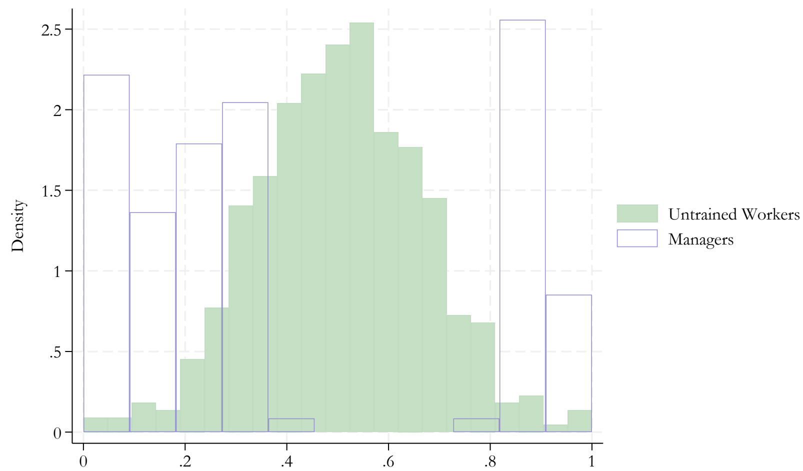
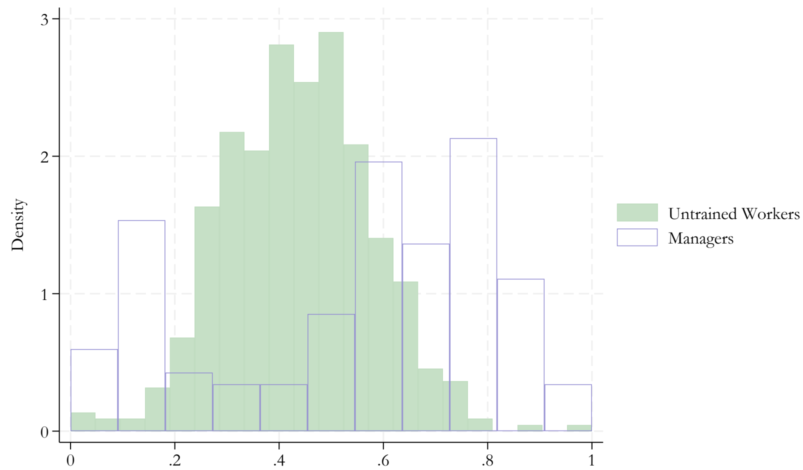


Figure A1: Distribution of Survey Responses to Questions Regarding the Mechanism

Note: This figure displays answers to an ex-post survey conducted in Fall of 2020 that was designed to understand the mechanisms behind our results. From top to bottom and left to right, the questions are as follows: 1. "Remember your work environment in 2018 and 2019. Consider all the people you interacted with by email every week. How frequently did you interact with them face to face? (choose only one option)." 2. "In your opinion, relative to 2018, monitoring from your managers in 2019 increased, decreased, or remained the same?". 3. "Remember your work environment in 2018 and 2019. What was the main reason that you emailed workers from a higher wage band (choose only one option)."



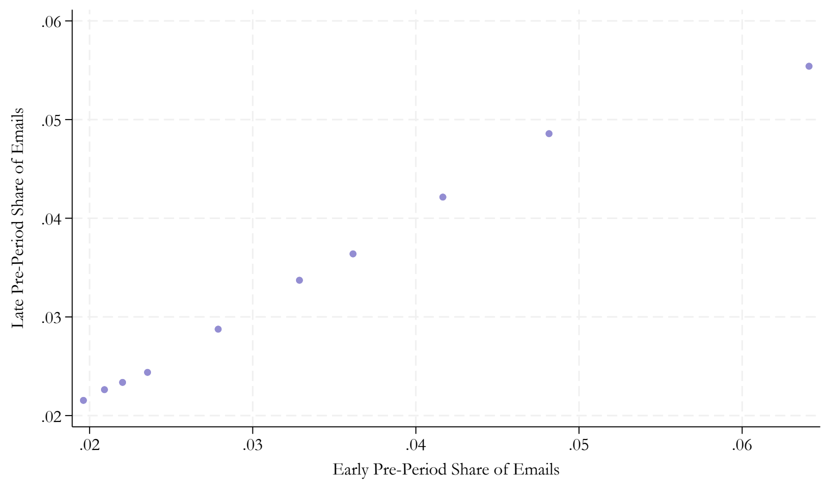
(a) Exposure Based on the Level of Pre-Period Emails from Trained Workers



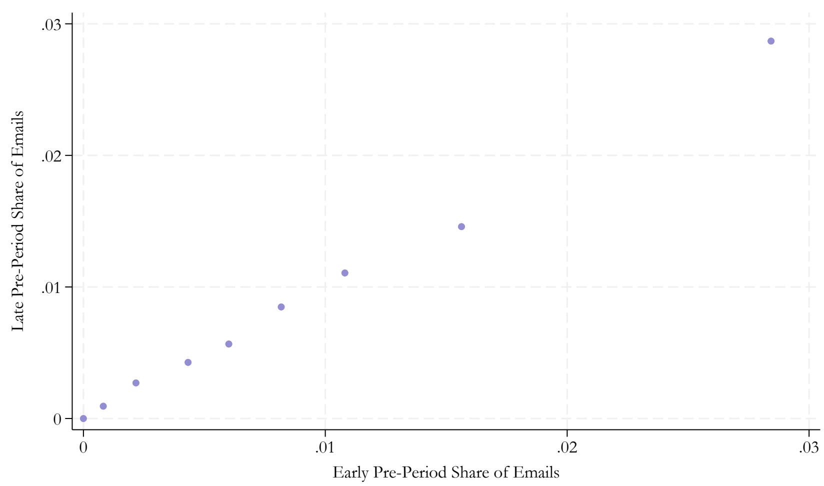
(b) Exposure Based on the Share of Pre-Period Emails from Trained Workers

Figure A2: Distribution of Exposure Measure to Trained Workers Based on the Level or Share of Emails Received from from Eventually Trained Workers in the Pre-Period

Note: The top figure shows a histogram of the transformed exposure measure that is calculated from the level of pre-period emails that managers receive from trained workers (which we interpret as help requests incoming to managers) and emails sent to trained workers by untrained workers (which is a proxy for outgoing help requests). The transformation is defined in equation (1), where the least exposed individual has a value of 0 and the most exposed a value of 1. The figure plots separate histograms for untrained workers and managers. The bottom figure shows the histogram when email shares are used to construct the exposure measure.



(a) Emails from Workers to Managers



(b) Emails from Workers to Workers

Figure A3: Persistence of Email Connections Between the First and Last Month of the Pre-Period

Note: This figure displays the share of emails sent in worker-manager dyads (Panel A) or worker-worker dyads (Panel B) in the first 4 weeks of the pre-period and the last 4 weeks of the pre-period. There is a 4 week gap between these periods.

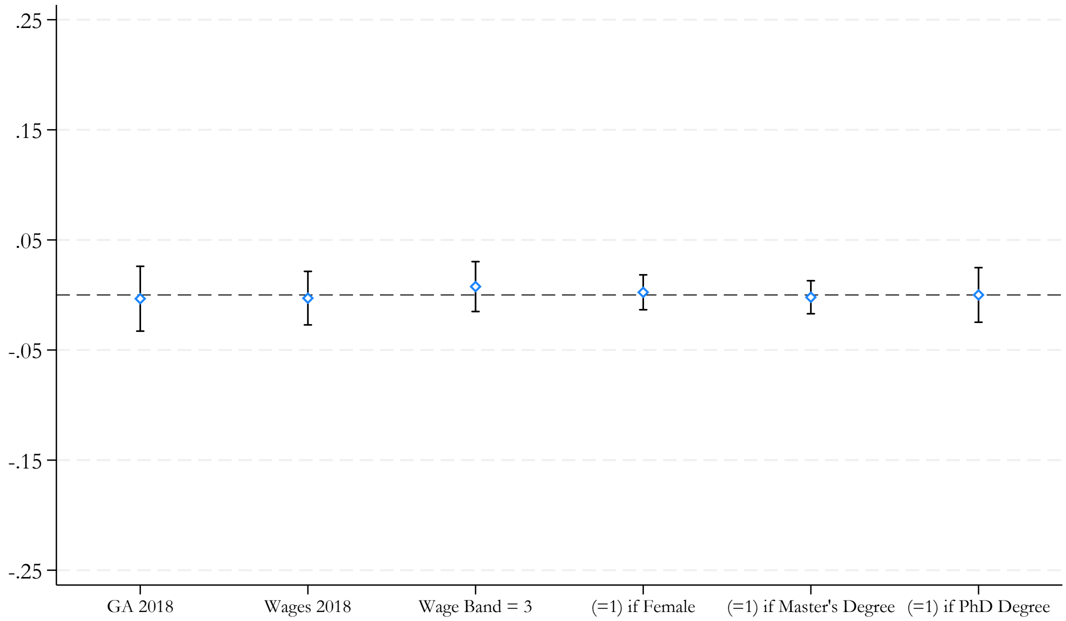
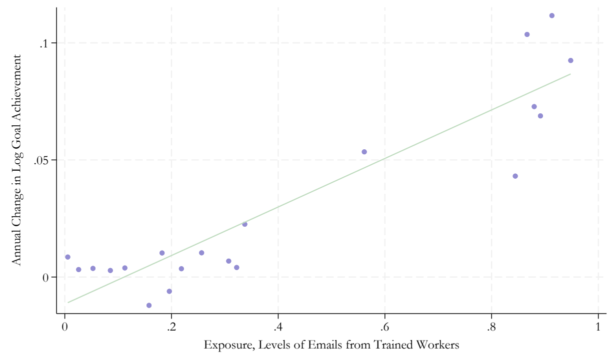
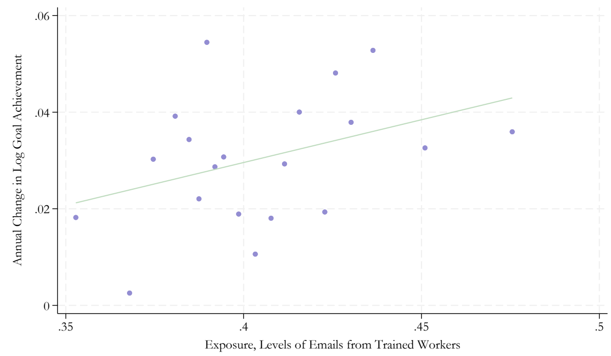


