NBER WORKING PAPER SERIES

TRAINING, COMMUNICATIONS PATTERNS, AND SPILLOVERS INSIDE ORGANIZATIONS

Miguel Espinosa
Christopher T. Stanton

Working Paper 30224
http://www.nber.org/papers/w30224

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2022

We thank Oriana Bandiera, Morten Bennedsen, Jordi Blanes-i-vidal, Ronald Burt, Lorenzo Caliendo, Alexia Delfino, Florian Englmaier, Guido Friebel, 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, 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 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2022 by Miguel Espinosa and Christopher T. Stanton. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Training, Communications Patterns, and Spillovers Inside Organizations
Miguel Espinosa and Christopher T. Stanton
NBER Working Paper No. 30224
July 2022
JEL No. J24,L2,M5,M53

ABSTRACT

We study direct productivity changes and spillovers after a randomized training program for the frontline workers in a Colombian government agency. While trained workers improved their individual production, we also find substantial spillovers that affected managers' productivity. We use email data and a survey to explore the mechanisms behind these spillovers and find that managers' increased output arises from reductions in the need to help lower level employees. Accounting for spillovers to manager productivity changes the organization's implied return on investment from the training program, expanding the set of training investments that can be supported.

Miguel Espinosa
Ramon Trias Fargas 25
miguel.espinosa@unibocconi.it

Christopher T. Stanton
210 Rock Center
Harvard University
Harvard Business School
Boston, MA 02163
and NBER
christopher.t.stanton@gmail.com

1 Introduction

Labor market frictions, like monopsony or imperfect information, provide a rationale for employers to sponsor training even if it provides general skills (Acemoglu and Pischke, 1998, 1999). In environments where organizations capture rents from workers, efficiency likely requires employer subsidized training because workers are not the full residual claimants on their skills investments. 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. According to Training Industry Magazine, when organizations do attempt to quantify training performance, they tend to focus on individual-level outcomes.[1] As a result, some potential benefits are likely missed, namely spillovers to others. In this paper, we study the direct returns and the spillovers from an employer-sponsored training program. We provide estimates of the magnitudes of spillovers relative to direct productivity increases for trained workers, discuss the potential mechanisms driving spillovers, and consider how accounting for spillovers would change an organizations' willingness to invest in worker training relative to calculations that only consider benefits for the trained workers.

The setting for our study is a Colombian government investigative agency where 63 frontline workers (12% of those eligible) were randomly allocated to participate in a training program. The program occurred between August and December of 2018 and entailed 120 hours of classes covering computer skills, principles of goal setting and management, legal analysis, written communication, and specific topics related to each participant's own work. Besides the random assignment to training, the setting has a number of attractive features.

First, each worker has goals set and evaluated every week by an independent, separate unit of the organization that is responsible for oversight and performance evaluation. The organization is structured this way because the main function of the employees that we study entails sensitive work for the public interest, and the separation of oversight is designed to provide accountability. Although we only observe aggregate measures of goal

[1] See: https://trainingindustry.com/articles/measurement-and-analytics/how-to-identify-the-right-training-kpis-for-your-learning-and-development-programs-spon-eidesign/

achievement, rather than the details of individual goals, our contacts in the organization indicate that goals might range from case processing metrics to the completion of strategy documents or the implementation of process changes. Similar uses of goals and objectives, with outside measurement against them, is common in the public sector (Rasul and Rogger, 2018; Rasul et al., 2018). Importantly for our context, the organization indicated that goal setting and attainment measurement did not depend on training status (due in part to the goals and measurement being set by an outside party). These features allow 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, the organization shared metadata on the quantity of emails between all employees–including emails between frontline workers and between frontline workers and those higher in the hierarchy. We label higher level employees as managers, although higher-level employees do not have the same hiring/firing authority or dedicated teams that would be typical of managers in some other settings. The communications data spans a 13 week period several months prior to when the training program began and includes the same 13 weeks in the following year, 2019, several months after the program's conclusion. From the email data, we infer connections between co-workers with each other and between workers and managers prior to the randomization into training. We then trace out differences in untrained co-workers' and managers' exposure to trained workers, allowing us to assess if productivity evolves differentially for those most closely connected to trained workers compared to more distant connections in the pre-period.

