Training, Communications Patterns, and Spillovers Inside Organizations*

Miguel Espinosa† and Christopher Stanton‡

March 19, 2023

Abstract

Most organizations utilize hierarchies to facilitate specialization, with managers assisting workers on tasks beyond their capabilities. As workers gain skills, they require less help, freeing up manager time. In this paper, we estimate direct productivity treatment effects for individual workers and spillovers to managers after a randomly assigned training program for frontline workers in a Colombian government agency. Trained workers improved their individual production, while help requests to managers declined, enabling managers to focus on higher-value work. Accounting for vertical spillovers to managers meaningfully changes the organization's implied return on investment from training.

*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, Michela Giorcelli, Mitchell Hoffman, Jonas Hjort, Hideo Owan, Andrea Prat, Canice Prendergast, Eric Quintane, Imran Rasul, Ray Reagans, Raffaella Sadun, Antoinette Schoar, 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 and Juan Herrera 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

Hierarchies that partition employees by skill or task are ubiquitous features of organizations, as they allow for worker specialization. This is achieved through a theoretical mechanism known as management by exception, where frontline workers handle common tasks, while higher-level managers solve relatively rare problems (Garicano, 2000; Garicano and Rossi-Hansberg, 2004; Caliendo and Rossi-Hansberg, 2012). One of the core features of organizational hierarchy models is that as frontline workers become more skilled, they can autonomously address a broader range of problems, leading to fewer problems qualifying as exceptional. This, in turn, reduces the need to ask managers for help, freeing up their time to focus on higher-level or strategic tasks that may be more valuable than solving individual workers' production problems.

The extent to which managers must help subordinates as a function of their skill is a crucial parameter for determining optimal organizational structures, yet direct evidence on this channel has been elusive. Most tests of hierarchical models are indirect, studying how layers in an organization vary with shocks to opportunities, differences in market size, or proxies for changes in communication costs or information (Garicano and Hubbard, 2007, 2009; Bloom et al., 2014; Caliendo et al., 2015, 2020; Friedrich, 2022; Gumpert et al., 2022).[1] Yet the direct test of the link between help requirements, worker skill, and manager productivity is important, as there are distinct implications when the cost of acquiring knowledge (training) changes and managers are required to provide less help. In particular, reductions in training costs can lead to de-layering, broader spans of managerial control, and widening wage inequality if knowledge acquisition becomes cheaper across the skill distribution (Garicano and Rossi-Hansberg, 2006). These implications are distinct from those arising from changes in scale or communications costs.[2]

This insight also has implications for the measurement of the returns to training

[1]In fact, later empirical applications model organizational structure (and resulting inequality inside firms) as a function of the scale of operations and/or communications frictions across establishments (Friedrich, 2022; Gumpert et al., 2022).

[2]Changes in the cost of acquiring commodity skills for lower level workers, say through online education or MOOCs, without changes in the cost of acquiring frontier skills may offset some of the increase in inequality.

programs. Training program evaluations where the unit of analysis is an individual worker will omit any benefits to managers, understating aggregate program benefits.[3] Failure to account for spillovers across a hierarchy can have potential welfare consequences. Under a variety of labor market frictions, efficiency likely requires employers to invest in training because workers are not the full residual claimants on their skills investments (Acemoglu and Pischke, 1998, 1999; Lazear, 2009; Cavounidis and Lang, 2020). Yet one prominent view is that firms under-provide training (Cappelli, 2012), and a potential reason is that the full return is difficult to quantify.[4] Thus, accounting for the spillovers from training may change an organization's perceived return on training investment, but the magnitude of these spillovers and whether they are consequential is an empirical question.

In this paper, we directly study how manager productivity changes as a result of frontline worker training, which we term vertical training spillovers. The setting for our study is a Colombian federal government investigative agency. In late 2018, 12% of the organization's frontline workers were randomly allocated to participate in a 120 hour training program covering computer skills, goal setting and management, legal analysis, written communication, and other topics related to each participant's work. Besides the random assignment to training, several features of the setting enable us to examine vertical training spillovers.

First, we have individual-level, time-varying productivity data on goal achievement for both workers and managers. Each employee in the organization, including managers, has goals set and evaluated every week by an independent, separate unit that is responsible for oversight and performance evaluation. Frontline workers' goals typically entail case processing or functional execution, while managers' goals involve establishing processes or formulating strategic direction. Similar uses of goals and objectives, with outside

[3] Analyses that proxy for productivity changes using workers' wage changes may be biased downward even for understanding worker output changes if the incidence of workers' skill upgrading partially falls on managers or other layers of a hierarchy. Due to data limitations, the early literature evaluating on-the-job training largely inferred efficacy based on changes in individual wage gains and qualitative measures of job performance (Bartel, 1995).

[4] According to Training Industry Magazine, when organizations do attempt to quantify training performance, they tend to focus on individual-level outcomes. See: https://trainingindustry.com/articles/measurement-and-analytics/how-to-identify-the-right-training-kpis-for-your-learning-and-development-programs-spon-eidesign/ .

measurement against them, is common in the public sector (Rasul and Rogger, 2018; Rasul et al., 2018). Our study organization has adopted this evaluation structure because its main function entails sensitive public-interest work, and the separation of oversight from on-the-job tasks is thought to provide accountability. The organization gave us access to 13 weeks of goal achievement data from early 2018, before the training period, and the same 13 weeks of data from early 2019, a few months after the end of training. This allows us to estimate the direct effect of training on productivity for those randomized into the program relative to controls who did not receive training.

Second, we have metadata on the quantity of email communications between all workers (frontline workers and managers) from the periods pre- and post-dating the training program. Although the organization does not have formal teams, workers and managers have stable relationships. As a result, some untrained workers and managers were relatively more exposed to workers who received training than others. Our direct evidence on vertical training spillovers entails an assessment of whether managers who were relatively more connected or exposed to trained workers had productivity that evolved differentially compared to managers who were more isolated from trained workers.

Turning to our main empirical results, we first document that the training program raised productivity for trainees, which is not obvious given the literature on training programs (Card et al., 2018). Average goal achievement among trained frontline workers increased from 71.9% per week to 78.5% per week between the pre- and post-periods. Trained workers' increase in goal achievement was positive across the pre-period productivity distribution, while effect sizes were slightly larger for lower performers. Untrained frontline workers' goal achievement remained at 72%, and their average goal achievement change was approximately zero across the pre-period productivity distribution. In this setting, everyone who was randomized into the program participated and attended at least 85% of the training sessions (which was the organization's threshold for determining successful completion). Thus, a simple comparison between treated and control workers suggests the average treatment effect of the 120 hour program is a 10% goal achievement increase in the medium-run (4-6 months after training completion).

Second, we show that vertical training spillovers to managers are economically significant. They are large enough to potentially alter our study organizations' assessment of whether the return on investment from the training program is positive. In the raw data, managers' average goal achievement increased from 70.8% to 73% between the pre- and post-periods. Our preferred measure of managers' exposure to trained workers is the log number of pre-period emails received from workers who later get randomized into the training program. If emails from workers proxy for help requests, this measure will capture managers' who have the greatest potential to see a reduction in help requests after workers receive training. Difference-in-differences regressions of managers' goal achievement on this exposure measure indicate that spillovers from training are responsible for an approximately 2 percentage point increase in manager goal achievement.

Our estimates of vertical spillovers to managers are robust to a variety of measurement strategies and controls, including using the share of emails from eventually trained workers as the exposure measure, controls for reallocation of communication networks in the firm, and mean reversion in productivity. A LASSO procedure selects managers' exposure to incoming emails from trained workers (in logs or as shares) as the variables that best explain changes in their goal achievement over the panel of data we observe.

We have also probed the general robustness of our estimates of direct returns to workers, as comparing the magnitude of the direct returns to the vertical spillovers to managers is important for our evaluation of the program. A similar strategy to the vertical spillover estimation allows us to assess and control for spillovers to coworkers, accounting for violations of the Stable Unit Treatment Value Assumption (SUTVA). These estimates yield qualitatively similar conclusions to the simple difference-in-differences estimator.[5] In addition, although the organization indicated that goal setting and attainment measurement did not depend on workers' training status (due in part to the goals and measurement being set by an outside party that was not aware of workers' treatment status), our inference about the importance of vertical spillovers would likely result in a null finding if frontline workers' actual productivity did not change but reported productivity increased

[5]Measures of potential spillovers to untrained workers do not survive LASSO procedures to select variables that matter for the evolution of productivity.

due to manipulation.

To evaluate the training program, we use the insights from a modified model of organizational hierarchy to compare the magnitude of the direct returns and the spillover to managers. Because managers receive higher wages, freeing up their time from helping to focusing on broad organizational priorities may have a greater impact than improvements in frontline workers' processing time. In this setting, managers' assigned goals and evaluations are based on their own work, which is assigned directly to them.[6] We illustrate how to compute the impact of training on the organization by calculating the labor cost savings to achieve a fixed number of frontline workers' goals and the labor cost savings (for higher earning managers) to achieve a fixed number of manager goals. The crucial parameter is the vertical training spillover that arises from frontline workers' increasing skills.

Based on the wage bill for managers versus workers, the 2 percentage point increase in managers' goal achievement due to vertical training spillovers is equivalent to changes in labor costs equal to the wage bill of approximately 8 frontline workers. The improvement in frontline workers' goal achievement is equivalent to the wage bill of 5.6 frontline workers, meaning spillovers to managers are substantially larger in light of their higher wage bill. These topline estimates of frontline workers' cost savings from training do not include the direct costs of the program and the opportunity cost of workers' time in the classroom. When these factors are included, the direct return calculation is highly dependent on the time horizon over which the organization expects to be able to capture productivity gains. However, including vertical spillovers to managers can change the sign on the return on investment from the program from negative to positive if the time horizon for the organization to capture benefits is under 1 year.[7]

To provide context for the source of vertical spillovers to managers, we evaluate several

[6] Managers' goals, like strategic planning and process improvement, are not dependent on frontline worker output, but managers are still tasked with helping workers when questions arise.

[7] In our context, the organization captures all gains for workers who remain, as trained workers had no increase in salary relative to untrained workers, likely because of rigid rules about compensation setting in this sector. Outside of changes in government, jobs in this organization are very stable, and turnover is minimal.

possible explanations for why managers may benefit from trained workers. Specifically, we compare the fit of models where workers and managers are complements in a production process to a Garicano (2000) inspired hierarchies mechanism, where manager help substitutes for workers' skills in production tasks. We find support for the hierarchies model based on two pieces of evidence. First, when managers handle exceptional problems, worker-manager communication is predicted to decline as workers become more productive. We find that trained workers simultaneously increase their productivity and communicate less frequently with managers; at the same time, managers connected to trained workers become more productive. If workers and managers were instead complements, an email reduction might instead signal a decline in collaboration with a productive colleague. Second, workers' survey responses support the hierarchies mechanism. Their responses indicate that emails from workers to managers are sent for the purpose of seeking out help. These surveys also suggest that emails are positively correlated with non-electronic communications, suggesting that email evidence is useful as a proxy for the totality of communications between employees. Given that email communications between trained workers and managers decline, we interpret the causal force behind our results as a reduction in managers' needing to provide help on exceptional problems.

When prior work has attempted to estimate human capital spillovers, the primary focus has been on peers at the same level. Prominent examples are Adhvaryu et al. (2018) and De Grip and Sauermann (2012). Adhvaryu et al. (2018) randomize whether any garment production workers are eligible for soft skills training at the line level and then randomize a subset of workers into the program within each eligible line. Using this saturation design, they examine spillovers to same-level coworkers. They find larger spillovers on teams where managers have more autonomy and smaller spillovers when managers are more attentive. Using an experimental design that varies the timing of training within teams, De Grip and Sauermann (2012) estimate that a 10 percentage point increase in the share of trained peers increases performance by 0.5% among call center workers.[8] There are two key differences between these analyses and ours. First, we

[8]Other relevant papers are Levitt et al. (2013), who examine learning by doing and how it cascades across workers, and Sandvik et al. (2020), who run an experiment showing the power of knowledge spillovers

examine spillovers across levels of the firm hierarchy, and second, our approach to detect spillovers is based on communication patterns. We are aware of no other papers that estimate the spillovers from training between workers and managers across the vertical hierarchy of an organization.

Instead, the work that considers vertical or multi-layer organizations examines the impact of managers on their subordinates (Lazear et al., 2015; Hoffman and Tadelis, 2021), or how managers' performance pay changes effort and the importance of social connections with workers (Bandiera et al., 2007, 2009). Relative to this literature, our focus is on bottom-up spillovers rather than top-down impacts of bosses. Our results suggest that the individual returns to training may fail to account for a significant fraction of the surplus generated because having more productive workers allows managers to focus on higher value work. While we caveat that both the direct returns and spillovers may be more ephemeral in other settings where the ability to capture the value from training programs may differ, we believe these results are relevant for a large class of public sector entities and firms with some market power or differentiated organizational structures. Like the organization we examine, many public sector organizations feature relatively low turnover and limited head-to-head competition among workers, suggesting that spillovers may be substantial and that the gains from training may significantly improve organizational performance and the quality of government (Besley and Persson, 2010; Dal Bó et al., 2013; Rasul and Rogger, 2018; Rasul et al., 2018; Bandiera et al., 2021).

2 A Motivating Framework

We begin with a motivating framework to present the simplest comparative statics from organizational models of hierarchy, *à la* Garicano (2000). The core purpose of the framework is to derive a cost function for the organization that enables a comparison of the

by increasing contacts between coworkers. Other work, like Kugler et al. (2022), estimates spillovers from training to relatives, which may provide another wedge between the social and private returns to training.

value of the direct training returns to the vertical spillovers to managers across the hierarchy. We have loosely modified the setup and assumptions to match some key features of our setting.