Figure A4: Predictors of Manager Exposure to Trained Workers

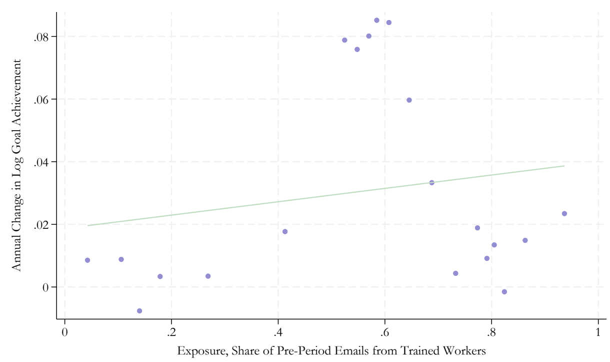
Note: This Figure displays a plot of regression coefficients and confidence intervals that test whether manager characteristics predict their exposure to trained workers. The coefficient plot comes from the regression $Exposure_{Levels} = \beta_1 GoalAchievement_{2018,i} + \beta_2 Wages_{2018,i} + \beta_3 WageBand3_i + \beta_4 Female_i + \beta_5 Master_i + \beta_6 PhD_i + Division_i + u_i$. The unit of observation is a manager. The joint test has F(6,118) = 0.17 (p-value = 0.985).



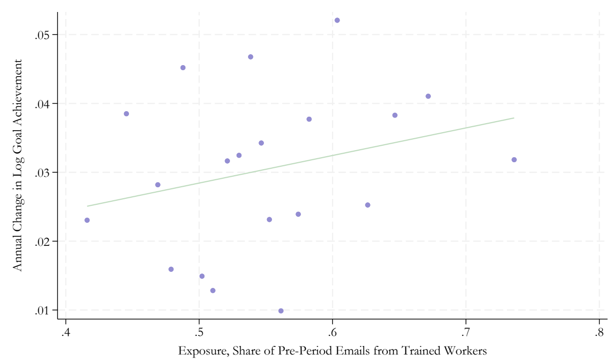
(a) Email Levels, No Fixed Effects



(b) Email Levels, Division Fixed Effects



(c) Email Shares, No Fixed Effects



(d) Email Shares, Division Fixed Effects

Figure A5: Evidence of Monotonicity for Vertical Spillovers Estimates: Annual Goal Achievement Changes for Managers based on Exposure Measures

Note: This figure shows the relationship between the exposure variable and the annual change in log goal achievement. The top panel use the level of emails as the exposure variable while the bottom panel use the share of emails. Panels (a) and (c) plot the binscatter of this relationship when we do not control for division fixed effects. Panels (b) and (d) control for Division fixed effects.

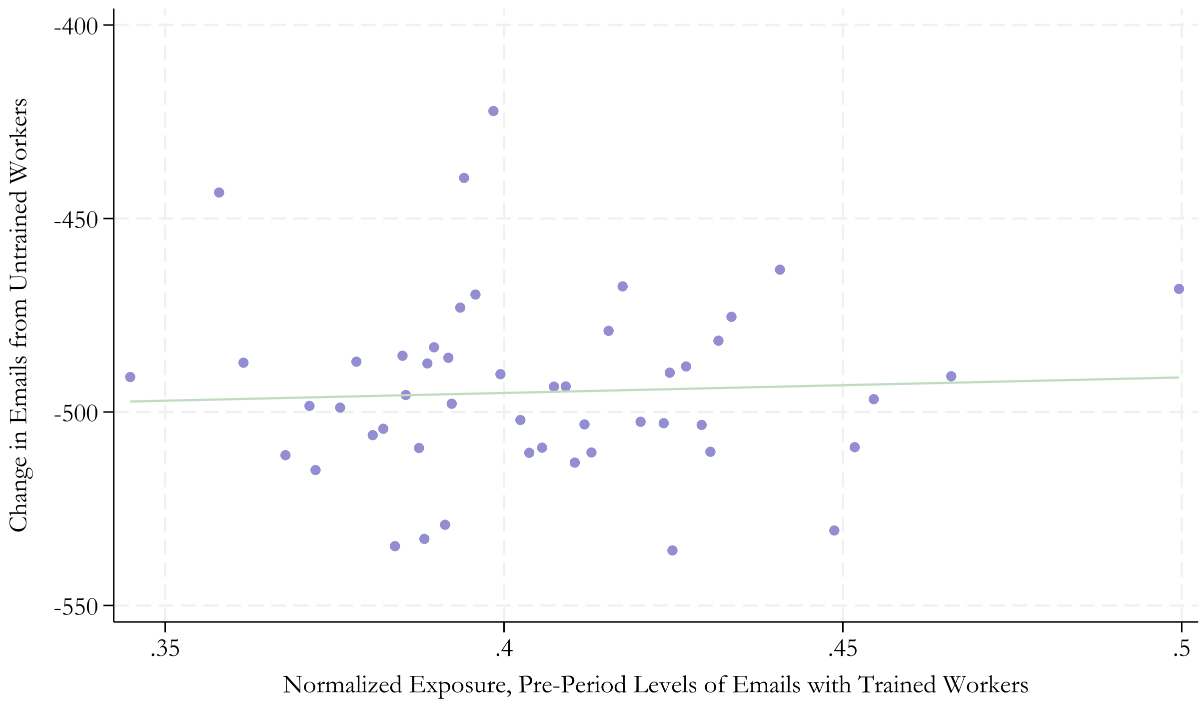


Figure A6: Tests for Manager SUTVA Violations Based on Changes in Emails with Untrained Workers

Note: A potential SUTVA violation is that managers with more exposure to trained workers may change their communication patterns with untrained workers (e.g. untrained workers seek help from them because they are less busy). This figure shows how changes in the average yearly number of emails vary with respect to managers' exposure to trained workers, and we net out division fixed effects to capture potential rebalancing of workloads within division. The y-axis is the normalized change in log emails between the pre- and post-periods and the x-axis is exposure based on levels of emails with trained workers. The regression coefficient and standard error (N=129) is 40.11 (98.35) .