We begin by documenting that the training program raised productivity, a finding that is not obvious given the literature on other forms of training (Card et al., 2018). Average goal achievement among trained frontline workers increased from 72% per week to 79% per week between the pre- and post-periods. The increase in goal achievement for trained workers was positive across the pre-period productivity distribution, with slightly larger increases for lower performers. Untrained frontline workers' goal achievement remained at 72%, and average goal achievement changes were approximately zero across the pre-period productivity distribution for untrained workers. The relative increase in goal achievement for trained workers does not appear to be driven by changes in labor supply or retention. Trained workers in the post-period are actually more likely to have days with no measured work activities, a proxy for absenteeism, and there was minimal turnover among either group during this period.

Putting these ingredients together, the estimated average treatment effect of the 120 hours program is a 10% productivity increase in the medium-run (4-6 months after training completion). There are two substantive reasons that we interpret the raw differences in goal achievement as the approximate average treatment effect. These are: 1) The organization indicated that everyone who was randomized into the program participated and attended at least 85% of the sessions (the benchmark for successful completion). 2) Estimation approaches that account for violations of the Stable Unit Treatment Value Assumption (SUTVA) yield qualitatively similar conclusions to the simple difference-in-differences estimator.

These estimates are inputs into calculating the direct returns to training. Under an assumption that labor demand is elastic, we use data on wages and pre-period goal achievement to get a money-metric for what the organization is paying (and willing to pay) per goal. This money metric allows us to convert goal achievement gains from the program into units that are comparable to program costs. We then compare benefits to costs under different scenarios for the short-run and long-run evolution of productivity that goes beyond our data (our data does not coincide with the program dates nor does it extend beyond 6 months after the end of the training period). We also include various assumptions about the opportunity cost of trained workers' classroom time, and we net out overhead and administrative costs of the program. After doing so, we find direct returns to the program are negative if they persist for only 6 months post-training (which is the period of our data), while the ROI to the organization from the direct benefit was 24% if gains persist for 1 year post-training.

Second, we find that the most important spillovers from the training program are to those workers in the management layers of the organization. In the raw data, managers' average goal achievement increased year-over-year from 71% to 73%. In our ex-ante exposure design, we calculate each manager's degree of connection to trained workers as a function either of the level or share of emails received from workers who eventually receive training. In our most conservative specification where exposure is defined as the log of total emails received from eventually trained workers, we estimate that spillovers from training are responsible for an approximately 1.5 percentage point increase in manager goal achievement, accounting for about 65% of the time series change in managers' aggregate goal achievement. In specifications where the exposure measure is the pre-period share of emails with eventually trained workers, we can explain the entirety (or more) of the 2.2 percentage point increase in manager productivity.

Third, the spillovers to managers are large enough to alter the evaluation of the training program. With even a 1 percentage point increase in manager goal achievement, the training program breaks even for the organization if gains persist for only 6 months, whereas at the 6 month horizon, the direct returns are negative. While a cursory comparison of the 1-2 percentage point manager goal achievement gain relative to the 7 percentage point increase for trained workers might suggest the manager spillovers are immaterial for the organization's choice to train, closer examination reveals this intuition is mistaken. There are two reasons that manager productivity gains were meaningful for the organization: i) The smaller per-capita percentage gains in output for managers are spread over more people (129 managers versus 63 trained workers). ii) By a revealed preference argument, it is likely that manager goal achievement is worth more to the organization than frontline workers' goal achievement because managers earn significantly more than lower-level workers. When we weight each workers' and managers' goal achievement gains by their compensation (a measure of the cost that the organization is willing to pay for each goal), we find that the total increase in compensation-weighted manager productivity was between 66% to 133% as large as the direct compensation-weighted productivity increase for workers.