We focus on the problem of a single manager's span of control over n workers and their choice of skill level – in the same spirit as Garicano and Hubbard (2007) – who fix the number of layers in a hierarchy to focus on a manager's span. This simple setup abstracts from multi-layer hierarchies and the choice of the number of layers. This closely matches our setting, while more general formulations are straightforward applications of Caliendo et al. (2015). The problem for the organization is to consider how the n workers and the manager allocate their 1 unit of time endowment between individual problem solving or seeking/providing help on tasks. There are two types of tasks. The first, which we call production, entails executing against cases or deliverables, represented as a goal to be achieved by a frontline worker. Production tasks vary in their difficulty. For simple illustrative purposes, we assume the distribution of production task difficulty is uniform on the unit interval. Frontline worker skills determine which production tasks can be done in the absence of help. A frontline worker with skill level z_w can solve all tasks less difficult than z_w. The probability that a production task can be accomplished by the frontline worker alone is thus z_w, which is a knowledge threshold for producing a solution. Workers are required to work on problems sequentially – they must solve a problem or seek help before moving on to the next one. This is akin to saying that workers cannot ignore difficult goals, they will still be evaluated against them.

We abstract from managers' skill investment in this setup, as manager skills were not affected by the program that motivates our inquiry. For this simple setup, we assume that managers know how to solve workers' problems (their skill level is 1), but there may be congestion in accessing manager time. In the case that a frontline worker needs the manager's help on a task, the worker reaches out for help, meets with the manager and explains the issue. The manager then provides the worker with a solution. The time it takes to ask a manager for help (and the time it takes the manager to provide help) is denoted h. That is, tasks in the range $z \in (z_w, 1]$ require an additional h units of time from both workers and managers to come to a solution.

The second task type entails strategy rather than production. These tasks involve planning and direction setting, which only the manager can do. While most models have all tasks that originate with lower-level employees, having certain task types that only managers do matches our setup.[9]

Workers have a unit of time to produce with time constraint:

$$1 = p + h(1 - z_w),$$

where p is the baseline time required to solve a production task and h is the additional time spent getting help from a manager on tasks that are more difficult than z_w. The manager's time constraint is

$$1 = s + \frac{nh(1 - z_w)}{p + h(1 - z_w)}, \tag{1}$$

where s is the time that the manager spends on strategy tasks and $\frac{nh(1-z_w)}{p+h(1-z_w)}$ is the time spent helping the n workers on production tasks. The denominator captures the expected number of tasks a worker with skill z_w undertakes.[10] The numerator is the probability a task requires help multiplied by the time cost of help and the number of workers.

Equation (1) has an important implication: decreases in the manager's helping time commitment as workers gain skills free up time for strategic activities, as $\partial \frac{nh(1-z_w)}{p+h(1-z_w)} / \partial z_w = \frac{-nhp}{(p+h(1-z_w))^2} < 0$. It is this core tradeoff between helping workers and spending time on strategic tasks that motivates our empirical exercise. When workers become more skilled, managers can devote time to strategic undertakings because workers require less help.

The organization's problem is to choose the number of workers per manager and their investment in skills. Managers are paid $\tilde{W}_m$, which is taken as given, while workers are paid $\tilde{W}_w$, which may vary with their skill. The cost of upskilling is c, which is proportional to

[9] This setup has attractive features for valuing spillovers to managers versus direct productivity changes because it allows us to compare two dimensions of costs. We can accommodate the same analysis with a single type of problem that gets passed from lower to higher layers in the organization if we parameterize the value of more difficult problems to the organization.

[10] To keep the problem simple, we abstract from variation due to the stochastic draws of difficulty. Depending on the production function, accounting for this variation might lead an organization to build in more slack to accommodate shifts in the demand for manager assistance.

the increment of workers' skill acquisition. The organization's task is to balance frontline worker output with an individual manager's time for strategy. The cost function is the solution to

$$\begin{aligned}
&\min_{n,z_w} n(\tilde{W}_w + cz_w) + \tilde{W}_m \\
&s.t. \\
&\frac{n}{p + h(1 - z_w)} \geq \overline{Q} \\
&1 - \frac{1}{p + h(1 - z_w)} nh(1 - z_w) \geq \overline{S}.
\end{aligned}$$

The first constraint says that the organization needs output from all frontline workers to total at least $\overline{Q}$ problems solved. The denominator is the expected amount of time per problem after a worker is trained and has skills z_w. This expected time includes the probability that a worker will require help, which addsadditional h units of time for problems with difficulty above z_w.[11] The second constraint says the organization requires $\overline{S}$ units of manager time devoted to strategy, where there is some linear, increasing function mapping time inputs to outputs that is innocuous because the strategy task is performed autonomously.

Suppose we begin with some initial level of frontline workers n^0 and training z_w^0. At an optimum, we know the first constraint binds because producing more than $\overline{Q}$ requires hiring more workers at wage $\tilde{W}_w$ or increasing their skill level, which costs c. Conditional on $\overline{Q}$, the second constraint binds because the only way to increase output above $\overline{S}$ is to increase z_w, which costs c. At an optimum, frontline worker output equals $\overline{Q}$, and total strategic output per manager is $\overline{S}$. Now consider a change in z_w from z_w^0 to z_w^1, which may result from a training program. Holding fixed n, manager output increases to S^1,

[11] As we discuss later, although some of the training happened during work time in our setting, we derive the constraint as one on long-run output, while the opportunity cost of training is a short-term reduction in worker availability. We will adjust for the opportunity cost of production time in our later calculations, but we have not done so here to simplify the exposition.

and the total expenditure required to obtain $\overline{S}$ falls by

$$\tilde{W}_m(1 - \frac{\overline{S}}{S^1}). \tag{2}$$

Similarly, total expenditures to achieve $\overline{Q}$ units of frontline output (among frontline workers) falls by

$$n\tilde{W}_w(1 - \frac{\overline{Q}}{Q^1}) \tag{3}$$

where Q^1 is the new output for frontline workers after increasing skills to z_w^1. In these derivations, we assume that wages are fixed/invariant to productivity (which fits our setting in the public sector), but these assumptions can easily be relaxed for other contexts.

Our approach to compare the relative importance of the direct benefit of training to frontline workers with the spillover benefit is based on a comparison of the respective changes in costs needed to achieve the pre-training level of output observed by frontline workers and managers, respectively. When frontline workers become more skilled, they can handle more tasks. Their processing time improves by avoiding the need to wait for managerial help. Managers can change their focus to strategic tasks or high-level concerns.

3 The Setting

The context for our study is a public sector organization in Colombia, a middle-income country with a growing economy when our data was collected. Our agreement with the organization prevents us from identifying it, but we can say that it is one of several control, oversight, inspection, surveillance or investigative institutions among the country's federal government. We have obtained anonymous email, productivity, and personnel records for each of the 655 employees from the core area of the organization.

The core area of the organization has 5 divisions, and workers perform slightly different functions depending on their division. Each division has the following responsibilities: 1. The "Execution Division" (36.9% of employees) is code named to preserve anonymity for the organization. This division answers citizen requests, conducts investigations, and issues findings that can be used in disciplinary proceedings. 2. The Administration Division (19.3%) controls acquisitions, inventory, storage, and the supply of goods and services required by the entity. 3. The Finance Division (13.7%) manages the budget and treasury. 4. The Human Talent Division (14.9%) handles the creation and implementation of recruiting policies, onboarding, payroll, and other HR tasks. 5. The Planning Division (14.9%) advises top management on the creation of policies, procedures, and resource allocation.

At entry into the organization, workers are assigned to a division and a wage band according to their education and experience in the government sector. There are 5 wage bands, with 5 being the highest. Wage band 1 and 2 employees consist of high school graduates and those with bachelors degrees and are “frontline” workers in our terminology. Workers from wage band 3 to 5 hold bachelors, masters or PhD degrees and are “managers.” Salaries make up the totality of compensation and are determined by wage band and experience in the public sector. The employees that we study are in stable, white collar occupations where turnover is limited.[12]

The organization also has other employees in non-core areas, the most important of which is an oversight group that is separate and independent of the core and serves to check and monitor employee performance. Employees outside of the core have limited direct interactions with the core employees and were not eligible for the training program.

The organization measures weekly individual performance. Weekly goals for each worker are set by the separate oversight group charged with performance evaluation. Because of the independent nature of the performance monitor, the organization’s leadership has confirmed that goal setting or performance evaluation did not (at least formally) take into account workers’ training status. There was no ratcheting of expectations either in response to past performance or training attainment. Organization leaders reported to us that there existed limited scope for the oversight group to increase trained workers’

[12]During the period of our data, the organization had minimal hiring and negligible turnover. In fact, we only observe two workers who leave the organization during our 2 years of data, one untrained frontline worker and one manager. Although unusual for other contexts, lack of mobility outside of elections and periods of government turnover is common in Colombian government agencies.

performance evaluations due to manipulation or demand effects.

Because of this oversight function, we clarify the particular managerial roles that high wage band workers play. Managers in our setting do not monitor or conduct performance evaluations. Managers' main role is planning and setting strategic priorities. When needed, they also help frontline workers, but managers' performance evaluations are based on their own strategic work – rather than on the performance of frontline workers whom receive help. Surveys and interviews with the organization indicate that managers main role is to provide help to frontline workers, but managers may also provide authorizations for a lower level worker to take a new course of action if higher level input into a decision is required.

We were given data that covers the same 13-week window from April to June in two adjacent years, 2018 and 2019. As we discuss in more detail below, the organization randomized frontline workers into a training program in the Fall of 2018, and our data spans the pre- and post-periods. The data contain individual weekly goal achievement for both workers and managers (our productivity measures), absenteeism, and demographic and personnel information, including gender, education, monthly wage, wage band, and division.

We supplement these administrative records with metadata on email communications between all 655 employees. We have data on daily bilateral email counts between every pair of employees over the 13 weeks in 2018 and the same 13 weeks in 2019.[13] We expect that the largest share of email communication is related to work matters, but we do not have the subject or the text of any emails. As such, we rely on results of surveys (provided in section 6.2) that confirm that emails proxy for the totality of communications between individuals.[14]

[13] The data contain the quantity of emails at the daily level, not the thread or message level, so we cannot observe whether sent emails contain multiple recipients.

[14] Emails are a good proxy for total communications if electronic and other communications are complementary (i.e., you are more likely to email people who you also talk with face-to-face), rather than substitutes. We surveyed workers in the organization, and their responses indicate that emails and other forms of communication are complements.

3.1 Goal Achievement and Evaluation

Every worker, including managers, has goals set and evaluated weekly. We do not observe the content of the individual goals, but interviews gave us examples of the goal setting process and the qualitative nature of goals that are specified each week. For example, a weekly goal for frontline workers in the Execution Division would typically entail progress on one or multiple cases or investigations. Weekly goals for managers in this division would typically include filing reports on case audits, planning for future investigations, and establishing contingencies if case execution is not going according to plan.[15]

Goals evaluation has 4 components, but we only observe the aggregate score out of 100. The components are: a target completion factor that is quality weighted (35%), a resource use efficiency factor (35%), an orientation factor that assesses whether the work output is in line with organizational objectives or guidelines (15%), and a processes factor (15%) that assesses whether appropriate procedures were used. Our main measure of productivity is the overall goal achievement score in each week. We only have the aggregate, numeric measure of goal achievement; we do not observe the individual components or the specific goals that each employee has every week.

3.2 Training Program

At the end of July 2018, the organization decided to run a training program that would last for a 16-week period from August to December of 2018. Although the original aim was to train the entire workforce, budget considerations meant only 63 employees could

[15]Examples for other divisions are similar. Workers in Administration handle procurement, inventory management, and policies and procedures. Workers' goals in the Administration division will typically involve satisfactory procurement execution or implementing compliance procedures for the organization. Managers in this division are involved in devising procedures and in strategic planning around inventory, property, and equipment. Workers' goals in the Human Talent division typically involve execution of HR functions, including acquisition of data for reporting processes. Managers will typically be measured against initiatives and analyses affecting the organization's human capital planning. Workers' tasks in the Finance division tend to focus on conducting transactions and adhering to budgets. Managers are responsible for budgeting and monitoring payments and cash inflows in the accounting system while ensuring that the legal requirements related to payments are fulfilled. Workers' goals in the Planning division tend to focus on strategy execution–gathering information and using it for planning purposes, whereas managers broadly oversee setting the direction for how plans will be produced and communicated.

participate. These employees were chosen randomly from frontline workers (wage band 1 and 2), without stratification. A lottery was conducted to determine eligibility. All employees were informed of this selection method and were aware that no other sponsored training programs of this type were planned for the future. Section 3.3.1 shows that randomization into training is balanced on observable characteristics.

Selected participants attended classes three days per month. Each day of training had 8 hours of classes, with a total training time of 120 hours. The program covered five different thematic areas. Four areas focused on the acquisition of general-skills and one focused on division-specific skills. The general-skills topics included: (i) Principles of goal setting, scheduling, and time management, (ii) Computer Skills, with a focus on Microsoft Excel, (iii) Legal Analysis, specifically on the Colombian constitution, and (iv) Principles of good written communication.

The final module contained specific topics related to the employee's division. Employees in the Finance division studied principles of banking, accounting, and public finance. Those in the Execution division studied national and international law. Administration division workers learned principles of operations research analysis. Human Talent division workers studied how to motivate workers and keep them satisfied in the workplace, while Planning division employees took a mini-course on impact evaluation and policy decision making.

3.3 Descriptive Statistics

3.3.1 Data on Workers and Treatment Balance

Table 1 shows descriptive statistics for the sample, split by frontline workers and managers in Columns 1-3. The organization effectively has two layers. The lower layer contains frontline workers in the first two wage bands, with wage band 2 workers having relatively higher levels of education or experience than those in wage band 1. The upper layer contains managers in wage bands 3 and onwards. There are 526 frontline workers and 129 managers. Columns 4-5 focus only on frontline workers and split the sample based on training status, while Column 6 presents a test of differences in characteristics of the null that treatment assignment is balanced.

Females make up 48% of wage band 1 workers, 29% of wage band 2 workers and just 18% of managers. Trained workers are also more likely to be female than untrained frontline workers, which is one of the only sources of imbalance that we find in treatment assignment. Some imbalance in individual covariates is expected in multivariate tests of balance in treatment assignment. The F-statistic at the bottom of the table presents the joint test of balance in treatment assignment based on worker characteristics. The joint test rejects systematic imbalance in treatment assignment to training. We will later control for characteristics, like whether the worker is female, and show that our results are insensitive to these controls for potential imbalance in treatment assignment.