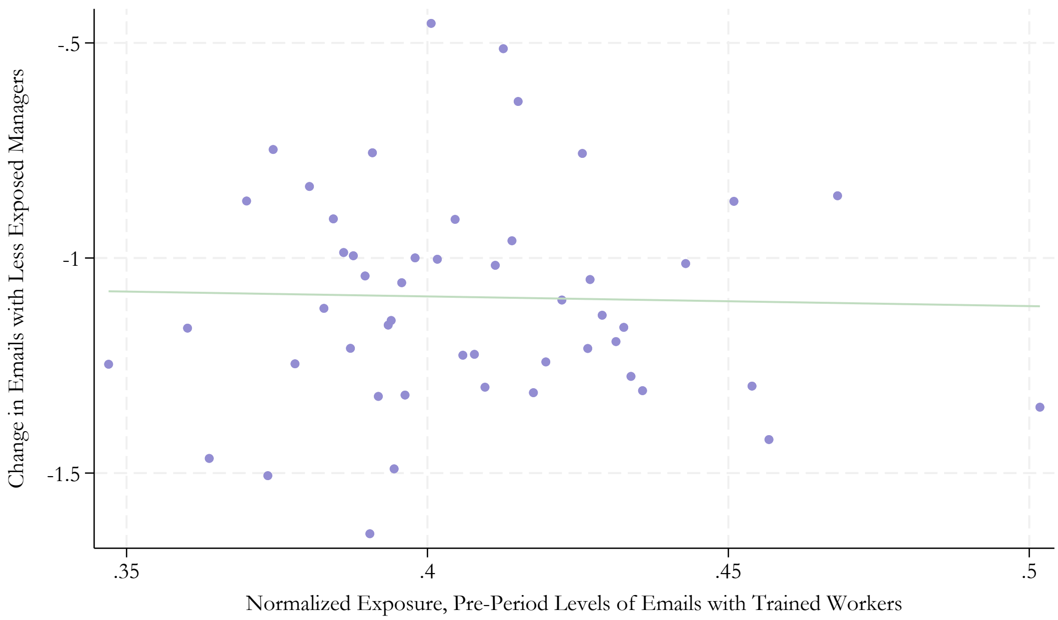
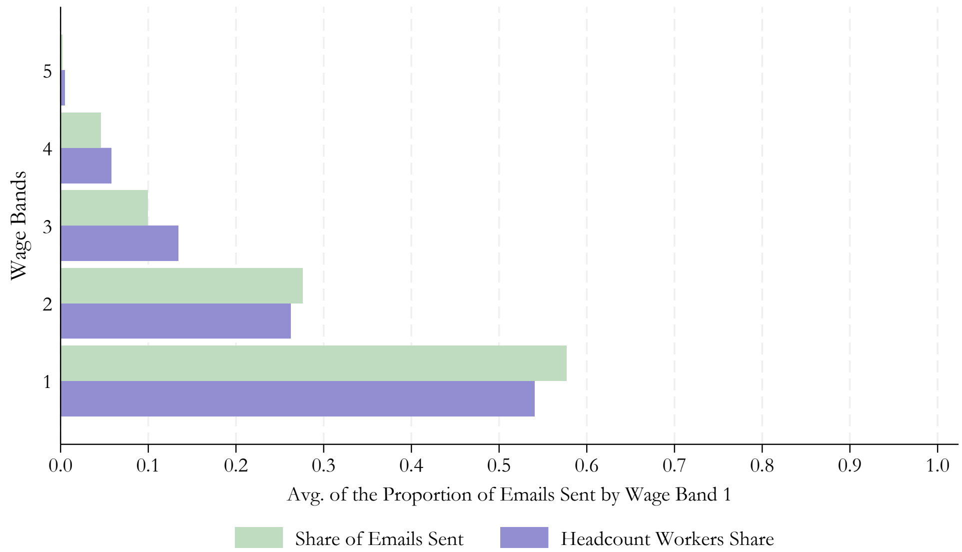
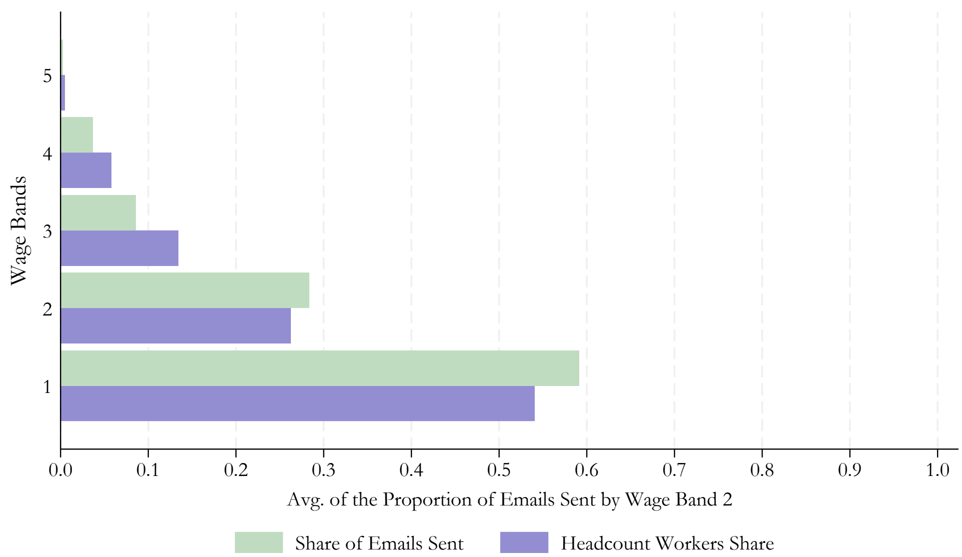


Figure A7: Tests for Manager SUTVA Violations Based on Changes in Emails with Other Managers

Note: This figure displays the relationship between the normalized exposure to trained workers (based on the normalized number of pre-period emails with trained workers) and the average yearly number of emails with managers who are less exposed to trained workers. Less exposed managers have a below-median number of pre-period emails with trained workers. The sample is manager-dyads (N=8192) and we net out division fixed effects. The regression coefficient is -0.008 (0.038).



(a) Email/Headcount Shares from the Perspective of Wage Band 1 Workers

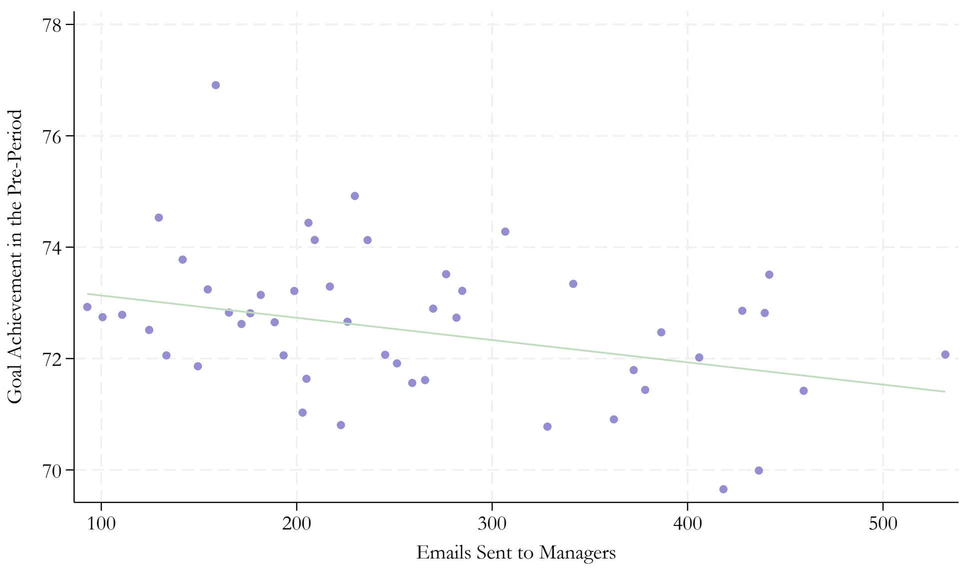


(b) Email/Headcount Shares from the Perspective of Wage Band 2 Workers

Figure A8: Email Communication Patterns Relative to Headcount in the Pre-Period

Note: This figure shows the distribution of email communications and headcount shares by wage band for workers in the pre-period (2018). Panel (a) shows the shares of emails relative to headcount shares from the perspective of wage band 1 workers. Panel (b) shows the same plot from the perspective of wage band 2 workers.

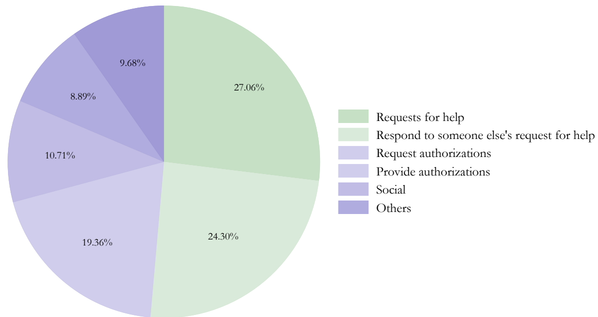
(a) Emails Sent to Other Frontline Workers



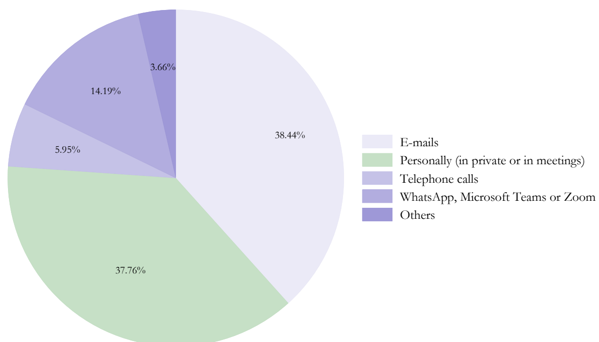
(b) Emails Sent to Managers

Figure A9: Relationship Between Frontline Workers' Goal Achievement in the Pre-Period and Emails Sent to Other Frontline Workers and Managers

Note: This figure displays the relationship between average weekly goal achievement and the number of emails sent to other frontline workers in wage bands 1 and 2 during the pre-period and Managers.



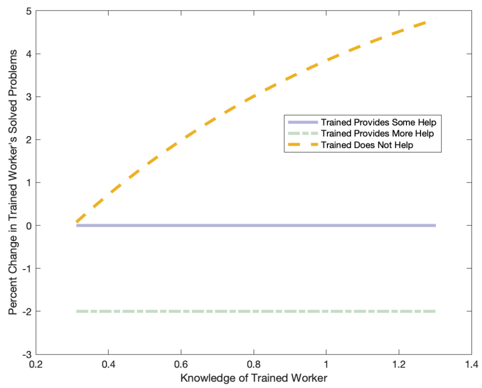
(a) What proportion of your outgoing emails would you place in the following categories?