Fourth, the channel for vertical spillovers to manager appears to favor a Garicano (2000) hierarchies mechanism, where managers and workers' are substitutes in production, compared to a model where managers and workers are complements. We arrive at this conclusion by examining the predictions of two different models. In the base model where workers in different layers of the hierarchy are complements, we would expect a positive covariance between workers' skills and emails received from more skilled workers. In alternative models where managers handle exceptional problems that workers can't, manager productivity is negatively related to email volume because as a worker's skill increases, she handles more problems, freeing up manager time. We find in OLS and IV regressions that the year-over-year change in manager productivity is negatively related to changes in the share of emails from eventually trained workers – as workers become more skilled, they appear to rely less on managers for help. If emails instead signaled productive connections across hierarchical layers, we would have expected to find that the shift away from emails with more productive trained workers would predict declining manager productivity. Survey evidence supports the hierarchies mechanism: responses indicate that emails across vertical layers in the hierarchy of this organization are often used to seek out or provide 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.

Fifth, we find that the net effect of spillovers to untrained frontline workers is approximately zero. We use a LASSO variable selection procedure to identify the pre-period connections that predict out-of-sample goal achievement. Using a battery of possible connection measures between trained and untrained workers, we find positive but insignificant measures of spillovers to untrained workers.

We have also probed whether our findings can be explained by different rationales. Our results do not appear to be driven by changes in monitoring, career concerns, or worker motivation. Trained and untrained workers report similar levels of monitoring before and after training and surveys show no differences in perceived career paths by training status. The most plausible alternative explanations, therefore, do not affect our results.

Our results have implications for understanding the economics of intra-firm spillovers, especially in the context of training programs. Due to data limitations, the approach in most of the literature that evaluates on-the-job training measures efficacy based on individual wage or performance gains (Bartel, 1995; Konings and Vanormelingen, 2015; Black and Lynch, 1996). When prior work has attempted to estimate spillovers, the focus has been on peers at the same level. Two 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 design, they examine spillovers to coworkers (but do not include managers) and find that they are significant, with larger spillovers found 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 positive peer spillovers from training in a call center. They find a 10 percentage point increase in the share of trained coworkers increases performance by 0.5%. There are two key differences in our context. First, we examine spillovers across levels of the firm hierarchy, and second our approach to detect spillovers is based on communication patterns (which may also respond to training). We are aware of few other papers that estimate the spillovers from training inside the firm, and none that do so across the vertical hierarchy of an organization.[2]

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 top-down effort targeting and the importance of social connections across levels of the hierarchy (Bandiera et al., 2007, 2009). Much less is known about how spillovers from more skilled lower level workers flow up the hierarchy to affeect managers. Our results suggest that the individual returns to training may fail to account for a significant fraction of the surplus generated from offering training programs because more productive workers allow managers to become more productive. While we

[2]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 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.

caveat that both the direct returns and spillovers may be more ephemeral in other types of organizations, 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 (Acemoglu, 2005; Besley and Persson, 2010; Dal Bó et al., 2013; Rasul and Rogger, 2018; Rasul et al., 2018; Bandiera et al., 2021).

The rest of the paper is organized as follows. Section 2 describes the institutional setting. While Section 3 provides a framework for our empirical methodology, Section 4 contains the main analysis of the direct effects and spillovers of the training program. Section 5 addresses potential confounding explanations and Section 6 concludes.

2 Study Setting

The context for our study is a public sector organization in Colombia, a middle-income country with a growing economy at the time of our data collection. Our agreement with the organization prevents us from disclosing further details beyond the fact 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 of the organization.

The employees that we study are in stable, white collar occupations where turnover is limited.[3] At entry into the organization, workers are assigned to a division and a 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 only have bachelors or secondary school attainment and are "frontline" workers in our terminology. Workers from wage band 3 to

[3]During the period of our data collection, the organization had minimal hiring and negligible turnover. In fact, we only observe two workers 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 organizations.