Moving down rows, we see that managers are more educated, with 64% holding a Bachelor's degree and 36% a Masters or PhD, while over half of the frontline workers have only a secondary (high school) education. The next few rows of Table 1 show the allocation of workers and managers across divisions. Forty-five percent of wage band 1 workers and 24 percent of wage band 2 workers are in the Execution Division, compared to 31 percent of managers. The ratio of managers to frontline workers is greatest in the Execution division, suggesting that small individual changes in worker productivity may add up to have a larger effect on managers in this division because of their greater spans of control.

The next few rows deal with wage bands and wages. The row labeled Wage Band is mechanical in Columns 1 and 2, but is relevant as a randomization check for frontline workers into training (Columns 4 and 5). All rows reporting wages are normalized relative to the pre-period average for Wage Band 1 workers. In the pre-period, the average wage band 2 worker earned 20% more than the wage band 1 mean, while the average manager earned 2.16 times more than wage band 1 worker. Comparing pre-period and post-period wages, there is an increase for all employees, including managers. Baseline wage increases are larger for higher wage bands year-over-year.

Of particular relevance is whether trained workers capture returns from training via higher wages. There are no abnormal wage increases for trained workers in Column 5 compared to Column 4.[16] Relying on wages to capture the effects of training would have yielded null results in our setting. On the other hand, managers do have greater wage increases than frontline workers. However, manager wage increases are orthogonal to their individual year-over-year changes in goal achievement. That suggests that manager pay changes should not be considered a cost of the training program.

The final rows of Table 1 show the most important results. Comparing Columns 4 and 5, we see significant changes in average goal achievement for trained workers, with goal achievement increasing from 71.9% to 78.5%. The increase in goal achievement is about 6.6 percentage points for both Wage Band 1 and Wage Band 2 trained workers. Goal achievement for untrained workers was essentially flat, averaging 72.6% in the pre-period and 72.1% in the post-period in 2019. While goal achievement was flat for untrained frontline workers, Column 3 shows that goal achievement for managers increased by 2.2 percentage points, from 70.8% to 73%. This change motivates our estimates of spillovers from training across the organization's vertical hierarchy.

3.3.2 Email Data

We have metadata on emails between pairs of workers for the same time period as our data on goal achievement. We base our analysis on quantities of emails between senders and individual receivers, as we cannot distinguish whether email threads are to teams or multiple recipients.

We use these email data for three purposes: a) to capture connections between coworkers, allowing us to account for same-level (horizontal) spillovers from training; b) to capture connections between workers and managers, allowing us to estimate vertical spillovers across from worker training to manager productivity; and c) to capture how communications patterns change after training, allowing us to provide evidence on the mechanisms behind the spillovers we estimate.

Our strategy for identifying vertical and horizontal spillovers is based on the idea that some managers or untrained workers will be more connected to workers who will eventually

[16] It is possible that wage increases lag beyond the end of our post-period data.

become trained compared to others. Our theory predicts that managers who spend more time helping workers– which we capture with the level of emails with eventually trained workers – should have larger spillover effects because at baseline they have less time to devote to their own strategic tasks. Table 2 provides detail about the email connections measures and correlations between alternate measures of exposure to trained workers. For example, untrained workers receive an average of 674 emails from eventually trained workers from their own division in the pre-period, with a standard deviation of 385. Managers average 1670 pre-period emails from eventually trained workers in their own division, with a substantial standard deviation relative to the mean of 893 emails.[17]

Some managers may be more exposed to trained workers because they are more central in communications with all frontline workers, not just those who are randomized into training. This is not a problem if the level of email communications is stable, but may overstate a managers' exposure to training if a manager's engagement with frontline workers is based on transitory spikes in activity. To account for this, we also compute exposure based on shares of emails. For managers, the average share of emails with eventually trained workers is 12.1%, with a standard deviation of 2%. For untrained workers, the average share of emails with trained workers is about 12%, with a standard deviation of 3.4%.[18]

Our identification approach relies on connections between workers being stable over time in the absence of the training program. We cannot test for connection persistence between the pre- and post-period directly because one channel for vertical spillovers to benefit managers is a reduction in help requests coming from frontline workers. We can, however, provide evidence that email connections in the pre-period are highly persistent. Figure A1 plots average dyad-level shares of emails sent in the "Late Pre-Period" against email shares in the "Early Pre-Period" share of emails. The early and late periods each contain 4 weeks of data, with a 5 week gap between them. In Panel A, email shares

[17]Because of the bureaucratic nature of this organization, managers do not receive emails from frontline workers in other divisions.

[18]Figure A2 displays the distribution of exposure to trained workers, as calculated by log pre-period emails received from eventually trained workers (Panel A) and the share of emails received from eventually trained workers (Panel B).

are shown to be highly persistent between workers and managers. A similar pattern of persistence is evident in Panel B, which examines emails between worker-dyads. These figures show that workers and managers who have a high share of emails with one another early on continue to do so many weeks later, suggesting that communications networks are relatively stable in this organization. If our connections measures instead were to reflect transitory communications networks rather than stable relationships, then we would be incapable of detecting a spillover that propagates through the connection network.[19]

4 Direct Returns to Training and Spillovers to Coworkers

Because of experimental variation, standard intuition when the Stable Unit Treatment Value Assumption (SUTVA) holds suggests that the direct benefits to training can be estimated by comparing goal achievement for trained workers versus untrained workers in the post-period. This estimate and the corresponding standard error come directly from Table 1.

Figure 1 shows the distribution of goal achievement changes for trained workers relative to untrained workers between the pre- and post-periods. Two key points are that: i) the density of goal achievement in the pre-period is similar for trained and untrained workers when looking across the horizontal axis, as each data point represents an equally sized bin. ii) there is a positive shift in productivity for trained workers relative to the untrained across the support of the pre-period productivity distribution. The upward shift for trained workers averages about 7 percentage points (or 10 percent), while percent gains are greater for lower performers in the pre-period. It is also apparent from Figure 1 that there are several distinct clusters of goal achievement scores, which could be due to rounding/lumpiness or to heterogeneity across groups, like divisions. The plot is similar when we include division fixed effects, suggesting the lumpiness is due to rounding.[20]

[19] Any mean reversion in connection strength will therefore likely bias our estimates to zero.

[20] The lumpiness does not occur at the expected round numbers, but this reflects that goal achievement is a weighted average of sub-components that each may be rounded.

We also estimate direct training returns using difference-in-differences regressions. This enables us to incorporate worker fixed effects and division-by-time fixed effects, which allows us to account for one possible source of lumpiness or clustering in the goal achievement data. Our simplest estimator is a two-way fixed effects model of the form:

$$\log(y_{it}) = \beta_i + \beta_t + \delta_1 Trained \times Post + \delta_2 Trained \times Post \times X_i + \varepsilon_{it} \tag{4}$$

where the main coefficient of interest is δ_1. In addition, δ_2 captures potential treatment effect heterogeneity through interactions with characteristics X_i. In practice, because we only have 63 trained workers, the ability to detect heterogeneous treatment effects will be limited to coarse characteristics. Individual fixed effects are captured through β_i and time fixed effects through β_t. We cluster standard errors by worker.

Table 3 contains difference-in-differences estimates confirming the increase in goal achievement. Because the dependent variable is log goal achievement, the coefficients can be interpreted roughly as percentage changes. The coefficient on *Trained x Post* of 0.105 indicates that goal achievement for trained workers increased by about 11 percent from a baseline of 72 percent, implying that training raised goal achievement by nearly 8 percentage points. The magnitude of the implied change is slightly larger than the cross-sectional estimate in the summary statistics. Columns 3 and 4 add interactions to test for heterogeneity by wage band. In the absence of division-by-time fixed effects (Column 3), there is no differential effect of training on wage band 2 workers based on the insignificant coefficient on *Wage Band 2 x Trained x Post*. With division-by-time fixed effects in Column 4, the coefficient of -0.030 indicates that trained Wage Band 2 workers had slightly smaller goal achievement increases than wage band 1 workers. We cannot precisely identify why wage band 2 workers might have a heterogeneous response to training, but later we will show that trained wage band 2 workers became more focal in communications with other workers, which may have reduced time for their own work. High-performers, those with above-median goal achievement in the pre-period, have lower returns from training in the post-period. Finally, Columns 5 and 6 report a post-LASSO OLS specification where we select regressors that determine treatment effect heterogeneity.

Only the main effect survives the LASSO.

4.1 Robustness

SUTVA Violations: In the potential outcomes framework, equation (4) stipulates that counterfactual expected log productivity in the post-period for workers who are not trained equals $\beta_i + \beta_t$. This imposes the assumption that there are no spillovers to untreated workers. To account for potential SUTVA violations, we follow De Grip and Sauermann (2012) and modify the model to allow a general form of spillovers to untrained workers so long as the mechanism for spillovers is through pre-period connections with trained workers. Let $g(Connections, \theta)$ be a function that captures the impact of connections between trained coworkers and untrained workers with parameters γ. Then

$$\begin{aligned} \log(y_{it}) &= \beta_i + \beta_t + \delta_1 Trained \times Post \\ &+ (1 - Trained) \times Post \times g(Connections, \theta) + \varepsilon_{it} \end{aligned} \tag{5}$$

captures potential spillover impacts that will influence untrained workers' outcomes because of SUTVA violations. Under this specification, the estimate of δ_1 is the effect of training relative to an untrained worker who is unconnected to those who become trained.

Appendix Table A1 contains the results. In all specifications, the main treatment effect of training remains positive and significant, with point estimates that are larger than those in Table 3. In the first 4 columns, we find positive spillovers to coworkers using a parsimonious set of variables for capturing co-worker connections. We have also experimented with various combinations of measures to capture connections and exposure to training. The main result, which can be succinctly summarized on the last 2 columns of the table, is that none of the connections measures changes the interpretation that trained workers had positive productivity treatment effects from the program. In addition, the possible connection variables do not survive LASSO variable selection procedures for choosing a sparse model of productivity evolution. To the extent that spillovers to peers are present, they will make training investments look more favorable. As these spillovers have been shown elsewhere, we document that there is some evidence of positive spillovers. However, in our context, the empirical evidence is not especially robust. Later, when we examine vertical spillovers to managers, we will check the robustness of our results by including the effects of horizontal spillovers to coworkers that may also influence managers' interactions with untrained workers.

Attendance: The increased goal achievement of trained workers does not appear to be driven by changes in absenteeism rates. Appendix Table A3 shows that there is no differential absenteeism in the post-period for trained workers relative to untrained workers.

4.2 Cost Savings for Workers

Based on the formula for cost savings in equation (3), we have that $\tilde{W}_w = 1.052 \times 63$, while the estimates indicate that $(1 - \frac{\overline{Q}}{Q^1}) = (1 - \frac{.719}{.785})$, yielding a direct cost savings from worker training that can be interpreted as the equivalent of 5.6 frontline workers' normalized wages. This is equivalent to about 1.06% of the wage bill of frontline workers. That is, after training, if the organization could adjust headcount, it would be able to save costs equal to the compensation of about 1% of their frontline worker wage bill in every period and would still achieve the pre-period level of output. These figures are flow cost savings (meaning that the organization could realize these savings in every period), while their present value will depend on the persistence of the gains from training.

4.3 Discussion

Finding direct returns to training for workers is not trivial, as the empirical evidence on the efficacy of training programs is mixed. McKenzie (2021) and Card et al. (2018) suggest that, among other characteristics, the programs that find positive returns are usually intensive training programs that emphasize human capital accumulation.[21]

There are two possible reasons why our estimates of frontline workers' goal achievement

[21] Most of the existing studies consider programs sponsored by governments for business/entrepreneurs (McKenzie (2021)) or the unemployed (see the meta-analysis in Card et al. (2018)).

changes may overstate changes in productivity or output. First, despite the separation of the evaluation/oversight function and the rest of the organization, there may be demand effects – where evaluators rate trained workers more highly than is warranted based on actual output. Second, the increase in goal achievement may reflect trained workers' improved ability to game the evaluation system. Both possibilities are difficult to evaluate from data on frontline workers alone, but our upcoming analysis of spillovers to managers and changes in communication patterns help to assuage these concerns. If training does not raise workers' productivity, we would not expect to find evidence of vertical spillovers to managers (where we would expect manipulation of goal achievement metrics to be much more difficult).

5 Vertical Training Spillovers to Managers

As mentioned previously, there are no defined teams in our setting, so we infer connections between managers and workers from emails. Our preferred measure from the theory is the log number of emails with trained workers, which is likely to be closely related to managers' time commitment spent interacting with frontline workers. Figure 2 shows the evolution of goal achievement between the pre-period and the post-period, split by whether managers are below the 25th percentile in connections or above the 75th percentile in connections to trained frontline workers. There is a noticeable shift in the distribution for more connected managers.

In the figure, there are relatively few managers with average pre-period goal achievement below 60%, but the density of connected and not connected managers is similar for these lower performing managers. Across the support of the pre-period productivity distribution, there are highly connected managers and those with less exposure to trained workers. This provides some preliminary evidence on the plausibility of our identification strategy, which assumes that in the absence of training that managers with greater and lesser degrees of connections with eventually trained workers would have a similar evolution of goal achievement. Figure A3 examines manager characteristics that might be correlated with their exposure to trained workers. The figure shows that manager pre-period goal achievement, wages, wage band, gender, and education do not predict pre-period connections with eventually trained workers either individually or jointly.

To measure vertical training spillovers to managers, we regress managers' log goal achievement on pre-period connection measures with trained workers, interacted with a post-period indicator. Our main estimating equation is

$$\log(y_{it}) = \gamma_i + \gamma_t + \delta_1 C_{Pre} \times Post + \varepsilon_{it}, \tag{6}$$

where C_{Pre} is a measure of pre-period connections with trained workers, either the log of pre-period emails with eventually trained workers or the share of pre-period emails with eventually trained workers, γ_i is an individual fixed effect, and γ_t is a time fixed effect. Some specifications also include division-by-time fixed effects, which accounts for the possibility that goal achievement in different divisions my have evolved on different trends.