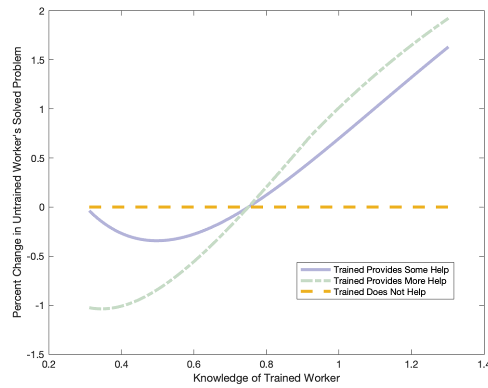
(b) What proportion of your work-related questions do you ask your boss or immediate superior using the following methods?

Figure A10: Questions from a 2024 Survey on Email Use and Help Requests

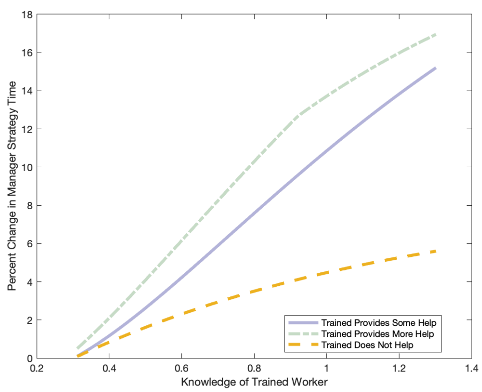
Note: The figure displays answers to survey questions 5 and 6 (in the figure subtitles) from a survey conducted in 2024. There were 22 responses after the survey was sent to a specific group of 65 workers in the agency.



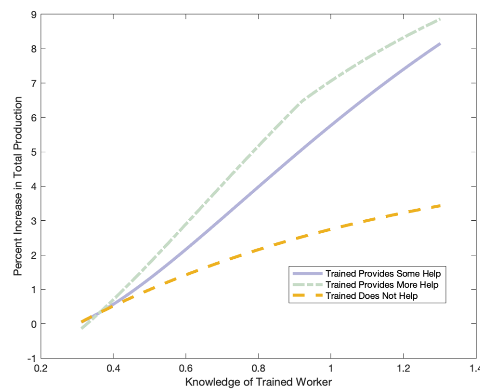
(a) Trained Solutions



(b) Untrained Solutions



(c) Manager Time for Strategy



(d) Total Output

Figure A11: Theoretical Results When Co-Worker Helping Costs Fall

Note: These figures are generated using the same setup as in Figure 4, but with lower relative costs to get help from a trained worker. In this figure, $H_t = 0.4 \times H_m$.

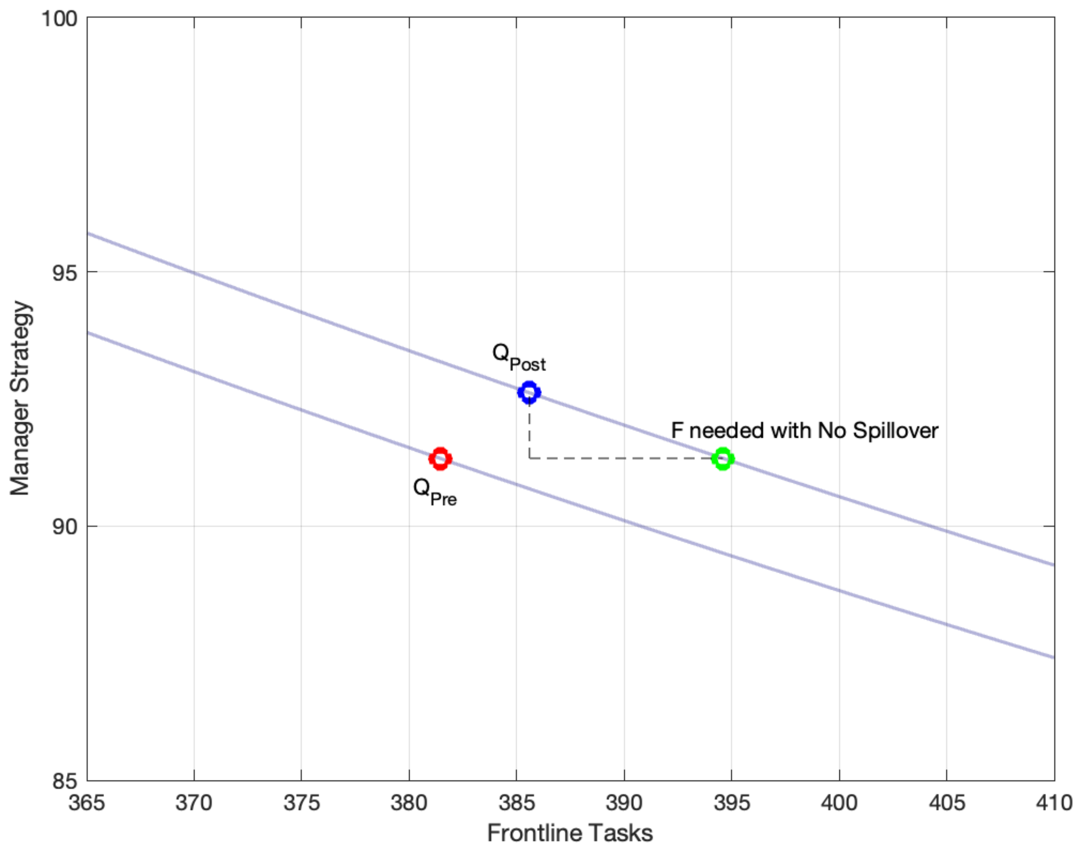
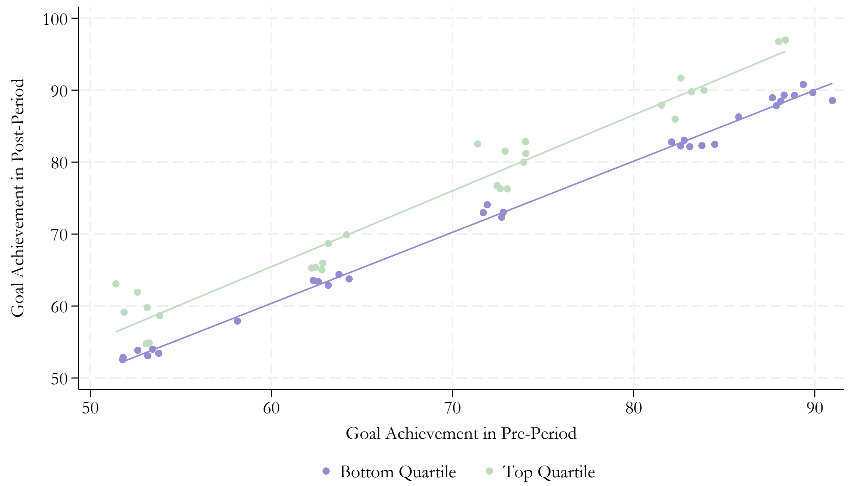


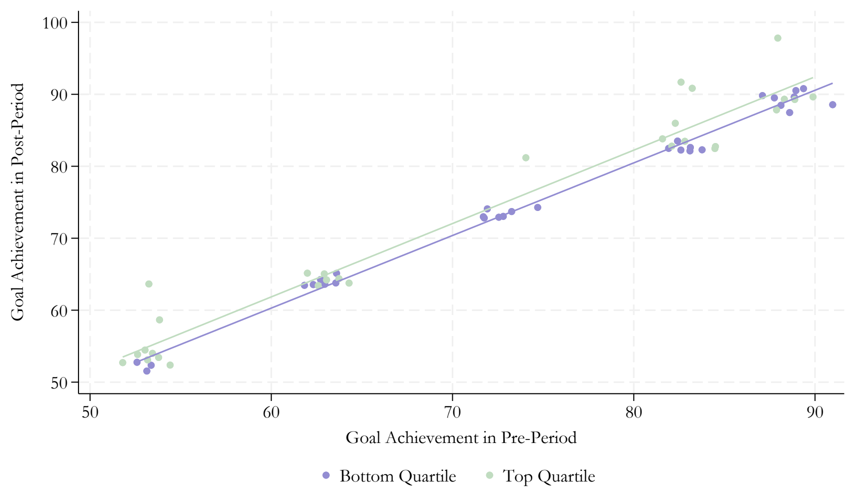
Figure A12: Illustration of Output Changes With and Without Spillovers

Note: This figure plots isoquants illustrating the organization's output in the pre-period and the post-period using observed values of the data. The point labeled "F needed with No Spillovers" illustrates the intuition for assessing how many more frontline workers would need to be trained in the absence of spillovers to managers in order for the organization to remain on the same post-period isoquant.