5 hold bachelors, masters or PhD degrees and are "managers." Wages are determined by wage band and organization-specific experience.

Workers perform slightly different functions depending on their division. Each division has the following responsibilities: 1. The "Execution Division" (36.9% of employees and code named to preserve anonymity for the organization) 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. Finance (13.7%) manages the budget and treasury. 4. Human Talent (14.9%) handles the creation and implementation of internal policies, inductions, permissions, fulfillment of requirements, payroll supervision, and other Human Resources tasks. 5. Planning (14.9%) advises the top management unit on the creation of policies, procedures, and resource allocations to accomplish the organization's objectives.

The organization also has other employees in non-core functions, the most important of which is an oversight group that is firewalled and independent, serving 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 oversight group that is 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 does not take into account workers' training status, and there is no ratcheting of expectations either in response to past performance or training attainment.

Despite our labeling of higher level workers as managers, they do not have direct authority over workers. Managers' main role is planning and setting priorities rather than executing frontline work. When needed, they also help frontline workers. Surveys and interviews with the organization indicate that managers may provide authorizations if higher level input into a decision is required; on occasion managers may also allocate tasks to different workers.

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 (our productivity measure), absenteeism, and demographic and personnel information, including gender, education, monthly wage, wage band, and division.

We supplement these administrative records with information on email communications between the 655 employees. We have data on daily bilateral email counts between every pair of workers over the 13 weeks in 2018 and 13 weeks in 2019.[4] 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 5.2) that confirm that emails proxy for the totality of communications between individuals.[5]

2.1 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 they could train only 63 employees. 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. Table 1 shows that randomization into training is balanced on observables. It also provides additional descriptive statistics about the sample.

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

[4] 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.

[5] 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.

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.[6]

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.

2.2 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 we gained qualitative insights through interviews about the typical goal setting process. For example, a weekly goal for frontline workers in the Execution Division would typically entail progress on one or multiple cases/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.[7]

[6]Trained workers may have become better at communication through emails, potentially decreasing the number of emails sent. This interpretation is consistent with the decrease in emails to managers. However, trained workers increased their emails with peers, which is inconsistent with communications becoming more concise.

[7]Examples for other divisions are similar. Workers in Administration handle procurement, inventory management, and policies and procedures. Managers in the division are in charge of the design of safety and security procedures and employees' compliance. The 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, properties, 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 analysis affecting the organization's human capital planning. Workers' tasks in the Finance division tend to focus on conducting transactions and adhering to budgets, whereas managers are responsible for budgeting and monitoring payments and cash inflows in

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.

2.3 Descriptive Statistics

2.3.1 Data on Workers, Wages, and Goals

The vertical division of the firm can roughly be described as containing 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. Table 1 shows descriptive statistics for the sample. Frontline workers are more likely to be female than managers (48% in wage band 1, 29% in wage band 2, versus 18% of managers). Managers are more educated, the 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. Comparing the ratio of managers to frontline workers across divisions, there are many more workers per manager in the Execution division than in others. As a result, small changes in worker productivity may get magnified for those managers when the ratio of workers to managers is highest. We will return to this point later when examining spillovers from worker training across the vertical hierarchy.

The next few rows deal with wages and wage bands. The row labeled Wage Band

the accounting system while ensuring that the legal requirements related to those 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.

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 average pre-period wage of Wage Band 1 workers. On average, managers earned 2.16 times more than wage band 1 workers while workers in wage band 2 earned 20% more than those in wage band 1 in the pre-period. 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. Our data on monthly compensation shows no abnormal wage increase for trained workers.[8] As a result, 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 do not appear to result from spillover gains, as their wage changes are orthogonal to the year-over-year change 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 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.