Panel A of Table 4 displays the estimates when our connection measure is the log of pre-period emails received from trained workers, while Panel B shows results when we use the share of emails from eventually trained workers as the connection measure. In both Panels A and B, and across all columns, managers who have stronger pre-period connections with eventually trained workers have greater productivity gains in the post-period. In Panel A, average implied effects for the level of goal achievement range from a 2.01 to a 2.11 percentage point increase.[22] The interquartile range (IQR) of the change in goal achievement due to connections in Panel A is about 4 percentage points. The qualitative patterns are similar in Panel B, with estimates of the average spillover to managers' goal achievement of between 2.11 and 5.13 percentage points.

[22]Because no manager has zero connections, we must extrapolate to get the effect for completely unconnected manager in the baseline. As a result, these calculations take the fitted values and subtract off the post-period indicator from Column 1. The negative coefficient on the post-period indicator suggests that our model is good only locally, as all managers are somewhat connected to trained workers.

5.1 Cost Savings for Managers

Based on the formula for cost savings in equation (2), we have that $\tilde{W}_m = 2.15 \times 129$, while the most conservative point estimate indicates that $(1 - \frac{\overline{S}}{S^1}) = (1 - \frac{.708}{.729})$, yielding an indirect cost savings from vertical training spillovers that is the equivalent of 8 frontline workers' normalized wages. This indirect cost savings is itself equal to about 1.5% of the wage bill for frontline workers. The estimates of the indirect spillovers to managers are thus 42% larger than the direct (gross) cost savings for frontline workers.

5.2 Robustness and Alternative Measures and Explanations

Robustness of Main Estimates: Table 5 presents results of a variety of robustness exercises to assess the sensitivity of our estimates. In Column 1 we probe for whether controls for spillovers to other frontline workers may influence the estimates of vertical spillovers to managers. We include controls that weight email connections with untrained workers by their connections with trained workers and managers. In Column 2 we consider mean reversion in manager productivity – which may arise if some managers were artificially busy with worker requests in the pre-period or were working on difficult tasks but later have those shocks subside. We interact deciles of pre-period goal achievement with the post dummy to account for the possibility that lower performing managers' productivity may naturally change over time. In Column 3 we include the log of emails or the share of emails that managers receive from females and those with a college degree, accounting for potential imbalance in treatment. While our connections/exposure measures utilize emails received, in Column 4 we include measures of managers' sending emails to trained workers in the pre-period and an interaction with the post-period dummy. In Column 5 we allow for the possibility that connected managers' may have changed their communications networks with untrained workers, potentially altering the stability of communications that we rely on to infer connections. To capture this, we include contemporaneous emails with untrained workers as a control. In Column 6 we control for total log emails with frontline workers in the pre-period interacted with the post-period dummy, as managers who are busier may be on different trends. Across these specifications, our point estimates are quite stable and within the confidence intervals implied by the estimates in Table 4.

These results suggest that potential violations of the Stable Unit Treatment Value for managers are not influencing our estimates. Figure A4 explores another possible change that may affect our inference, in particular whether managers who are more connected to eventually trained workers receive a different number of emails from untrained workers in the post-period. The obvious concern is that connections with trained workers may mean that these managers stop helping workers altogether, and receive fewer emails from the untrained, which would overstate the impact of training on connected managers' workload. Instead, if anything the figure shows that managers with greater connections to trained workers receive relative more emails from untrained workers, but the effect size is small and the regression coefficient is imprecisely estimated. If the channel for our estimates is manager workload, this suggests that the raw estimates of manager goal achievement changes may be biased downward if managers who are more exposed to trained workers take on an increase share of helping responsibilities for untrained workers. This form of SUTVA violation likely means that our estimates of training spillovers to managers are conservative. We also check for and find no evidence of changes in communication patterns between managers in the post-period (see Figure A5).

Alternative Email Measures: Although incoming help requests from workers, and thus inbound emails to managers, are the natural measure of connections from the theory, there is no ex-ante connection measure that comes from the environment. As a result, in Table A2 we explore alternative definitions of manager connections to trained workers by including connections based on emails sent from managers to workers. As measures of received and sent emails are correlated, it is an empirical question regarding which best explains managers' evolution of goal achievement. The table shows that it is only connection measures based on incoming emails that survive a LASSO variable selection procedure that penalizes overfitting. The table notes provide details about how the LASSO procedure is implemented with cross-validation that accounts for clustering of the data at the manager level.

Perceived Versus Actual Productivity: We would overstate the vertical spillover gains if managers do not actually become more productive but are instead perceived to achieve more because their connected workers' goal achievement increases. Given the institutional setup in this organization has managers evaluated against different tasks than frontline workers, there is little scope for claiming credit for connected trained workers' achievements. We also attempt to provide evidence evaluating whether managers and workers' goal achievement co-move in the pre-period, which would provide some suggestive evidence on whether our manager goal achievement estimates are overstated. However, testing for co-movement in goal achievement is difficult because managers are expected to become more productive if their connected workers need little help, which may show up as increases in their own goal achievement. As a result, our test is designed to assess whether manager goal achievement increases with connected workers' goal even after we control for contemporaneous email volume between workers and managers. Table 6 presents the results of this exercise. Across the 4 columns in the table, we find that manager goal achievement is negatively related to weekly email volume from frontline workers, with elasticities of around -0.09. Connected worker goal achievement, weighted based on the leave-out-mean of emails from worker j to manager i in all other weeks, is not statistically different from zero and actually has a negative coefficient. An alternative measure that looks at transitory changes in workers' goal achievement, rather than the contemporaneous level, is also not statistically different from zero after controlling for email volume. These findings help to validate that workers and managers are evaluated on separate, distinct goals and that our estimates of vertical spillovers are not mechanical.

5.3 ROI from Training: Benefits Relative to Costs for the Organization

Simple ROI Calculations: Computing the ROI from the training program is subtle, as there are no output prices for our study organization and headcount cannot be adjusted to realize cost savings due to employment protections. However, our cost minimization approach from the theory allows us to compute a ROI as-if the organization could adjust headcount. In this case, the program benefit is simply the hypothetical cost savings if the organization could re-optimize, which we compare to program costs. To compute the as-if ROI, we sum the benefits over all affected workers and net out the fixed and administrative costs of the program.[23]

Accounting for spillovers to managers meaningfully changes the implied attractiveness of the program when we impose very conservative assumptions about program costs and the persistence of gains. Table 7 presents calculations of program return on investment under a variety of scenarios that alter the assumptions about the persistence of training plus spillover gains and the size of the spillovers to managers. We present three separate scenarios for the size of the spillover to managers. The base scenario is analogous to using workers alone as the unit of analysis, where an organization would assume spillovers are 0. The second scenario assumes moderate spillovers of 1 percentage point to managers. The third scenario uses our treatment estimates of a 2 percentage point spillover to managers. Within each scenario, we consider different horizons for the persistence of treatment gains. Our most conservative assumption is that the program gains last through 6 months post-training and then depreciate completely. We then consider gain horizons of 1 year and 18 months. For each scenario, we assume that the opportunity cost of the program is equal to workers' wages when they are in training rather than producing output, as these wages are what the organization was revealed to be willing to pay for the output produced in the pre-period.

The second to last column presents the ROI from the direct training benefits, inclusive of the fixed administrative costs of the program and the opportunity cost of worker time, but excluding the vertical spillovers to managers. The last column includes the vertical spillovers into the ROI calculation. The ROI is -37% when considering only the direct

[23] A different conceptual setup for this analysis assumes that marginal output gains from the program are valued by the organization at average wages paid to managers and workers, respectively. Using this assumption, the program benefit in monetary terms is $(GA_{Post} - GA_{Pre}) * \frac{W_{Pre}}{GA_{Pre}}$ where GA_t is the average goal achievement in year t and W_{Pre} is the wage bill for the worker in the pre-period. The expression W_{Pre}/GA_{Pre} is the price-per-goal paid in the pre-period and $GA_{Post} - GA_{Pre}$ is the year-over-year change in goal achievement. This approach and the cost minimization approach yield similar conclusions. In an organization with rent-sharing between workers and firms, this approach would yield a lower bound on aggregate program benefits.

returns at a 6 month horizon. Adding just a 1 percentage point gain in goal achievement for managers turns the ROI positive, to 3%, at this short horizon. At a 1 year horizon, the direct return ROI is 25% and adding a vertical spillover to managers increases the ROI to 106% and 187% for a 1 and 2 percentage point spillover, respectively. At longer horizons, the program ROI is positive with and without the vertical spillover to managers, but the spillover greatly increases the implied return to the organization.

At first glance it wouldn't be obvious that vertical spillovers could be so valuable, but the large gains come from two sources. First, there are more managers than trained workers, so smaller gains in goal achievement are spread over more people. Second, from Table 1, managers earn more than twice as much as trained workers, so the money metric gives them more weight because the organization is willing to pay more for each goal they achieve.

Alternative Uses of Funds Spent on Training: A different approach asks whether the funds spent on the training program could have been put to better use by hiring additional managers, which may have had even greater returns if these new hires could have freed up time for incumbent managers. We attempt to calculate under what conditions the training investment would have been preferred to direct changes in the firm hierarchy. As we will show in Section 6, the channel of manager gains appears to be a reduction in demands on their time (busyness). As a result, a simple proportional rule for how one additional manager increases incumbent productivity suggests that the gains will be approximately equal to the number of new hires over the number of incumbent managers. The funds spent on training could have been used to hire about 1 manager for 12 months. A proportional increase in productivity implies that average incumbent manager's goal achievement would increase from 70.8% to 71.3%, which is smaller than our estimated gains from training spillovers. However, this increase for managers has no opportunity cost for workers, and arguably, hiring an additional manager may also increase the speed and quality of answers provided to workers. To obtain the same benefit from training workers as what we calculate in Table 7 (using the 1 year horizon with a 1 percentage point increase in manager productivity and with opportunity costs of worker training time), the hiring of an additional manager would need to increase the productivity of *all frontline workers* by about 1.1 percentage point. In other words, training 63 workers is equivalent to hiring one additional manager if each lower-level workers increase their productivity by more than 1 percentage point. Because the reduction in manager busyness after training did not raise productivity for untrained workers, it is doubtful that alternative uses of funds would have been more effective than the training program.

6 Evidence on Mechanisms and Discussion

6.1 Communication Pattern Changes Support the Hierarchies Mechanism

This section explores changes in communication patterns, which provides context for our findings while enabling us to examine mechanisms. The core comparative static in hierarchical organizations models is that more skilled workers stop asking managers to help on tasks that they can handle themselves. As a result, the prediction is a negative relationship between changes in worker skills and outgoing communications to managers. Figure 3 shows changes in log emails between the pre- and post-periods according to sender and recipient type, which we extract from difference-in-differences regression coefficients. For each sender and recipient type, we distinguish between the untrained baseline change in log emails (purple) and the change for trained workers (light green).

Apparent in this figure is that emails sent to managers from trained workers drop dramatically. There is also a smaller decline in emails to managers from untrained frontline workers in wage band 1. The large reduction in trained workers' emails to managers, both in absolute terms and relative to untrained workers, is consistent with the hierarchies mechanism such that an increase in workers' skills reduces requests for managerial help. Differences across wage bands provide a possible explanation for why emails to managers also decline for untrained workers. Untrained wage band 1 workers reduce their emails to managers – but they send more emails to trained wage band 2 workers. Untrained wage band 2 workers have almost no reduction in emails sent to managers. This pattern suggests trained wage band 2 workers begin to substitute for managers amongst untrained wage band 1, but not wage band 2, workers.[24] These changes are consistent with untrained workers beginning to rely on trained Wage Band 2 workers for help, representing something like an informal additional layer in the hierarchy of the organization. While most studies of hierarchies use external shocks to study the number of layers in an organization, this suggests that augmenting worker skills may lead to informal layers emerging as a response to new capabilities. We note, however, that our identification of spillovers to managers is not impacted by the emergence of the Wage Band 2 workers as a source of help, as managers' exposure to trained workers captures the direct effect based on direct cross-sectional differences, while our estimates are little changed when we control for potential spillovers to untrained workers from exposure to trained workers (see Table 5).

An additional test involves the relationship between manager productivity and changes in emails from eventually trained workers. Consider two potential mechanisms that correspond to different production processes. In the first, managers and frontline workers collaborate on tasks, and thus need to communicate to work together. As in Kremer (1993), when one member of a team becomes more productive, the team can do more if that member was originally the bottleneck in production. Under this mechanism, raising frontline worker productivity can lead to collaboration on additional tasks, which would predict an increase in communications. In the second mechanism, managers help workers only when needed. The helping mechanism in the hierarchies model predicts that as workers gain skill, manager productivity will increase as email volume from frontline workers falls.

To test between these two mechanisms, we regress changes in manager log goal achievement on changes in the number of emails from eventually trained workers (who become

[24] Appendix Table A4 presents the full set of changes in emails between sender and receiver pairs. Emails originating from managers and sent to frontline workers also decline, with a 17% reduction for trained Wage Band 1 workers and a 14% reduction for untrained Wage Band 1 workers. Managers reduce emails to Wage Band 2 workers, both trained and untrained, by between 4 and 5%. The smaller changes in emails sent from managers to workers may arise because not all help requests receive an email response – that is, emails may serve as a ticket system to alert managers to a problem, while they respond using different mediums.

more productive).[25] Table 8 presents this test. Columns 1 and 2 display OLS regressions. The coefficients range from -0.07 to -0.15 per 100 email change for trained workers. These results are substantial and imply large increases in goal achievement for managers who receive fewer emails from trained workers. In fact, the -0.07 coefficient implies that the average manager would have increased goal achievement by about 6% (or about 4 percentage points) as a result of the average reduction in emails from trained workers. To match the actual change in goal achievement, we note that the constant term is negative. The coefficient is sensitive to division fixed effects, which likely reflects some treatment effect heterogeneity. This heterogeneity appears to come from the baseline level of manager busyness. When we run the regression division-by-division, we find the largest effects in the Execution division, which is the division with the largest ratio of workers to managers and is the division where managers received the most emails from frontline workers in the pre-period.[26]

It is also possible that the change in emails with trained workers reflect endogenous choices to reallocate work. To deal with endogeneity, Columns 3 and 4 report IV regressions. The instrument for the change in emails with trained workers is the pre-period number of emails with eventually trained workers. The IV coefficients range from -0.07 to -0.095, indicating that managers who had the largest declines in emails with eventually trained workers had the largest increases in goal achievement. The IV estimate without division fixed effects is nearly identical to OLS, while it is smaller than the OLS coefficient when division fixed effects are included. The last two columns report the first stage regressions, which have effective F-statistics of over 4000 and 59, respectively, implying a

[25] To this point, we have used the log number of emails with trained workers as a measure of exposure to training. This eases the interpretation of parameter estimates, but using the change in log emails with trained workers as the right hand side variable is problematic in IV regressions that use exposure to trained to instrument for changes in emails. The reason is that trained workers' emails fall by about the same percent, so instrumenting the change in log emails with pre-period log emails is a weak predictor of percent changes that appear independent of the original pre-period baseline. Instrumenting for the change in the number of emails with the pre-period level of emails with trained workers alleviates this issue.