B Appendix: Multiple Frontline Layers



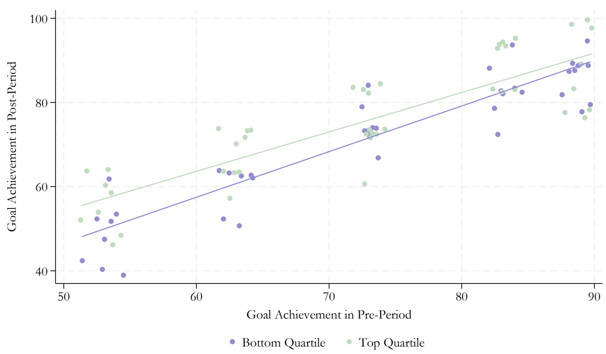
(a) Strength of Connections to Wage Band 1



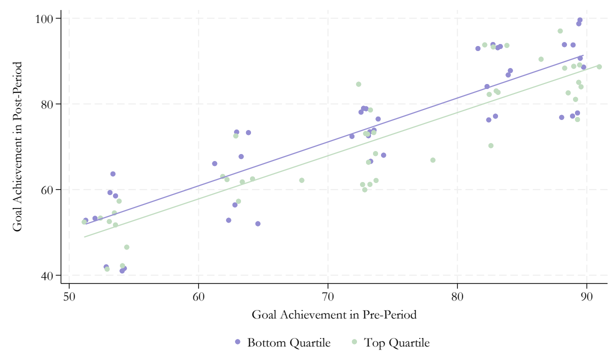
(b) Strength of Connections to Wage Band 2

Figure B1: Goal Achievement Changes For Managers by Connection Strength to Trained Workers

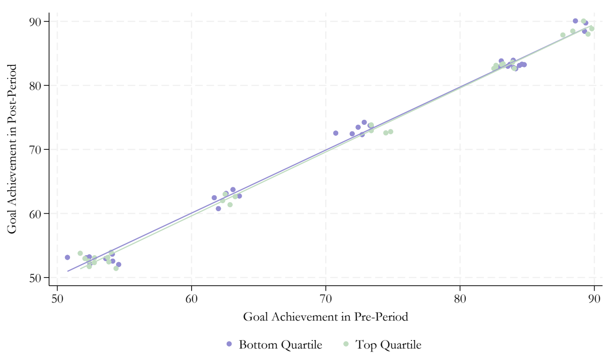
Note: These figures plot the relationship between average post-period and average pre-period goal achievement for individual managers. Each figure separately plots relationships by quartiles of managers' connection strength. Panel (a) considers connections with trained wage band 1 workers. Panel (b) considers connections with trained wage band 2 workers. Note that there are fewer trained wage band 2 workers overall, and this plot captures aggregate connection strength to all trainees, which may be different than the impact of per-capita connections with an individual trained worker. Connections to wage band 2 trained workers also matter, but the interquartile differences are smaller. One possibility is that connections to wage band 2 workers provide less total time savings for managers, as there are only 18 trained wage band 2 workers compared to 45 trained wage band 1 workers.



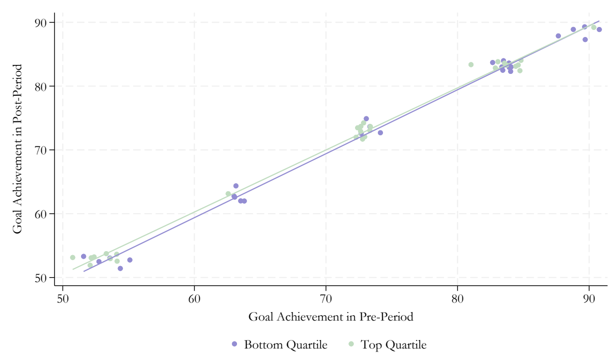
(a) Untrained WB1 to Trained WB2



(b) Untrained WB1 to Trained WB1



(c) Untrained WB2 to Trained WB1



(d) Untrained WB2 to Trained WB2

Figure B2: Goal Achievement Changes For Untrained Workers by Connection Strength to Trained Workers. Figures Are Split by Wage Band 1 (WB1) and Wage Band 2 (WB2) Workers and Connections

Note: This figure plots the relationship between average post-period and average pre-period individual goal achievement for untrained workers. We plot the relationship separately based on the strength of an untrained worker's connections to eventually trained workers in each wage band. Connection strength is based on the total number of emails with eventually trained workers in the pre-period.

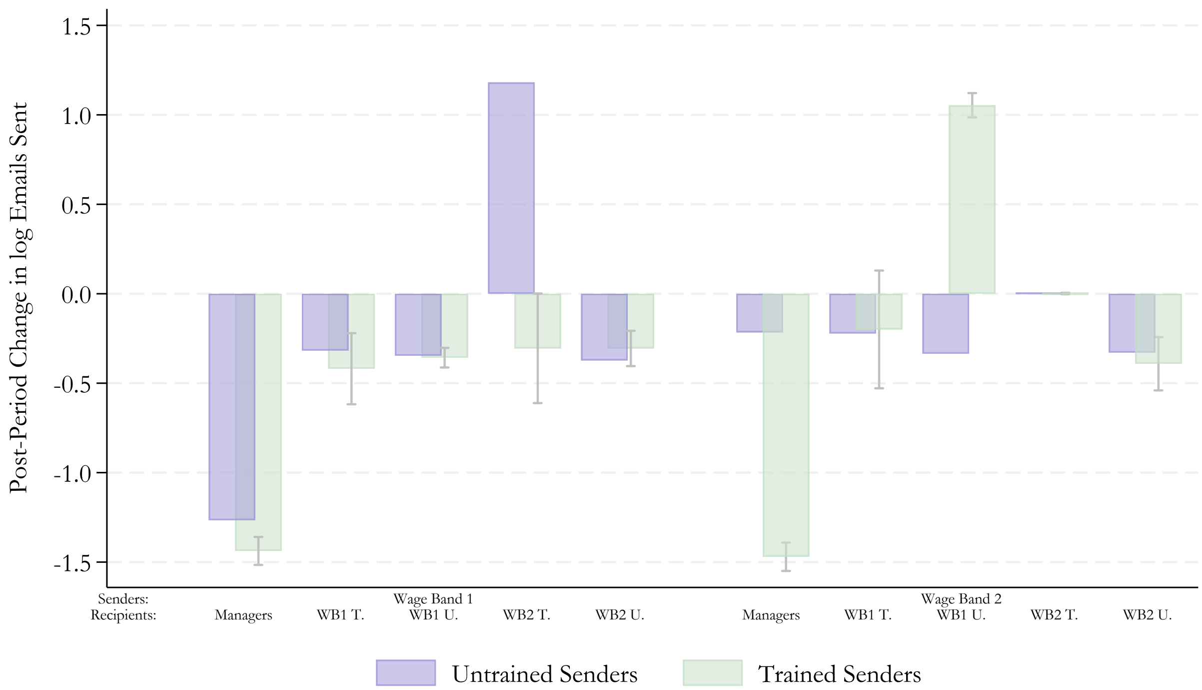


Figure B3: Changes in Log Emails Between the Pre- and Post-Period by Sender/Recipient Group

Note: This figure displays the average change in log emails between the pre- and post-periods by senders' wage band, recipient groups, and training status. Purple bars are the average changes in emails from untrained workers to each recipient group. Recipient groups are Managers, Wage Band 1 Trained Workers (abbreviated as WB1 T.), Wage Band 1 Untrained Workers (WB1 U.), Wage Band 2 Trained Workers (WB2 T.), and Wage Band 2 Untrained Workers (WB2 U.). Green bars are the change for trained workers, with standard errors computed from a difference-in-differences regression of log emails on a post-period-by-trained dummy. The regression are run by origin (wage band 1 or 2) and recipient group and include fixed effects for workers and time. Standard errors are clustered by sender.

	(1)	(2)	(3)
Panel A: Exposure Based on the Level of Emails Sent to Eventually Trained Workers			
Exposure WB1 to WB1 × Post	0.028	0.028	0.021
	(0.057)	(0.057)	(0.049)
Exposure WB2 to WB1 × Post	0.055	0.055	0.058
	(0.047)	(0.047)	(0.043)
Exposure WB1 to WB2 × Post	0.062	0.062	0.087**
	(0.040)	(0.040)	(0.040)
Exposure WB2 to WB2 × Post	0.038*	0.038*	-0.019
	(0.023)	(0.023)	(0.035)
Post	-0.069*		
	(0.040)		
N	11295	11295	11295
R^2	.898	.898	.908
Worker Fixed Effects	✓	✓	✓
Time Fixed Effects	×	✓	✓
Division-Time Fixed Effects:	×	×	✓
Panel B: Exposure Based on the Share of Emails with Eventually Trained Workers			
Exposure WB1 to WB1 × Post	0.021	0.021	0.008
	(0.057)	(0.057)	(0.049)
Exposure WB2 to WB1 × Post	0.052	0.052	0.051
	(0.047)	(0.047)	(0.044)
Exposure WB1 to WB2 × Post	0.054	0.054	0.072*
	(0.040)	(0.040)	(0.040)
Exposure WB2 to WB2 × Post	0.038	0.038	-0.023
	(0.026)	(0.026)	(0.037)
Post	-0.063		
	(0.042)		
N	11295	11295	11295
R^2	.898	.898	.908
Worker Fixed Effects	✓	✓	✓
Time Fixed Effects	×	✓	✓
Division-Time Fixed Effects:	×	×	✓

Table B1: Estimates of Horizontal Training Spillovers to Untrained Workers by Wage Band

Note: The dependent variable is log goal achievement and the sample consists of untrained workers. Measures of exposure to eventually trained frontline workers in the two wage bands are computed from emails in the pre- period. These measures are then normalized to a unit interval, ranging from a minimum value of zero (for the least exposed worker) to a maximum value of one (for the most exposed worker). In Panel A, the exposure measures are based on the level of emails sent from trained wage band 1 or wage band 2 workers. Panel B uses email shares to calculate Exposure. Effects are allowed to vary based on whether the focal worker is in wage band 1 or 2. For example, "Exposure WB1 to WB2 × Post" means the focal worker is in wage band 1 and the exposure measure is based on connections to trained wage band 2 workers. Column 1 includes worker fixed effects, Column 2 includes worker and time fixed effects, and column 3 includes worker and time-by-division fixed effects.