2.3.2 Email Data

We use email data to infer connections between coworkers, between workers and managers, and how connections and communications patterns change in the period after training. Note that our email data do not distinguish between emails sent to one person or to multiple recipients. Our analysis is thus based on quantities of emails between senders and individual receivers, but we cannot distinguish whether email threads are to teams or multiple recipients. Table 2 provides details about the email data.

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

Our strategy for identifying spillovers utilizes the idea that some managers or untrained workers are more connected to eventually trained workers than others. Table 2 shows these connections in the pre-period. For example, untrained workers have an average of 674 emails with eventually trained workers from their own division in the pre-period, with a standard deviation of 385. The average share is about 12%, with a standard deviation of 3.4%. 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. For managers, these numbers imply that the average share of emails with eventually trained workers is 12.1% with a standard deviation of 2%. The difference between levels and shares reflects that some managers are generally more active with emails than others. Because of the bureaucratic nature of this organization, managers do not receive emails from frontline workers in other divisions.

Our identification strategy also assumes that in the absence of the training program, communications patterns would have remained stable. To provide evidence that this assumption is reasonable, we utilize pre-period data to show that connections in the email data are highly persistent.[9] Figure 1 plots average dyad-level shares of emails sent in the "Late Pre-Period" for each decile of 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. When managers are the recipients, in Panel A, emails are shown to be highly persistent. A similar pattern of persistence is evident in Panel B, which examines emails to other workers both within and across division. These figures suggest that communications patterns are relatively stable in this organization, at least after a several week lag, suggesting that our exposure design is likely reasonable.

Returning to Table 2, one striking fact is the change in email shares with trained frontline workers across columns. For untrained workers, their email share increases from about 12% in the pre-period to 18% in the post-period. It falls from 12% to 6.7% for managers. This significant reduction in emails for managers will be useful for trying to

[9] It is difficult to test stability using post-period data because of endogenous changes in communications that resulted from training (and we later show evidence that training changed communications as part of the mechanism for our findings).

distinguish different mechanisms.

3 Measuring Direct and Spillover Returns to Training

3.1 Direct Returns and Spillovers to Frontline Coworkers

Because of experimental variation, standard intuition suggests that estimation of the direct benefits of training simply entails a comparison of goal achievement for trained workers versus untrained workers in the post-period. This estimate and the corresponding standard error come directly from Table 1. However, there are a few additional reasons to consider regression analysis. First, a difference-in-differences framework allows us to absorb some pre-period productivity heterogeneity with worker fixed effects, increasing statistical power. We are also able to test whether training has differential effects for workers who are likely to have higher baseline levels of human capital (i.e. those workers in Wage Band 2 who have greater average levels of education or work experience). Our simplest estimator is then a two-way fixed effects model:

$$\log(y_{it}) = \beta_i + \beta_t + \delta_1 Trained \times Post + \delta_2 Trained \times Post \times X_i + \varepsilon_{it} \tag{1}$$

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 very coarse characteristics. Individual fixed effects are captured through β_i and time fixed effects through β_t.

Viewed from a potential outcomes perspective, equation (1) stipulates that counterfactual expected log productivity in the post-period for workers who are not trained equals $\beta_i + \beta_t$. This imposes an assumption that there are no spillovers to untreated workers, known as the Stable Unit Treatment Value Assumption (SUTVA). 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 $g(Connections, \gamma)$ be a function that captures the impact of connections between trained coworkers and the untrained with parameters γ. Then

$$\begin{aligned} \log(y_{it}) = \beta_i + \beta_t + \delta_1 Trained \times Post + \delta_2 Trained \times Post \times X_i \\ +(1 - Trained) \times Post \times g(Connections, \gamma) + \varepsilon_{it} \end{aligned} \tag{2}$$

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.

3.2 Vertical Spillovers

We consider two models of potential vertical spillovers that help to clarify the sources of productivity changes for managers.