[26] We get similar results when we regress changes in log goal achievement on changes in shares of emails with trained workers and division fixed effects, but these estimates are harder to interpret if the channel is manager busyness. As mentioned in the discussion of Figure 3, changes in shares of emails with trained workers arise because the number of emails from trained workers falls relative to emails from untrained workers.

maximal bias of under 5% (Olea and Pflueger, 2013).

6.2 Survey Evidence

We also conducted a survey in August of 2020 to improve our understanding of mechanisms. The organization distributed the survey to 63 trained workers and 105 untrained workers who were present in the pre-period.[27]

One of the main concerns with analyzing interactions through email communication is that workers have alternative communication modes that may substitute for or complement emails, like face-to-face interaction or phone calls.[28] To proxy for other forms of communication, the survey asked the respondents about the frequency of face-to-face interaction with those that they interact with through email communications. Figure 4 shows that the majority of workers interact either several times a week or at least once a week with those that they send emails frequently, suggesting that electronic and face-to-face communication are complements.

The survey also allows us to assess the reasons for email contact between frontline workers and managers. Figure 4 shows that 3 out of 4 workers reported that the main reason they contact superiors is to ask for help, with the other responses split evenly between asking for authorizations and reporting on task progress.

The survey further asked respondents about their perceptions of changes over time between the pre-period and the post-period. Table 9 shows that trained workers reported much greater improvements in their knowledge of task requirements, their understanding of division-appropriate workflows, and general skills and knowledge relative to untrained workers. The table is inconsistent with changes in goal understanding, interdependence, changes in hours worked, or directives to reduce help requests to managers. The table

[27] The survey contained 7 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 not incentivized. Fifty-two percent of the trained workers (N=33 workers) and 54% of the untrained workers (N=57) took the survey. The completion rate is in line with average response rates in organizational research Baruch and Holtom (2008). Appendix B contains the English version of the survey.

[28] During the sample, the organization prohibited the use of other communication technologies such as WhatsApp and Skype.

transmits a simple message, trainees report that the training program improved their skills.

6.2.1 Alternative Explanations

Survey questions are also useful to understand potential alternative explanations. Reduced monitoring or supervision does not appear to explain the results, as the top-right panel of Figure 4 shows that among trained and untrained workers, 85% of workers think that the supervision level remained constant through the pre- and post-periods. Only 3.3% of respondents think that monitoring effort by managers decreased. The training program also did not appear to change career incentives. While we see no upward mobility through the organization during our sample, we asked directly whether survey respondents thought that their promotion possibilities increased from 2018 to 2019. Table 9 shows no differential changes in promotion prospects by worker training status. For this organization, promotion from within is rare, making career concerns unlikely to explain our results. A different possibility is that the nature of workers' tasks changed over time. Table 9 suggests tasks did not change. In particular, the vast majority of both trained and untrained workers thought there was no increase in task interdependence, with only 6.1% of trained workers and 5.3% of untrained workers reporting an increase in interdependent tasks. Finally, a different possibility to explain the productivity increase from trained workers is that they became more motivated, changing their labor supply. Table 9 shows that while 6.1% of trained workers increased their working hours in a week, 5.3% of untrained workers also did so. We cannot reject that these results differ, and the small mean differences indicates that internal incentives to work more are unlikely to explain the increase in goal achievement from trained workers.

The survey also provides context around why trained workers decreased communications with managers. The survey presented these patterns and then asked whether the change was a result of communication from the organization's leadership. Table 9 shows that there were no differential perceptions of directives to change help requests to managers from above. Under 3 percent of respondents believed that managers formally became a less important source of help. As a consequence, changes in communication patterns likely arise organically, rather than as a result of a formal dictate or reorganization from leadership.

6.3 Discussion

An area for future work is to consider how to target who gets training and how many workers should optimally be trained. Other literature suggests that returns to workplace programs are heterogeneous, so getting the targeting rules right may depend on understanding personalized returns as well as spillovers (Sandvik et al., 2021). In other contexts, these spillovers may be a function of social networks (Bandiera et al., 2010). Another implication is that training might be correlated with having relatively flat organizations, a conjecture which may provide fertile ground for further empirical work in the spirit of Rajan and Wulf (2006) and Guadalupe and Wulf (2010). All else equal, training liberates managers' time, allowing them to have larger spans of control.

7 Conclusion

There has been a growing interest in understanding the returns of training programs in different countries, industries and settings (Card et al., 2011; Attanasio et al., 2011; Hirshleifer et al., 2016; McKenzie, 2017; Card et al., 2018; Alfonsi et al., 2020). The literature has mainly focused on providing estimates of the effect of these programs on trained individuals, but more limited attention has been paid to the potential spillover effects of training. Using frontline workers' randomization into training in a Colombian government organization, we study changes in productivity for trained workers as well as spillovers to managers.

We find significant direct benefits to the training program for those workers randomized into it. Less appreciated, but of greater consequence to the calculation of the organization's returns from the program, are spillovers to managers higher in the organizational hierarchy. We find productivity spillovers to managers are economically significant and large enough to change the organizations decision rule to offer training to its workers. To understand the mechanism behind spillovers, we examine changes in email communications and a survey of employees. Both sources are suggestive that spillovers to managers arise from reduced needs to assist lower level workers. These results indicate the importance of considering production hierarchies and organizational structure when accounting for the returns to training or skill upgrading in organizations. To the best of our knowledge, this is the first paper to quantify this channel for different hierarchical layers in an organization. In doing so, we shed new light on the microeconomic foundations of organizational economics models of hierarchies that have become ubiquitious for explaining patterns of production and inequality at the individual and macro level. In testing the core comparative static of organizational economics models that increasing worker skill reduces the need to engage managers for help, we provide additional implications beyond existing tests that rely on shocks to opportunities or variation in the size of a market. In particular, as new technologies like massive online open courses (MOOCs) potentially reduce the cost of offering employer-sponsored training to frontline workers, our results suggest that managers may greatly benefit even if the cost of upgrading manager skills does not change.

At least since the Training Within Industry program in World War II, scholars have focused on studying how upgrading management quality influences employees on the bottom of a hierarchy (Bianchi and Giorcelli, 2022).[29] One of the main lessons from our study is that influence does not necessarily travel downward. In this paper, we have provided some of the first empirical evidence that employees in lower levels of a hierarchy can impact those at the top. Our work validates the core comparative static predictions in Garicano (2000) and helps to provide a potential rationale for the positive assortative matching between workers and managers documented across firms and other contexts (Bloom and Van Reenen, 2007).

[29] There is extensive research on how managers have an effect on lower level employees (Lazear et al., 2015; Bloom et al., 2015, 2020).

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., Kala, N., and Nyshadham, A. (2018). The skills to pay the bills: Returns to on-the-job soft skills training. Technical report, National Bureau of Economic Research.

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., and Vitali, A. (2020). Tackling youth unemployment: Evidence from a labor market experiment in uganda. *Econometrica*, 88(6):2369–2414.

Attanasio, O., Kugler, A., and Meghir, C. (2011). Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial. *American Economic Journal: Applied Economics*, 3(3):188–220.

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., Barankay, I., and Rasul, I. (2010). Social Incentives in the Workplace. *The Review of Economic Studies*, 77(2):417–458.

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.

Baruch, Y. and Holtom, B. C. (2008). Survey Response Rate Levels and Trends in Organizational Research. *Human Relations*, 61(8):1139–1160.

Besley, T. and Persson, T. (2010). State Capacity, Conflict, and Development. *Econometrica*, 78(1):1–34.

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.

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., Ibarraran, P., Regalia, F., Rosas-Shady, D., and Soares, Y. (2011). The Labor Market Impacts of Youth Training in the Dominican Republic. *Journal of Labor Economics*, 29(2):267–300.

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.

Cavounidis, C. and Lang, K. (2020). Ben-porath meets lazear: Microfoundations for dynamic skill formation. *Journal of Political Economy*, 128(4):1405–1435.

Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service. *The Quarterly Journal of Economics*, 128(3):1169–1218.

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.

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. (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.

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.

Hirshleifer, S., McKenzie, D., Almeida, R., and Ridao-Cano, C. (2016). The Impact of Vocational Training for the Unemployed: Experimental Evidence from Turkey. *The Economic Journal*, 126(597):2115–2146.

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.

Kremer, M. (1993). The o-ring theory of economic development. *The Quarterly Journal of Economics*, 108(3):551–575.

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. (2009). Firm-specific human capital: A skill-weights approach. *Journal of political economy*, 117(5):914–940.

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.

McKenzie, D. (2017). How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence. *The World Bank Research Observer*, 32(2):127–154.

McKenzie, D. (2021). Small business training to improve management practices in developing countries: re-assessing the evidence for 'training doesn't work'. *Oxford Review of Economic Policy*, 37(2):276–301.

Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369.

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. (2021). Should workplace programs be voluntary or mandatory? evidence from a field experiment on mentorship. Technical report, National Bureau of Economic Research.

Sandvik, J. J., Saouma, R. E., Seegert, N. T., and Stanton, C. T. (2020). Workplace Knowledge Flows. *The Quarterly Journal of Economics*.

Figures and Tables

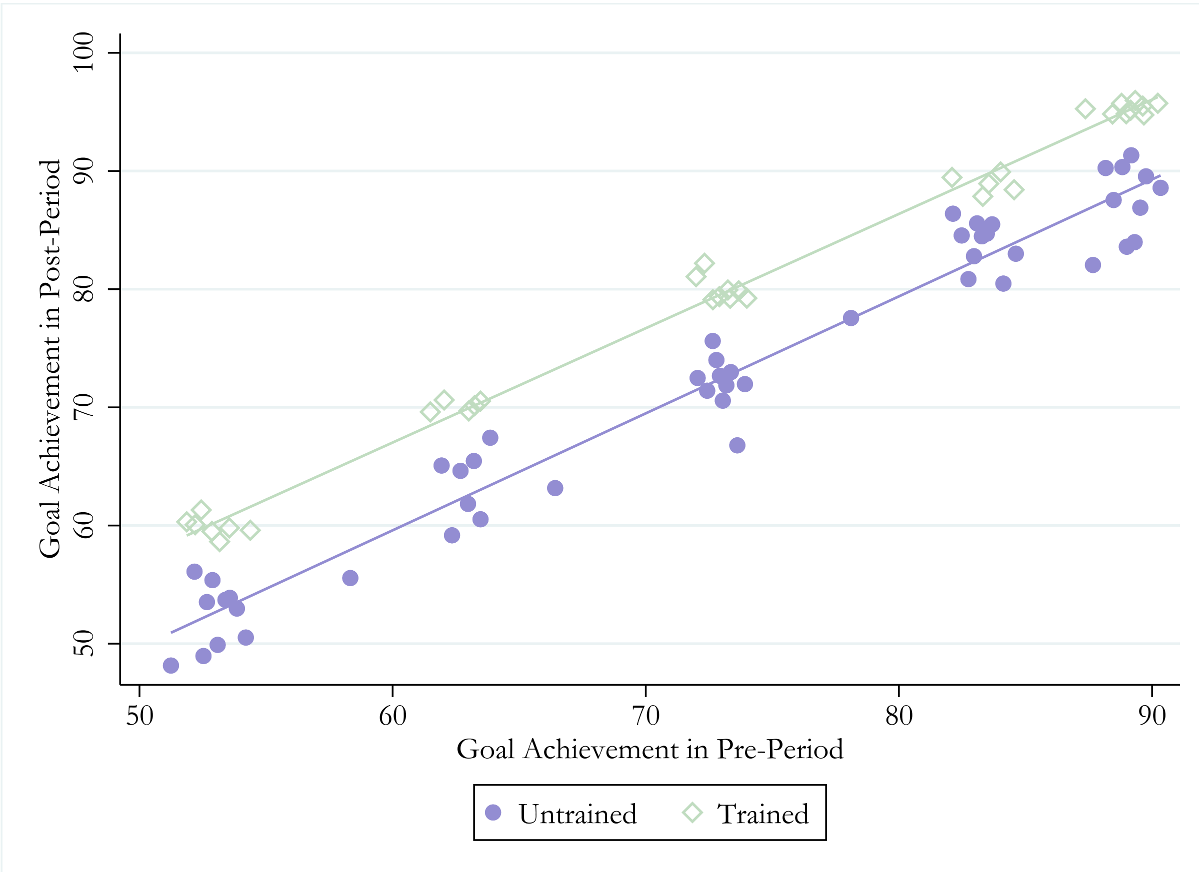


Figure 1: Goal Achievement Changes For Workers by Training Status

Note: This figure plots the relationship between average post-period and average pre-period individual goal achievement for frontline workers. We plot the relationship separately based on whether the worker was randomized into the training program or not.