C Theoretical Details

C.1 The Caliendo and Rossi-Hansberg (2012) and Gumpert, Steimer, and Antoni (2022) Background Models

Setup: To give readers background, we describe the setup of traditional models of hierarchy (based on Gumpert et al. (2022), which implements a "cumulative knowledge" version of Caliendo and Rossi-Hansberg (2012)). This approach characterizes an organization's choice of layers, workers, and their knowledge in response to a given level of demand.[35]

In this framework, the organization aims to minimize the knowledge acquisition cost required to solve Q problems, where Q represents total demand. It makes three key decisions to meet this demand: it selects the number of hierarchical layers, L, it chooses workers' knowledge in each layer, z_l for $l \in \{1, ..., L\}$, and it allocates the number of workers in each layer, n_l. Problems arrive from an exponential distribution. A worker at layer l can autonomously solve any problem whose difficulty does not exceed their knowledge z_l. If $z > z_l$, the worker escalates the problem to the next layer, which requires h units of helping time from the receiving worker. The most difficult problem the organization can solve is determined by the knowledge of the highest-layer worker, z_L. The organization's problem can be solved in 2 steps. Given a fixed number of layers L, choose n_l and z_l in each layer to minimize costs subject to output and time constraints. Second, search over L to find the optimal number of layers that minimizes total costs.

The formal cost-minimization problem given an exponential distribution of problems (with parameter λ) for a L-layer firm follows as

$$\min_{\{z_l, n_l\}_{l=1}^{L}} \sum_{l=1}^{L} n_l \times w(1 + c \times z_l) \tag{6}$$

[35] While the original papers in this literature, like Garicano (2000) and Garicano and Rossi-Hansberg (2004), use a profit maximization setup that is useful for proving properties of optimality, the cost minimization approach has advantages for working with data on organizations with a fixed number of layers. Cumulative knowledge means that to solve a problem with difficulty level z, workers must know how to solve all problems less difficult than z.

subject to

$$
\begin{aligned}
n_1 \times (1 - e^{-\lambda z_L}) &\geq Q \\
n_L = 1 = n_1 \times h \times e^{-\lambda z_{L-1}} & \\
n_l = n_1 \times h \times e^{-\lambda z_{l-1}} \text{ for } 1 < l < L &
\end{aligned}
\tag{7}
$$

The first constraint says that the n_1 problems drawn, and the share of problems ultimately solved, must be large enough to meet demand, Q. The second constraint says there is 1 person at the top of the organization and all problems that get passed up must be addressed, implying that $n_1 = e^{\lambda z_{L-1}}/h$. The third line is the time constraint for helpers below layer L. Plugging n_1 into this constraint yields $n_l = e^{\lambda(z_{L-1} - z_{l-1})}$ for $1 < l < L$.

Mapping to the Experiment: Although this model captures some features of the agency we study, its predictions about training effects – particularly the impact of increasing lower-level workers' knowledge on higher-level workers' performance – do not fully align with our stylized facts. In a 2-layer hierarchy, z_2 is the top manager's knowledge and z_1 is frontline worker's knowledge. In a 3-layer hierarchy, z_3 is the top manager's knowledge, z_2 is the knowledge of n_2 workers who only help workers in layer 1, and z_1 is the knowledge of n_1 layer 1 workers who draw problems. If considering a two-layer hierarchy, the experimental variation in our setting would be analogous to increasing a subset of the z_1s while holding fixed n_1, and z_2. If considering a hierarchy with 3 or more layers, the experimental variation is analogous to increasing a subset of z_1s and z_2s while holding fixed headcount and knowledge of all other workers in higher layers.

To study how these experimental changes map to concepts that we can measure, we need a notion of goal achievement for each individual. Define goal achievement of a worker in layer 1 as the share of problems they can solve autonomously. For layer $l > 1$ workers, define goal achievement as the share of problems solved conditional on the problem reaching that layer.[36]

With these definitions, we study what happens in our experiment. We make clear here that we are studying partial-equilibrium changes that the organization made through the

[36] That is, if a level-1 worker has knowledge level z_1, goal achievement is $1 - e^{-\lambda z_1}$. Given the memoryless property of the exponential distribution, goal achievement in layer $l > 1$ is $1 - e^{-\lambda(z_l - z_{l-1})}$.

training program. We always hold fixed n_l for each l, and we consider comparative statics when knowledge (z_l) increases for a subset of the frontline workers (layer 1 in the 2-layer hierarchy or layers 1 and 2 in the multi-layer hierarchy). We always hold fixed the values of z in higher layers. For example, in a hierarchy with 3 or more layers, z_k is fixed for $k \geq 3$.[37] Applying this partial-equilibrium logic yields the following predictions, some of which are inconsistent with the effects we estimate from the introduction of the training program. These are:

1. An increase in z_k increases goal achievement for workers in layer k.

2. An increase in z_k weakly reduces goal achievement in all higher layers ($l > k$) when z_l is held fixed. This is because increasing z_k makes the problems that get passed upward harder, on average, reducing higher-layer workers' propensity to solve them. To see this, consider an increase in z_1. Then $\frac{\partial 1-e^{-\lambda(z_2-z_1)}}{\partial z_1} < 0$ and $\frac{\partial 1-e^{-\lambda(z_3-z_2)}}{\partial z_1} = 0$. However, workers' time constraints in layers $l > k$ may no longer bind, as lower-level workers would be able to handle more problems autonomously, freeing up manager time. In equilibrium, the organization would adjust, but the mapping to our experiment shuts down that adjustment.

3. Increasing z_1 while holding fixed n_1, n_l, and z_l for $l > 1$ does not change aggregate output. To see this, the first order condition with respect to the Lagrange multiplier, μ, is $Q = \frac{e^{\lambda z_{L-1}}}{h}(1 - e^{-\lambda z_L})$. Differentiating again shows that $\frac{\partial Q}{\partial z_1} = 0$.

Items 2 and 3 are at odds with the main estimates from the experiment. In particular, we show that training frontline workers weakly increases output for connected workers in higher layers. In addition, if the definition of each layer is based on a unique partition of knowledge-levels, then changes in the knowledge of wage band 1 workers propagates up beyond just wage band 2 workers in the organization. In fact, wage band 1 and wage band 2 workers both draw production problems – whereas in the base model, all production problems originate from the workers with the lowest knowledge levels (and wages). As such, we require some minor modifications to study the consequences of the program.

[37] The ability for the organization to fully adjust may potentially weaken or reverse these predictions, but they are useful nonetheless for understanding how our setting potentially does or does not map to the standard model.

C.2 Proofs

Proof of Proposition 1. To show 1), re-write the manager's time constraint to yield an expression for time spent on strategic tasks when 1 worker becomes trained, S^{Train_1} as

$$S^{Train_1} = 1 - (n^* - 1) \times h_m \times \frac{e^{-\lambda z_1^*}}{p + H_m \times e^{-\lambda z_1^*}} - h_m \times \frac{e^{-\lambda z_1^t}}{p + H_m \times e^{-\lambda z_1^t}}. \tag{8}$$

To show that $S^{Train_1} > S^*$ when 1 frontline workers becomes more knowledgeable, it is sufficient to show that the last expression in (8) is declining in z_1, or $\frac{\partial \frac{e^{-\lambda z_1}}{p+H_m \times e^{-\lambda z_1}}}{\partial z_1} = \frac{-\lambda p e^{-\lambda z_1}}{(p+H_m \times e^{-\lambda z_1})^2} < 0$, which establishes that $S^{Train_1} > S^*$.

To show 2), the frontline worker's output is $\frac{1-e^{-\lambda z_m}}{p+H_m e^{-\lambda z_1}}$ and $\frac{\partial \frac{1-e^{-\lambda z_m}}{p+H_m e^{-\lambda z_1}}}{\partial z_1} = \frac{H_m \lambda e^{-\lambda z_1}(1-e^{-\lambda z_m})}{(p+H_m e^{-\lambda z_1})^2} > 0$, establishing that trained workers' output increases.

To show 3) is trivial, as untrained workers' output is constant.

To show 4), the production function is increasing in both of its arguments, and by parts 1) and 2), both arguments are increasing after training.

□

Proof of Proposition 2. Let the set-valued function $\eta(l)$ denote the sequence of requests a worker in layer l uses to seek help. We prove the claims out of order to build up results for trained and untrained workers before getting to the effects on managers. To show 3), untrained frontline worker output is $\frac{1-e^{-\lambda z_m}}{p+H_m e^{-\lambda z_1}}$ when $\rho = 0$ and the helping rule is $\eta(1) = \{m\}$. When $\rho = 1$, the helping rule is $\eta(1) = \{t, m\}$, untrained frontline worker output is $\frac{1-e^{-\lambda z_m}}{p+H_m e^{-\lambda z_1^t}+H_t e^{-\lambda z_1}}$. Untrained frontline workers become more productive under sequential help iff $\frac{H_m[e^{-\lambda z_1}-e^{-\lambda z_1^t}]}{e^{-\lambda z_1}} > H_t$. Because $0 < \frac{[e^{-\lambda z_1}-e^{-\lambda z_1^t}]}{e^{-\lambda z_1}} < 1$ when $z_1^t > z_1$, if there is a positive spillover to untrained workers, there must always be a sufficiently large time-advantage for accessing help from trained workers (the left-hand side) rather than going directly to a manager (the right-hand side). The right-hand side of the inequality is increasing in z_1^t, meaning that when skills for the trained worker increase, it is more likely that the inequality is fulfilled.

To show 2), a trained worker's time constraint is $1 = p + H_m e^{-\lambda z_1^t} + h_t \times \tilde{T}$ where h_t is the cost of help for a trained worker (which is distinct from h_m) and $\tilde{T}$ is the number of help requests a trained worker handles. Trained workers spend the remaining time after helping on drawing their own problems, or $\frac{(1-e^{-\lambda z_m})(1-h_t \times \tilde{T})}{p+H_m e^{-\lambda z_1^t}}$. For $\tilde{T}$ sufficiently large, this can be smaller than the worker's production before training. The trained worker's production falls when $\frac{(1-e^{-\lambda z_m})(1-h_t \times \tilde{H}_t)}{p+H_m e^{-\lambda z_1^t}} < \frac{(1-e^{-\lambda z_m})}{p+H_m e^{-\lambda z_1}}$. Total demand for help if trained workers handle all problems is $(n_1^* - 1)\frac{e^{-\lambda z_1}}{p+H_m e^{-\lambda z_1^t}+H_t e^{-\lambda z_1}}$. Total demand for help is a complicated function of the share of problems handled if $\rho < 1$, but we can derive a condition on the number of problems where trained workers help (determining ρ) such that the worker's productivity does not fall. To maintain their prior productivity, trained workers can handle at most $1 - (p + H_m e^{-\lambda z_1^t})/(p + H_m e^{-\lambda z_1})$ problems.

To show 1), note that help requests from an untrained worker to managers when the sequence is $\eta(1) = \{m\}$ are unchanged, yielding total help requests $\frac{e^{-\lambda z_1}}{p+H_m e^{-\lambda z_1}}$. When the help sequence is $\eta(1) = \{t, m\}$, managers receive $\frac{e^{-\lambda z_1^t}}{p+H_m e^{-\lambda z_1^t}+H_t e^{-\lambda z_1}}$ help requests. We have already established in the proof of Proposition 1 that help requests to managers from trained workers fall. Help requests from trained workers are weakly lower when they provide help. We can thus show that overall help requests fall by establishing that requests from untrained workers do not increase. For help requests to increase, we need $\frac{e^{-\lambda z_1^t}}{p+H_m e^{-\lambda z_1^t}+H_t e^{-\lambda z_1}} > \frac{e^{-\lambda z_1}}{p+H_m e^{-\lambda z_1}}$ where the left-hand side is the help-requests that reach managers under the sequential help rule and the right-hand side is help requests that reach managers when trained workers provide no help. The inequality can only hold when $p(e^{-\lambda z_1^t} - e^{-\lambda z_1}) > H_t e^{-2\lambda z_1})$. The left hand side is negative as $(e^{-\lambda z_1^t} - e^{-\lambda z_1}) < 0$ and the right hand side is positive, showing that help-requests to managers cannot increase.

To show 4), see the example in Figure 4.

$\square$

Proof of Proposition 3. The first order conditions for the organization's problem are:

$$\begin{aligned} \mu &: Y - (n\phi)^a (S)^b = 0 \\ z_1 &: n_1 \times w \times c - \mu \frac{\partial Q}{\partial z_1} = 0 \\ n_1 &: w(1 + c \times z_1) - \mu \frac{\partial Q}{\partial n_1} = 0 \\ z_m &: n_m \times w \times c - \mu \frac{\partial Q}{\partial z_m} = 0 \\ n_m &: w(1 + c \times z_m) - \mu \frac{\partial Q}{\partial n_m} = 0 \end{aligned} \tag{9}$$

The identification proof utilizes a standard rank condition for nonlinear GMM estimation, which is that the matrix of derivatives of the first order conditions has rank 3 after differentiating each with respect μ, a, and b. This follows from Newey and McFadden (1994). □

C.3 Extending the Model and Estimates to an Organization with More than 2 Layers

C.3.1 Robustness to the 2 Layer Assumption

Our results are not especially sensitive to a production function that assumes constant returns to scale (our results imply decreasing returns), which simplifies comparisons to models with more than 2 layers (see Table A12, Panel D). We have estimated models where workers in wage bands 1 and 2 draw problems from distinct distributions, with different λs, both types initially turn to managers for help, and we invert the first order conditions from a constant returns to scale production function to recover output elasticities. We reach similar conclusions as the 2 layer model, and the returns to vertical spillovers appear similar.

C.3.2 Estimating a Model with Different Knowledge Layers Among Frontline Workers

Our simple model with 2 layers generalizes to an organizational structure that can have different knowledge levels among frontline workers and where managers are assigned strategic tasks. We lay out the setup to a general problem and then focus on estimating a version of the model where frontline workers have 2 skill levels (corresponding to wage bands 1 and 2). Pinning down post-period spillovers to co-workers and different helping patterns across multiple layers is challenging because of thin cell sizes. Thus, we only estimate pre-period parameters to pin down the production aggregator from the first order conditions and we then plug in actual production changes (or assumptions).

In the general setup, the organization's problem is to choose: i) The number or layers, L, with 1 being the lowest layer and L being the top, where $L = M$ in our notation for the 2-layer setup. ii) A problem assignment rule $a : l \to k$ which assigns each layer problems from a distribution F_k (or a set of distributions) from which to draw problems.[38] iii) A set-valued helping rule for each layer $\eta(l) : l \to \{l+1, ..., L\}$ which determines the sequence in which problems should be passed through the organization if the original worker in layer l cannot solve it. All helping rules will be such that problems only go higher in the hierarchy (see the proof in Garicano (2000)), but it may be optimal to skip some layers if the organization can tailor the initial problem distribution to frontline workers with different knowledge. For example, it may be that $\eta(l) = \{l+2, l+3, ..., L\}$ or $\eta(l) = \{L\}$, where some layers may be skipped over when requesting help.

Point ii) extends the standard model that has all problems arise from the bottom layer. Workers in higher layers can be assigned their own problems, while part iii) allows for them to be skipped over in the helping sequence such that problems are passed above them.[39]

We let the production function be $Y = (n_1\phi_1)^{a_1} \times (n_2\phi_2)^{a_2} \times (n_m S)^{(1-a_1-a_2)}$. We then normalize by n_m and solve the system of first order conditions such that the production parameters a_1 and a_2 are the solutions to the system of equations:

$$\begin{aligned} \frac{1+cz_1}{1+cz_2} &= \frac{a_1/(n_1/n_m) + (1-a_1-a_2)(\tilde{h}_1/S)}{a_2/(n_2/n_m) + (1-a_1-a_2)(\tilde{h}_2/S)} \\ \frac{1+cz_1}{c} &= \frac{a_1/(n_1/n_m) - (1-a_1-a_2)(\tilde{h}_1/S)}{a_1 D_1 + a_2 D_2} \end{aligned} \tag{10}$$

where n_1 and n_2 are headcount of wage band 1 and wage band 2 workers, n_m is the number

[38] If the number of distributions $K > L$, then at least one layer will be assigned problems from multiple types of distributions.