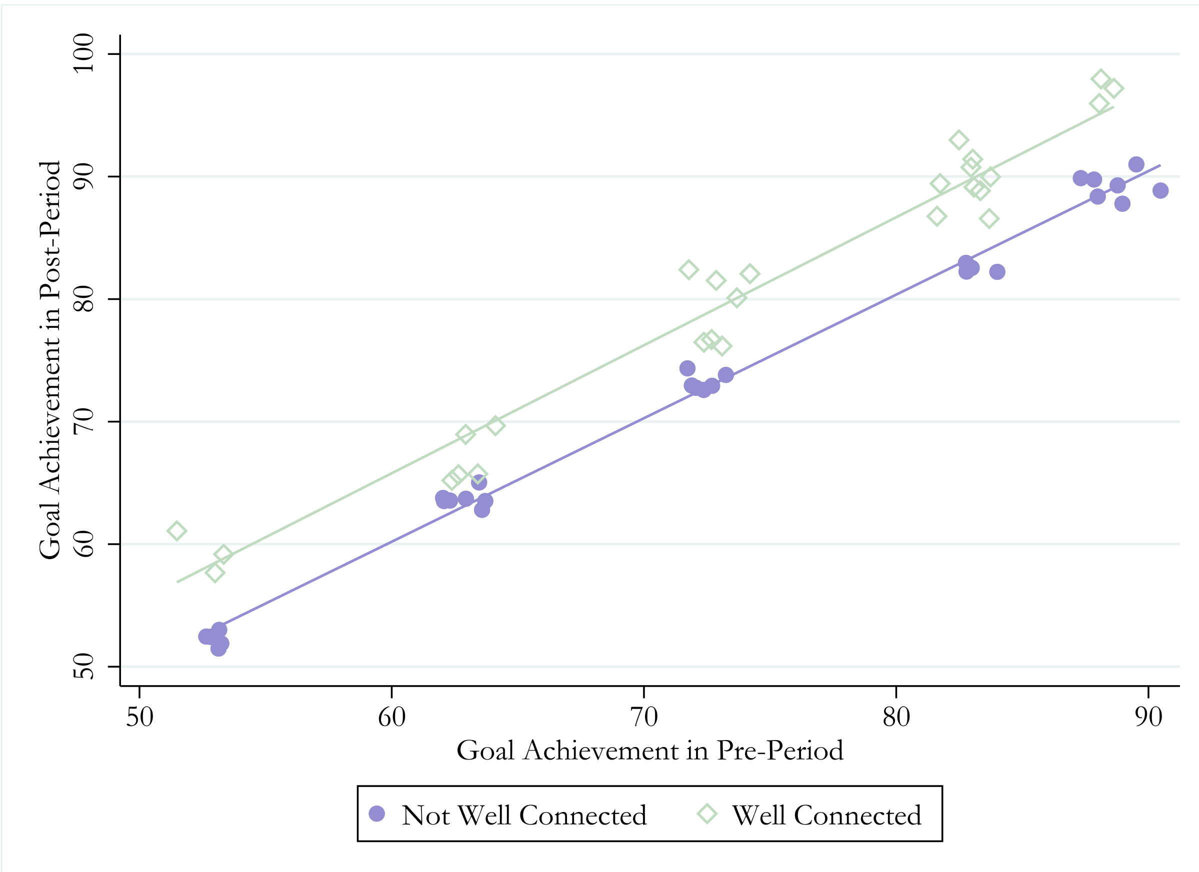


Figure 2: Goal Achievement Changes For Managers by Connection Strength to Trained Workers

Note: This figure plots the relationship between average post-period and average pre-period individual goal achievement for managers. We plot the relationship separately based on the strength of a manager's connections to eventually trained workers. A manager is defined as well-connected if he or she is above the 75th percentile of pre-period emails with eventually trained workers from the division and a manager is not well connected if he or she is below the 25th percentile.

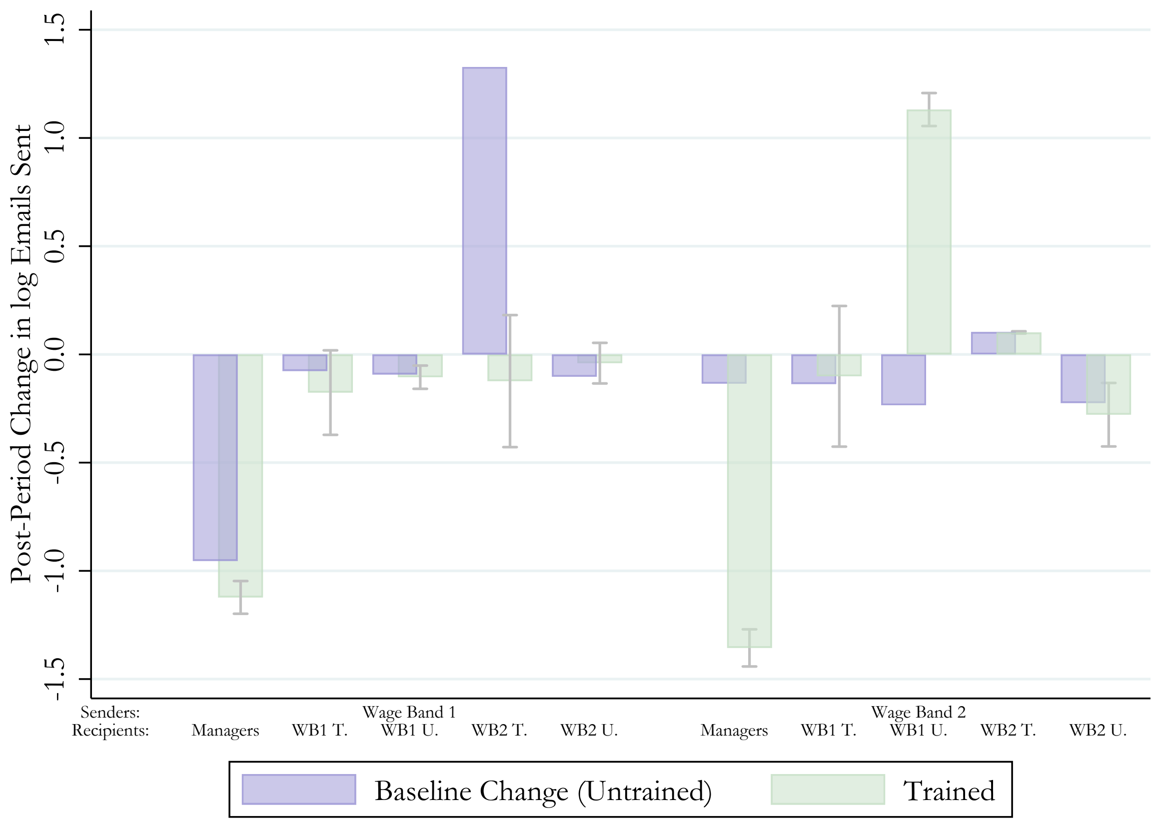


Figure 3: Changes in Log Emails Between the Pre- and Post-Period by Sender/Recipient Type

Note: This figure displays the average change in log emails between the pre- and post-periods by senders' wage band, recipient type, and training status. Purple bars are the average changes for untrained workers to each recipient type. 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.). Green bars are the change for trained workers, with standard errors 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 type, and include fixed effects for workers and time. Standard errors are clustered by sender.

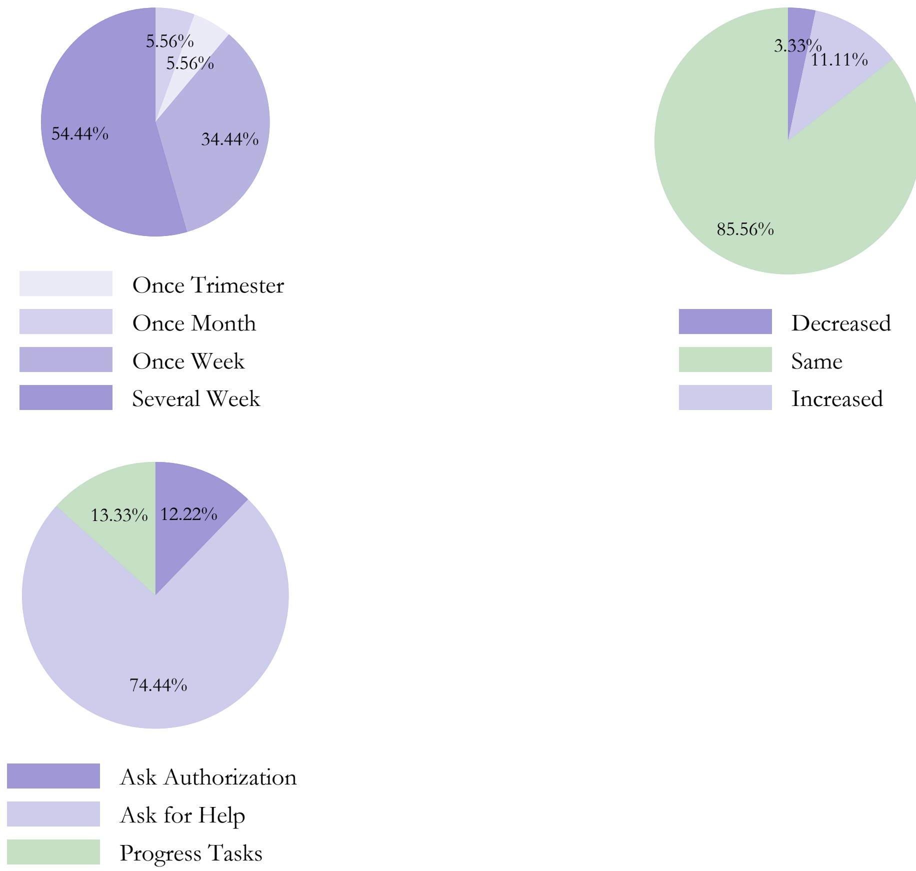


Figure 4: Distribution of Survey Responses to Questions Regarding the Mechanism

Note: This figure displays answers to an ex-post survey designed to understand the environment and mechanisms behind 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 used to interact with by e-mail 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)."

	(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	0.452	0.244	0.310	0.378	0.429	0.051
						(0.067)
Administration	0.181	0.203	0.225	0.188	0.190	0.003
						(0.053)
Finance	0.119	0.163	0.116	0.136	0.111	-0.025
						(0.043)
Human Talent	0.119	0.233	0.147	0.162	0.111	-0.051
						(0.043)
Planning	0.130	0.157	0.202	0.136	0.159	0.023
						(0.049)
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	1.195	2.155	1.065	1.052	-0.014
	(0.410)	(0.452)	(1.100)	(0.434)	(0.436)	(0.058)
Wages, Post-Period (Normalized)	1.045	1.249	2.252	1.113	1.099	-0.014
	(0.428)	(0.473)	(1.149)	(0.453)	(0.455)	(0.061)
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						1.075
						(0.379)

Table 1: Descriptive Statistics and Balance on Observable Characteristics

This table displays descriptive statistics for workers in Wage Band 1, Wage Band 2, and Wage Bands 3-5 (Managers). The table also provides balancing tests between trained and untrained frontline workers (columns 4-6). The last column displays t-tests of differences between trained and untrained workers across columns 4 and 5. The unit of observation is a worker. Secondary Education, Bachelors Degree and Masters-PhD are dummy variables for the highest educational level achieved. Execution Division, Administration, Finance, Human Talent and Planning are division dummy variables. Wage Band is either 1, 2, 3, 4 or 5. Monthly wages for 2018 and 2019 are normalized by taking the mean of 2018 wages for Wage Band 1 and dividing all wages by the 2018 Wage Band 1 mean. 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' pre-period observable characteristics.

	Panel A: Managers								
	(1)	(2)	(3)	(4)	(5)	(6)	Mean	SD	Obs
(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): Share of Manager Emails Sent to Trained Workers	0.33	0.32	1.00				0.12	0.02	129
(4): Share of Manager Emails Received from Trained Workers	0.53	0.25	0.62	1.00			0.12	0.03	129
(5): Log Emails from Managers to Trained Workers	0.95	0.84	0.46	0.67	1.00		7.25	0.59	129
(6): Log Emails from Trained Workers to Managers	0.91	0.98	0.47	0.37	0.88	1.00	7.28	0.54	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): Share of Emails Sent from Untrained to Trained Workers	0.23	0.49	1.00				0.12	0.03	463
(4): Share of Emails Sent from Trained to Untrained Workers	0.48	0.19	0.28	1.00			0.12	0.03	463
(5): Log Emails from Untrained to Trained	0.96	0.81	0.28	0.60	1.00		6.32	0.65	463
(6): Log Emails from Trained to Untrained	0.81	0.95	0.62	0.25	0.79	1.00	6.32	0.67	463

Table 2: 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 13 week pre-period. Email share measures divide by the total emails sent or received relative to emails with all frontline workers. The log email measures displayed in rows (5) and (6) are the measures we use in our later empirical analyses. Columns (1)-(6) display correlations between the various exposure measures.

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Goal Achievement					
Trained × Post	0.105***	0.107***	0.132***	0.152***	0.105***	0.107***
	(0.006)	(0.007)	(0.013)	(0.014)	(0.006)	(0.007)
Wage Band 2 × Trained × Post			-0.010	-0.030**		
			(0.010)	(0.014)		
Wage Band 2 × Post			0.015*	0.010		
			(0.008)	(0.008)		
Higher Education × Post			-0.003	-0.003		
			(0.010)	(0.009)		
Higher Education × Trained × Post			0.008	0.011		
			(0.012)	(0.014)		
Woman × Post			-0.001	-0.008		
			(0.010)	(0.011)		
Woman × Trained × Post			-0.007	-0.026*		
			(0.012)	(0.014)		
High Performer × Post			0.009	0.011		
			(0.010)	(0.009)		
High Performer × Trained × Post			-0.051***	-0.055***		
			(0.011)	(0.013)		
Observations	13327	13327	13327	13327	13327	13327
R-squared	0.903	0.909	0.904	0.910	0.903	0.909
Worker F.E.	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓
Division-Time F.E.	×	✓	×	✓	×	✓
Post-LASSO OLS	×	×	×	×	✓	✓
Socio-demographic Controls	×	×	✓	✓	✓	✓

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

Table 3: Treatment Effects of Training For Frontline Workers

Note: This table displays estimates of training treatment effects. The unit of observation is a 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 characteristics with the trained-by-post indicator. Sociodemographic controls in these columns are dummies for wage band 2, a college degree or more, female, and above-median pre-period goal achievement (defined as a High-Performer). Columns 5 and 6 report post-LASSO OLS regressions after selecting variables using LASSO with cross-validation of the penalty. 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)
Panel A: Log Pre-Period Emails with Eventually Trained Workers			
Log Pre-Period Emails from Trained × Post	0.063***	0.063***	0.063**
	(0.006)	(0.006)	(0.031)
Post	-0.430***		
	(0.043)		
Average. p.p. Δ GA	2.11	2.11	2.01
Interquartile Range p.p. Δ GA	4.18	4.18	4.15
N	3276	3276	3276
R^2	.95	.951	.953
Panel B: Share of Pre-Period Emails with Eventually Trained Workers			
Share of Pre-Period Emails from Trained × Post	0.294**	0.292**	0.647**
	(0.124)	(0.124)	(0.313)
Post	-0.005		
	(0.014)		
Average. p.p. Δ GA	2.13	2.11	5.13
Interquartile Range p.p. Δ GA	.769	.764	1.74
N	3276	3276	3276
R^2	.943	.943	.953
Time FE or Post-Indicator:	Post	Time	Time
Division-Time FE:	No	No	Yes

Table 4: Estimates of Vertical Training Spillovers to Managers Based on 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 log emails received from eventually trained workers. In Panel B, these measures are 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. 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 in logs and multiplies by the individual manager's average of pre-period goal achievement. These measures include the post-period constant term estimated from Column 1.

Control For	(1) Horizontal Spillovers	(2) Mean Reversion	(3) Imbalanced Treatment	(4) Sent Emails	(5) Contemp. w Untrained	(6) All Pre-Period Emails
Panel A: Log Pre-Period Emails with Eventually Trained Workers						
Log Pre-Period Emails from Trained × Post	0.064**	0.058*	0.065**	0.066**	0.066**	0.060*
	(0.032)	(0.031)	(0.031)	(0.033)	(0.032)	(0.035)
N	3276	3276	3276	3276	3276	3276
R^2	.953	.954	.953	.953	.953	.953
Panel B: Share of Pre-Period Emails with Eventually Trained Workers						
Share of Pre-Period Emails from Trained × Post	0.625*	0.715**	0.661**	0.666**	0.718**	0.648**
	(0.320)	(0.313)	(0.307)	(0.323)	(0.320)	(0.314)
N	3276	3276	3276	3276	3276	3276
R^2	.953	.954	.953	.953	.953	.953
Time FE or Post-Indicator:	Time	Time	Time	Time	Time	Time
Division-Time FE:	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: Robustness Analysis of Estimates 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 log emails received from eventually trained workers. In Panel B, these measures are 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, the connection measure for untrained workers is based on log pre-period emails, while Panel B uses 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 the log of emails or the share of emails from female frontline workers and those with a college degree or more education. Column 4 controls for log emails (share of emails) sent to trained workers in the pre-period interacted with the post-period indicator. Column 5 controls for contemporaneous weekly log emails from untrained workers, which captures changes in workload that may result from exposure to trained workers. Finally, Column 6 controls for pre-period log emails from all workers x Post. All columns include manager fixed effects and division x time fixed effects.

	(1)	(2)	(3)	(4)
Log Emails From Workers to Managers	-0.089**	-0.091**	-0.088**	-0.090**
	(0.041)	(0.042)	(0.041)	(0.042)
Leave-Out-Week Email-Weighted Worker Log GA	-0.005	-0.056		
	(0.010)	(0.067)		
Transitory Δ in Leave-Out-Week Weighted Worker Log GA			0.116	-0.843
			(0.173)	(1.262)
N	1569	1569	1569	1569
R^2	.956	.958	.956	.958
Division-Time FE:	No	Yes	No	Yes

Table 6: Regressions of Manager Log Goal Achievement on Log Emails and Connected Worker Goal Achievement in the Pre-Period

Note: The dependent variable is managers' weekly log goal in the pre-period. "Log Emails From Workers to Managers" is the number of emails received from frontline workers by manager i in week t. The "Leave-Out-Week Email-Weighted Worker Log GA" measure captures connected workers' contemporaneous 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.

Gains Horizon Spillover	Manager Spillover (Pct Points)	Direct Benefit (USD)	Vertical Spillover Benefit (USD)	ROI (%) From Direct	ROI (%) From Direct + Vertical Benefit
6 Months	1.00	49,855	32,005	-37	3
1 Year	1.00	99,710	64,010	25	106
18 Months	1.00	149,565	96,015	88	209
6 Months	2.01	49,855	64,304	-37	43
1 Year	2.01	99,710	128,608	25	187
18 Months	2.01	149,565	192,912	88	330

Table 7: Return on Investment Under Different Scenarios

Note: This table displays different scenarios for calculating program returns on investment. The table presents 2 different scenarios for the vertical spillover to managers (1 and 2 percentage points) and 3 different time horizons for the gains from training to persist (6 months, 1 year, and 18 months). The benefits from the program are equivalent to the costs the organization could save if they were able to reduce headcount while still achieving the same level of output as the pre-period (see the text for details). These estimated benefits are also approximately the implied gain if the organization values the additional output as a result of the direct gains and vertical spillovers at the average wages paid to workers and managers in the pre-period. The costs of the program include the opportunity cost, which are assumed to be the wages paid while trainees were in class for 120 hours and total $55,096, plus direct costs and overhead costs, which totalled $24,500.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS		IV		First Stage	
Change in Emails from Trained Workers / 100	-0.071***	-0.150***	-0.069***	-0.095***		
	(0.006)	(0.025)	(0.005)	(0.031)		
Pre-Period Emails Received from Trained / 100					-0.008***	-0.010***
					(0.000)	(0.001)
N	129	129	129	129	129	129
R^2	.663	.728	.662	.712	.971	.973
Division FE:	No	Yes	No	Yes	No	Yes

Table 8: Regressions of Changes in Managers' Log Goal Achievement on Changes in Emails from Trained Workers

Note: The dependent variable is the year-over-year change in manager log goal achievement. The main regressor is the year-over-year change in the number of emails from (eventually) trained workers. IV regressions instrument the change with the pre-period number of emails with eventually trained workers, as shown in the first stage regression columns. Robust standard errors are reported.

	Untrained Mean (SD)	Trained Mean (SD)	Difference (SE)
Increased 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)
Increased 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)
Increased Skills and Knowledge	0.035	0.909	0.874***
	(0.186)	(0.292)	(0.056)
Increased 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 9: Survey Results: 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 question had nine sub-components that each began with "Relative to 2018, in 2019 you:". These sub-options 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? Each sub-question had three option answers: Yes, No, Does not apply/Do not know.

A Appendix Tables

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Goal Achievement					
Trained × Post	0.170**	0.109	0.265***	0.225*	0.105***	0.107***
	(0.083)	(0.081)	(0.092)	(0.132)	(0.006)	(0.007)
Pre-Share of Emails from Trained workers × Untrained × Post	0.897***	0.636**	0.968***	0.763**		
	(0.334)	(0.304)	(0.334)	(0.316)		
Pre-Share of Emails from Untrained workers × Untrained × Post	-0.036	-0.084	0.067	0.046		
	(0.117)	(0.111)	(0.126)	(0.165)		
Wage Band 2 Worker × Untrained × Post			0.022**	0.016		
			(0.009)	(0.013)		
Advance Degree (Bachelor's or Master's) × Untrained × Post			-0.003	-0.003		
			(0.010)	(0.009)		
Woman × Untrained × Post			-0.000	-0.008		
			(0.010)	(0.011)		
Worker with high performance × Untrained × Post			0.006	0.008		
			(0.009)	(0.009)		
Avg. Horizontal Spillover	.093	.066	.1	.079	0	0
Spillover Std. Error	.036	.029	.037	.033	0	0
Observations	12834	12834	12834	12834	12834	12834
R-squared	.904	.91	.904	.91	.903	.91
Worker F.E.	✓	✓	✓	✓	✓	✓
Time F.E.	✓	✓	✓	✓	✓	✓
Division F.E.	×	✓	×	✓	×	✓
LASSO	×	×	×	×	✓	✓
Control Emails Flows	✓	✓	✓	✓	✓	✓
Control Sociodemographic	×	×	✓	✓	✓	✓

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

Table A1: 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 3. In addition, we add measures of each untrained workers' pre-period share of emails with trained and untrained workers. These measures capture the potential for horizontal spillovers, as shown in equation (5). We compute the average implied horizontal spillover and standard error below each column. None of these spillover measures survive a LASSO variable selection procedure where we cross-validate the penalty term. Columns 5 and 6 report post-LASSO OLS regressions on the variables that do survive this variable selection procedure.

	(1)	(2)	(3)
Panel A: Log Pre-Period Emails with Eventually Trained Workers			
Log Pre- Emails Received from Trained × Post	0.083***	0.066**	0.063**
	(7.69)	(2.03)	(2.01)
Log Pre- Emails Sent to Trained × Post	-0.021***	-0.005	
	(-2.76)	(-0.53)	
Average Vertical Spillover	1.83	1.46	2.01
IQR Vertical Spillover	4.27	4.11	4.15
N	3276	3276	3276
R^2	.951	.953	.953
Worker F.E.	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes
Division-Time F.E.	No	Yes	Yes
Post-LASSO OLS	No	No	Yes
Panel B: Share of Pre-Period Emails with Eventually Trained Workers			
Post x Pre-Share of Emails Received from Trained	0.299*	0.666**	0.647**
	(1.67)	(2.06)	(2.07)
Post x Pre-Share of Emails Sent to Trained	-0.006	-0.028	
	(-0.05)	(-0.29)	
Average Vertical Spillover	2.49	5.43	5.51
IQR Vertical Spillover	.849	1.83	1.86
N	3276	3276	3276
R^2	.943	.953	.953
Worker F.E.	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes
Division-Time F.E.	No	Yes	Yes
Post-LASSO OLS	No	No	Yes

Table A2: Regressions of Managers' Log Goal Achievement on Additional Measures of Email Connections with Eventually Trained Workers

Note: This table displays results of manager spillover regressions with different email measures. The dependent variable is managers' weekly log goal achievement. Panel A includes log pre-period emails sent and received from eventually trained workers and Panel B includes the pre-period share of emails sent and received from eventually trained workers. Column 1 includes Time Fixed Effects. Columns 2 and 3 include Time x Division Fixed effects. Column 3 reports post-LASSO OLS estimates of variables that survive a first-stage LASSO variable selection procedure that cross-validates the penalty. Standard errors are clustered at the manager level.

	(1)	(2)	(3)	(4)	(5)
Trained × Post	-0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	81,530	81,530	81,530	81,530	48,980
R-squared	1.000	1.000	1.000	1.000	1.000
Mean DV	.24	.24	.24	.24	.24
Worker FE	×	×	✓	✓	✓
Date FE	✓	×	✓	×	×
Division x Date FE	×	✓	×	✓	✓
Divisions	ALL	ALL	ALL	ALL	Execution Excluded

Table A3: Effects of the Training Program on Absenteeism.

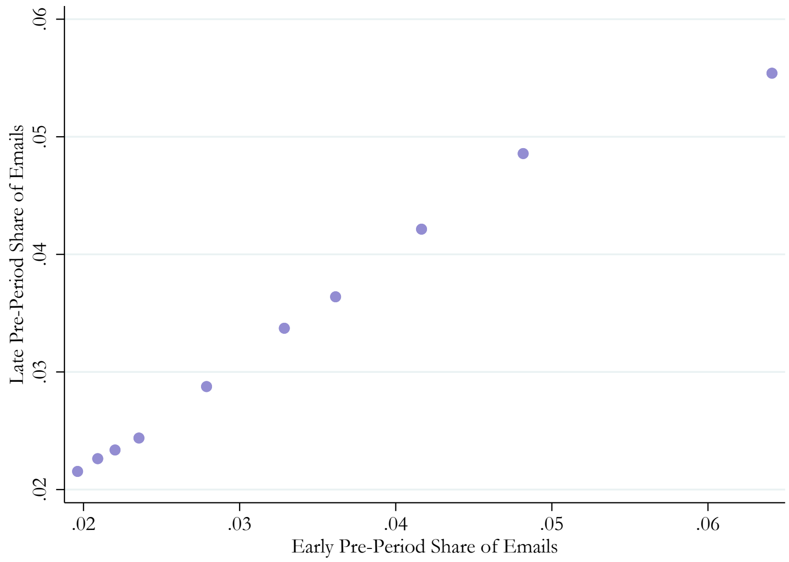
Note: Differences in differences regressions where the dependent variable is daily absenteeism. The sample includes all weekdays and Saturdays. Absenteeism is calculated as lack of email activity (as email is only available from office computers), and the dependent variable takes the value 1 if the worker did not send any email on a given workday. All models include worker and date fixed effects. The sample includes all frontline workers. Standard errors are clustered by worker.

Senders	Receipts:	Managers	Wage Band 1 Trained	Wage Band 1 Untrained	Wage Band 2 Trained	Wage Band 2 Untrained
Wage Band 1	Main Effect	-0.954	-0.076	-0.092	1.328	-0.103
	Trained Effect	-1.123	-0.176	-0.105	-0.123	-0.040
		(0.076)	(0.195)	(0.054)	(0.305)	(0.094)
Wage Band 2	Main Effect	-0.135	-0.137	-0.235	0.104	-0.225
	Trained Effect	-1.356	-0.101	1.131	0.102	-0.279
		(0.086)	(0.325)	(0.076)	(0.005)	(0.147)
Managers	Main Effect	-0.527	-0.166	-0.136	-0.043	-0.048

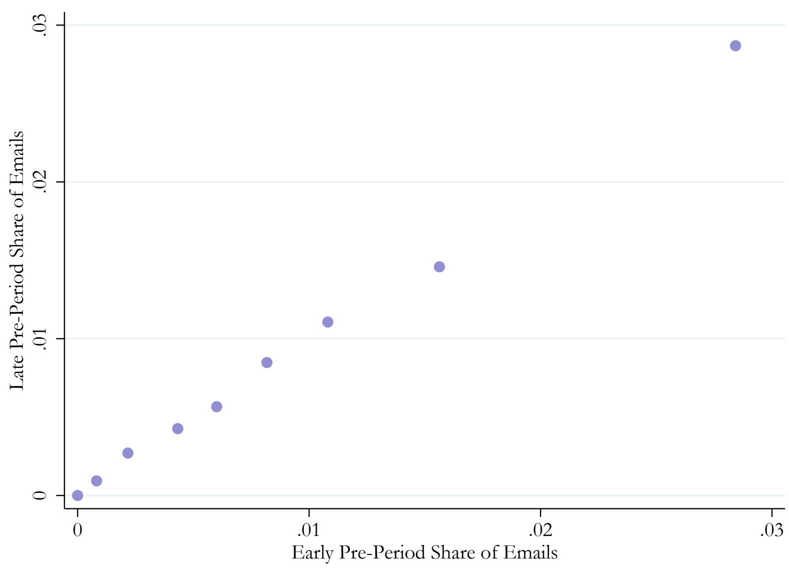
Table A4: Difference-in-differences Estimates of Changes in Log Emails

Note: This table displays coefficient estimates for the change in log emails that go into Figure 3. Columns show the receivers while rows display the senders. Each regression is run separately for an 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.