[39] Carmona and Laohakunakorn (2024) analyze a model where workers can sort problems to solvers who know the solution. Our model, instead, features problems that arise from different distributions, but lower-level employees do not have the ability to discern whether those in higher-level are capable of solving problems. In a more general problem (especially with more than 3 layers), when production workers in different layers can get assigned tasks from different distributions, we must also check that it is optimal to blend teams together such that the same higher-level workers handle help requests from all layers of lower-level workers. Standard intuition would suggest that there should be specialization among a hierarchy of helpers. Organizations that face costs of creating separate partitions of workers may find that helpers should assist on problems from different distributions.

of managers, S is average manager strategic time, $\tilde{h}_l = h_m e^{-\lambda_l z_l} / (1 + H_m e^{-\lambda_l z_l})$ for $l \in \{1, 2\}$ is the expected per-capita helping time required of a manager for each worker in layer l, and $D_l = \lambda_l e^{-\lambda_l z_m} / (1 - e^{-\lambda_l z_m})$ for $l \in \{1, 2\}$. The parameters of the production aggregator are identified up to estimates of the inputs that come from fitting the micro-moments using GMM on the pre-period data.

D Survey

1. What was your wage band in 2019? (*choose only one option*):

   (a) 1____.

   (b) 2____.

   (c) Greater than____.

2. Did you participate in the training program run in the second half of 2018?:

   (a) Yes___.

   (b) No___.

   (c) DK/NA___.[40]

3. Remember your work environment in 2018 and 2019. Consider all of the people who you interacted with via email every week. How frequently did you interact with them face to face? (*choose only one option*):

   (a) More than once a week____.

   (b) Once a week____.

   (c) Once a month____.

   (d) Once a quarter____.

   (e) Once a half-year____.

   (f) Never____.

4. In your opinion, relative to 2018, the monitoring from your managers in 2019?

   (a) Was greater___.

   (b) Was smaller___.

   (c) It remained the same___.

[40] DK means: does not know while NA means that the question does not apply.

5. Remember your work environment in 2018 and 2019. What is the main reason why you emailed workers from a higher wage band *(choose only one option)*:

    (a) Asking for help to solve tasks and projects_________.

    (b) To report progress in tasks and projects_________.

    (c) Ask for authorization or approval of tasks and projects_________.

    (d) Social events_________.

    (e) If any other reason, which one_________.

6. Relative to 2018, in 2019 you:

    (a) Improved your understanding of how goals are set and how they are evaluated weekly? Yes___ No___ DK/NA___.

    (b) Were told explicitly that you should ask more for help to colleagues and peers and less to managers? Yes___ No___ DK/NA___.

    (c) Increased your probability of promotion inside the organization? Yes___ No___ DK/NA___.

    (d) Improved your ability to distinguish if tasks and projects require large or small divisional knowledge? Yes___ No___ DK/NA___.

    (e) Improved your ability to recognize if the tasks and projects require the knowledge from your division or different divisions? Yes___ No___ DK/NA___.

    (f) Increased the knowledge and the skills required to satisfactorily achieve goals? Yes___ No___ DK/NA___.

    (g) Received a larger number of across-divisions interdependent tasks. That is, a larger flow of tasks, projects or goals that require interaction with other divisions. Yes___ No___ DK/NA___.

    (h) Worked a larger number of hours a week? Yes___ No___ DK/NA___.

If you belong to wage band 2 or greater in 2019, please reply questions 7 and 8. Otherwise, please jump to question 9.

7. The main reason why you emailed workers from lower wage bands from your same division was (*choose only one option*):

    (a) Ask for help to solve tasks _________.

    (b) Give help to solve tasks _________.

    (c) Monitoring _________.

    (d) Delegating_________.

    (e) Social events _________.

    (f) If any other reason, which one is?_________.

8. What percentage of your working time in a week did you spend helping workers from wage band 1 from your same division in 2019? ____________%.

    (a) This percentage (*choose only one option*):

        i. Increased relative to 2018____.

        ii. Decreased relative to 2018____.

        iii. It remained the same relative to 2018____.

9. Recent research has found that wage band 2 workers increased their electronic communication with those of wage band 1 from their same division. In your opinion this is due to (*choose only one option*):

    (a) Workers from wage band 2 helped workers from wage band 1 on a larger number of tasks.

    (b) Workers from wage band 2 had to supervise workers from wage band 1.

    (c) Workers from wage band 1 asked more questions to workers from wage band 2.

    (d) Workers from wage band 1 helped workers from wage band 2 on tasks.

E 2024 Survey

For the following questions, we want to understand your perceptions of how quickly the knowledge you need to perform your job is changing. Please think only about the last 12 months of your current job.

1. How often do the tasks you perform at your job change? (*choose only one option*)

   (a) The nature of my tasks changes frequently: I am always doing something different.

   (b) My job is a mix of steady and new tasks.

   (c) My assignments are mostly stable.

2. How often do you need to learn new things to perform your tasks? (*Please choose as many options as needed*)

   (a) I often learn new things because of the IT-related changes.

   (b) I often learn new things because of the NORMATIVE changes.

   (c) I often learn new things because of the changes in MANAGEMENT STYLES.

   (d) Sometimes I learn new things because of the IT-related changes.

   (e) Sometimes I learn new things because of NORMATIVE changes.

   (f) Sometimes I learn new things because of the changes in MANAGEMENT STYLES.

   (g) Generally, I don't need to learn much to perform my daily tasks, considering that the function manual is constant.

   In the following questions, we want to understand with whom you communicate to seek help, authorizations, or to interact within the organization.

3. Occasionally, staff members ask questions of another(s) when they need help with work. What proportion of requests for assistance do you direct to the following persons? (Please make sure your answers add up to 100)

   (a) Someone in the same grade or a similar position. [0-100]

   (b) Someone of some grade lower than yours. [0-100]

(c) Someone at a higher grade level but not your supervisor, boss, or coordinator. [0-100]

(d) Your supervisor, boss, or coordinator. [0-100]

(e) External Agencies. [0-100]

4. When you need authorizations to perform your work, what proportion of requests for authorization do you address to the following persons? (Please make sure your answers add up to 100)

   (a) Someone in the same grade or a similar position. [0-100]

   (b) Someone at a higher grade level but not your supervisor, boss, or coordinator. [0-100]

   (c) Your supervisor, boss, or coordinator. [0-100]

5. Consider the emails you exchange with others. What proportion of your outgoing emails would you place in the following categories? (Please make sure your answers add up to 100)

   (a) Requests for help sent to others [0-100].

   (b) Respond to someone else's request for help [0-100].

   (c) Request endorsements or authorizations [0-100].

   (d) Provide endorsements or authorizations [0-100].

   (e) Social relationship with an official [0-100].

   (f) Others [0-100].

6. Asking questions is a normal part of the workplace. You can ask them to your bosses or immediate superiors or to other people (e.g., colleagues in the same grade or higher but who are not your bosses). What proportion of your work-related questions do you ask your boss or immediate superior using the following methods? (Please make sure your answers add up to 100)

   (a) emails. [0-100]

(b) Personally (in private or in meetings). [0-100]

(c) Telephone calls [0-100]

(d) WhatsApp, Microsoft Teams or Zoom [0-100]

(e) Others [0-100]

7. What portion of your work-related questions do you ask people who are **NOT** your immediate bosses or superiors using the following methods? (Please make sure your answers add up to 100)

   (a) emails. [0-100]

   (b) Personally (in private or in meetings). [0-100]

   (c) Telephone calls [0-100]

   (d) WhatsApp, Microsoft Teams or Zoom [0-100]

   (e) Others [0-100]

8. For each of the following questions, please answer from 0 to 100.

   (a) What proportion of the **QUESTIONS** you ask your boss or immediate superiors are exclusively work-related? [0-100]

   (b) What proportion of the **EMAILS** you send to your boss or immediate superiors are exclusively work-related? [0-100]

   (c) What proportion of the **QUESTIONS** you ask people who are not your bosses or immediate superiors are exclusively work-related? [0-100]

   (d) What proportion of the **EMAILS** you send to people who are not your bosses or immediate superiors are exclusively work-related? [0-100]

9. Expectations of your work in this organization compared to other jobs. (*choose only one option*)

   (a) I am likely to remain with the organization for the long term and my salary and benefits compare favorably with the market.

(b) I am likely to remain with the organization for the long term despite concerns about my salary and benefits relative to the market.

(c) I am unsure of my long-term prospects with the organization even though my salary and benefits are favorable relative to the market.

(d) I am unsure of my long-term prospects with the organization and, in addition, my salary and benefits are not favorable relative to the market.

References: Appendix

Belloni, A., Chernozhukov, V., Hansen, C., and Kozbur, D. (2016). Inference in high-dimensional panel models with an application to gun control. *Journal of Business & Economic Statistics*, 34(4):590–605.

Caliendo, L. and Rossi-Hansberg, E. (2012). The impact of trade on organization and productivity. *The Quarterly Journal of Economics*, 127(3):1393–1467.

Carmona, G. and Laohakunakorn, K. (2024). Improving the organization of knowledge in production by screening problems. *Journal of Political Economy*, 132(4):1290–1326.

Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. *Journal of Political Economy*, 108(5):874–904.

Garicano, L. and Rossi-Hansberg, E. (2004). Inequality and the organization of knowledge. *American Economic Review*, 94(2):197–202.

Gumpert, A., Steimer, H., and Antoni, M. (2022). Firm organization with multiple establishments. *The Quarterly Journal of Economics*, 137(2):1091–1138.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245.