Appendix Figures



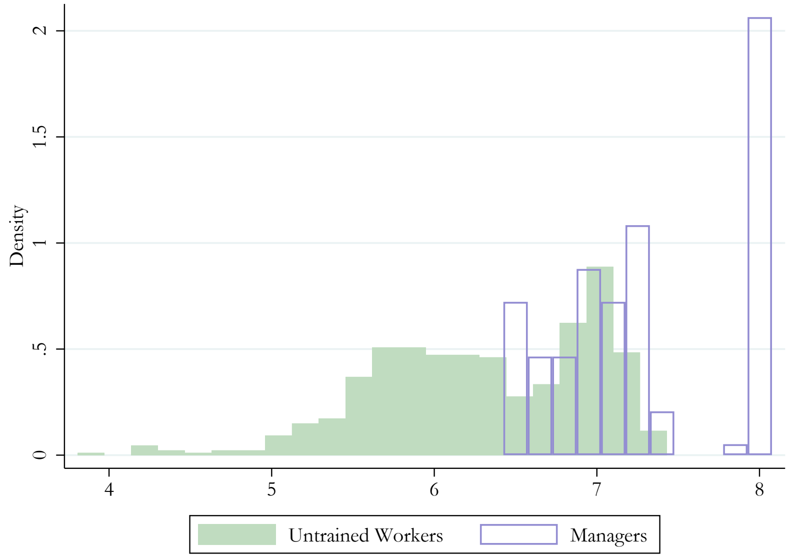
(a) Emails from Workers to Managers



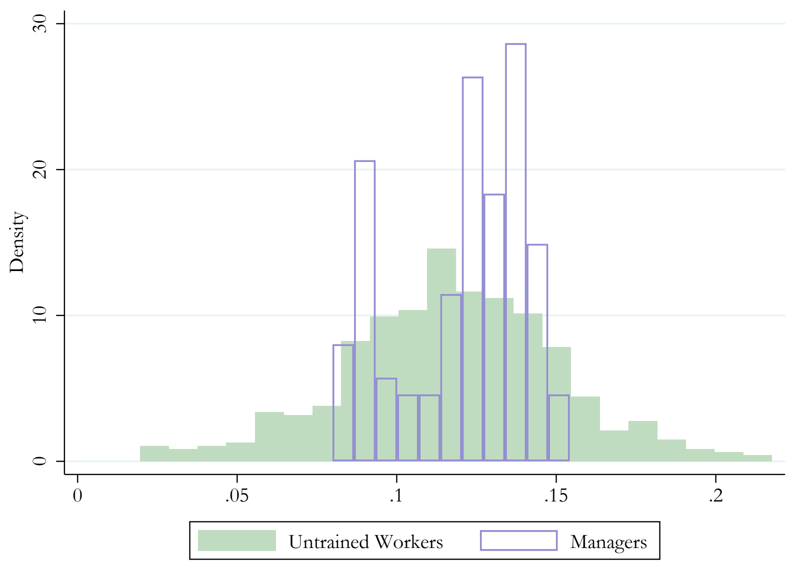
(b) Emails from Workers to Workers

Figure A1: 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 5 week gap between these periods.



(a) Log of Pre-Period Emails from Trained Workers to Managers and Untrained Workers



(b) Share of Pre-Period Emails from Trained Workers to Managers and Untrained Workers

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

Note: The top figure shows the distribution of log pre-period emails received from trained workers for untrained workers and managers. The bottom figure shows the distribution of the share of pre-period emails received from trained workers over all workers from the same division, plotted separately for untrained workers and managers.

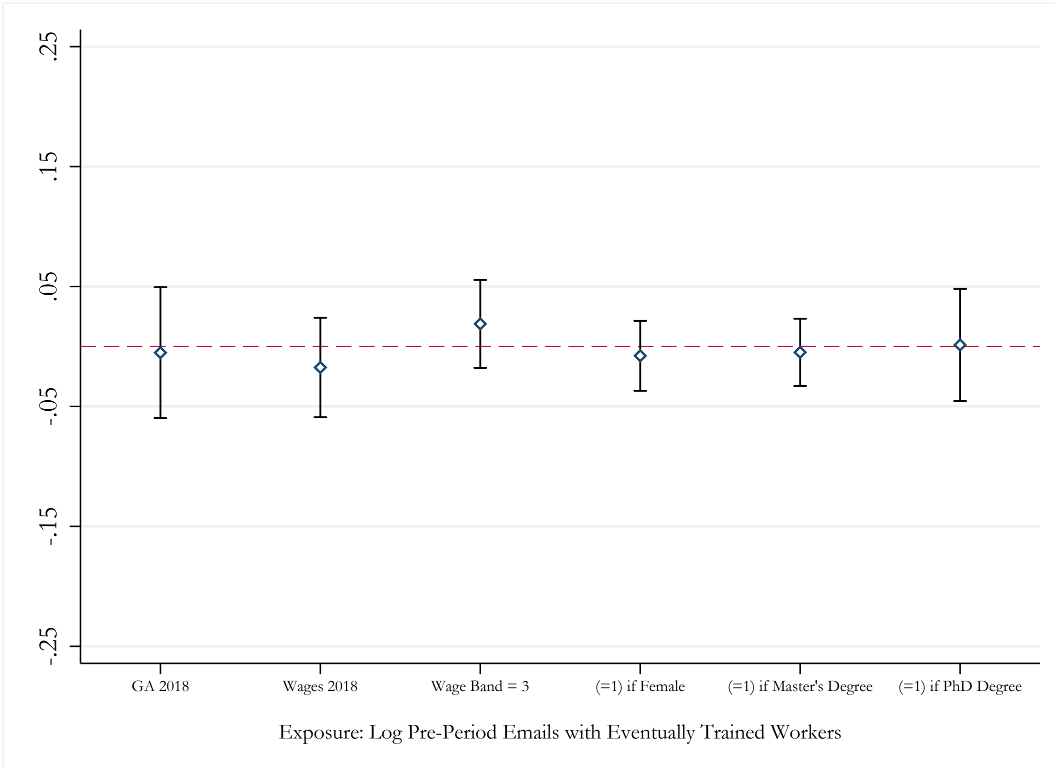


Figure A3: Predictors of Manager Exposure to Trained Workers

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 $log(EmailTrained_{i,Pre}) = \beta_1 GoalAchievement_{2018,i} + \beta_2 Wages_{2018,i} + \beta_3 WageBand3_i + \beta_4 Female_i + \beta_5 Master_i + \beta_6 PhD_i + Division_i + u_i$. The unit of observation is a manager. The joint test has F(6,18) = 0.23 (p-value = 0.968).

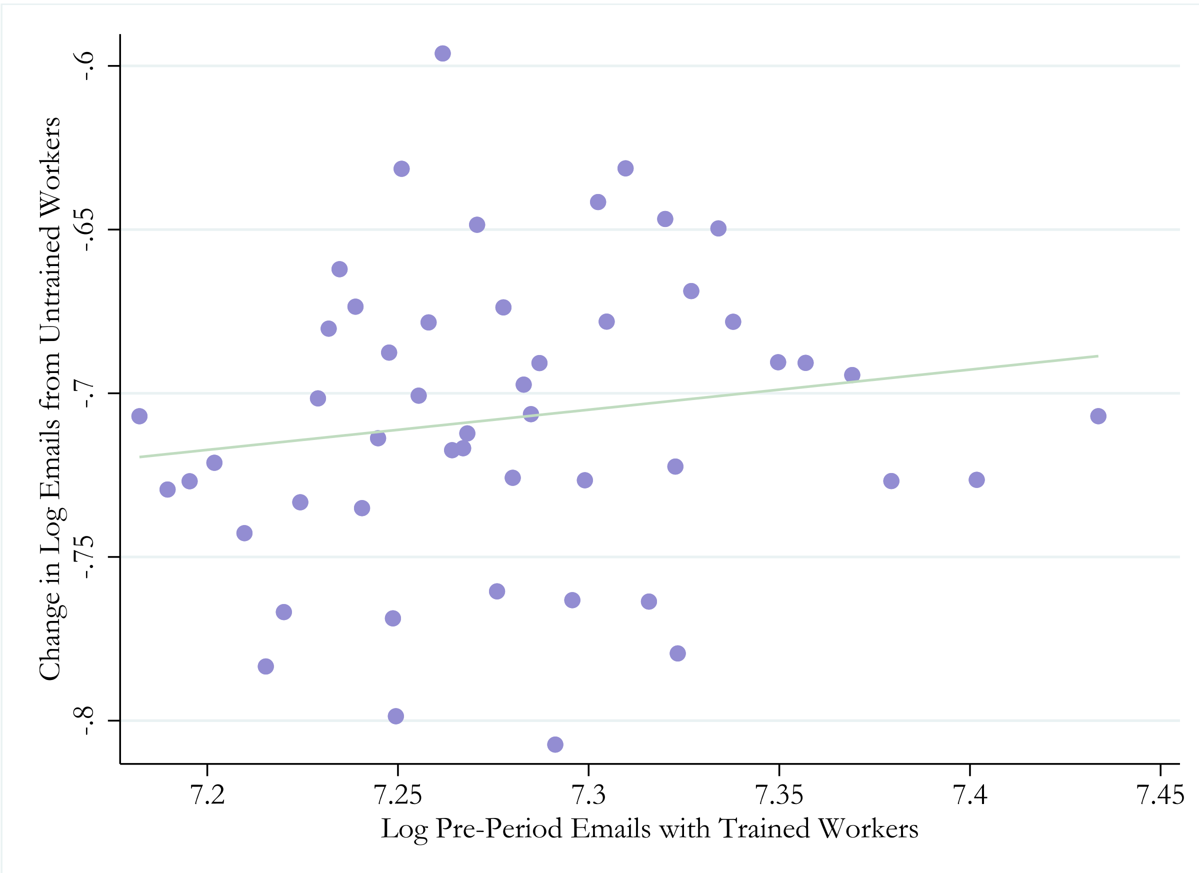


Figure A4: Tests for Manager SUTVA Violations Based on Changes in Emails with Untrained Workers

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 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 change in log emails between the pre- and post-periods and the x-axis is log pre-period emails with trained workers. The regression coefficient and standard error (N=129) is 0.12 (0.11) .

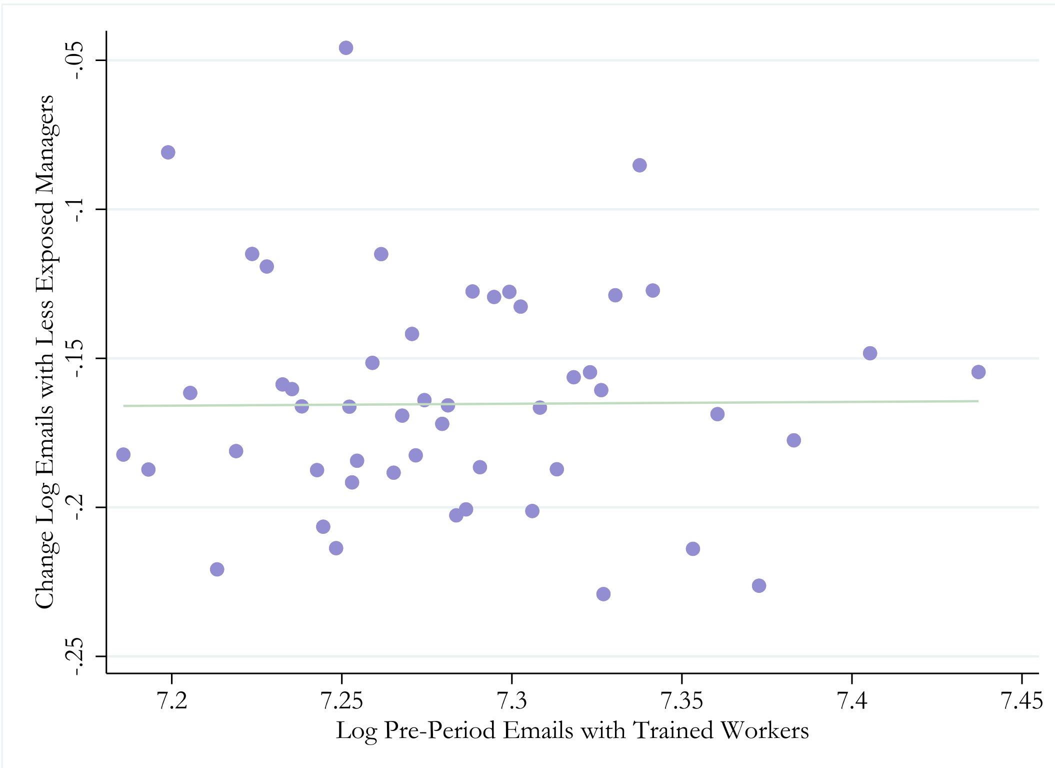


Figure A5: Tests for Manager SUTVA Violations Based on Changes in Emails with Other Managers

Note: This figure displays the relationship between the exposure to trained workers (based on the log number of pre-period emails with trained workers) and the change in log emails with managers who are less exposed to trained workers managers. 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 0.006 (0.086).

B 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___.[30]

3. Remember your work environment in 2018 and 2019. Consider all of the people who you interacted with via e-mail 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___.

[30] 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.