Model 1 The first model captures complements in production – i.e. where productivity of a managers increases if workers on the same projects or tasks become more skilled. In the complements case, workers and managers' goal achievement is positively correlated because they work together in ways that may be interdependent. Because in our setting there are no defined teams, we infer connections between managers and workers from emails. Using the email data, the simplest specification to capture production complements is a linear in means model of interaction effects. This can be specified as

$$\log(y_{it}) = \alpha_i + \beta_1 \frac{\sum_j C_{ij} s_{jt}}{\sum_j C_{ij}} + \varepsilon_{it} \tag{3}$$

where C_{ij} is a metric capturing the strength or degree of connections between focal manager i and eligible workers j. In the numerator, connection strength is multiplied by the baseline skill of worker j at time t, s_{jt}. The denominator normalizes by the strength of all connections. In what follows we suppress the time subscript on worker skills except when we specify changes between the pre- and post-periods for trained workers.

There are two testable comparative statics for the production complements model. First, consider an increase in skill for the kth worker. It is easy to see from (3) that

$$\frac{\partial \log(y_{it})}{\partial s_k} = \beta_1 \frac{C_{ik}}{\sum_j C_{ij}}$$

which is positive if $\beta_1 > 0$ and is increasing in the relative strength of connections between i and k. Second, consider a change in connections between manager i and worker k, yielding

$$\frac{\partial \log(y_{it})}{\partial C_{ik}} = \beta_1 \frac{s_k(\sum_j C_{ij}) - (\sum_j C_{ij} s_j)}{(\sum_j C_{ij})^2}$$

The sign of this comparative static depends on a comparison between s_k and the baseline skill of all other workers. When s_k is above the mean of other workers, an increase in connection strength C_{ik} between manager i and worker k positively impacts manager output. When s_k is below the mean of other workers, increasing the connection strength draws the manager away from higher performers, reducing output.

These comparative statics yield two testable hypotheses for the complementary interaction model.

C1 Managers who are better connected to workers that become more skilled or productive will have a greater increase in productivity compared to those who are less connected to workers that become more skilled.

C2 Managers who experience an increase (decrease) in relative connections with highly skilled workers should have increasing (decreasing) productivity. That is, manager productivity is positively related to the relative strength of simultaneous connections with skilled workers.

Testing these comparative statics is challenging, as identifying equation (3) in any setting is inherently difficult due to simultaneous unobservables and the reflection problem. In this setting there is an additional challenge because of a) potentially endogenous connections that change C_{ij} in response to training and b) the potential that the firm reorganizes or reoptimizes in a way that muddles the historical relationship between connections and output.

We instead focus on testing comparative static C1 by estimating versions of equation (3) that exploit randomization into training as shifters of s_j, while holding fixed connections in C_{ij} based on pre-training communication patterns. This assumption is based on the notion that there is some residual persistence in connections that remains after the training program. The reduced form for this equation is

$$\log(y_{it}) = \alpha_i + \beta_t + \beta_1 \frac{\sum_j C_{ij,Pre} \times Trained_j \times Post}{\sum_j C_{ij,Pre}} + \varepsilon_{it} \tag{4}$$

This test says nothing about concurrent connections after training, as it is based on historical communications. As a result, this reduced form does not allow us to test comparative static C2. Comparative static C2 can be tested against an alternative model that predicts a different sign of the relationship between connection strength, training, and manager output. Note, however, that comparative static C1 is consistent with the predictions in Model 2. Thus the reduced form estimate cannot discriminate between models, but it can reject both models. Ultimately we will find support for comparative static C1, yielding an estimate of spillovers to managers.

Model 2 The second model we consider is a hierarchical model with specialization that follows Garicano (2000). To illustrate, assume that managers allocate 1 unit of time between their own production tasks and helping subordinates. Based on conversations with the organization, managers' goal achievement evaluation is not linked to the contemporaneous goals of their subordinates, but there is some expectation that managers need to allocate time to subordinate requests because part of manager evaluation includes how they plan for and establish projects that subordinates staff. As a result, managers spend some amount of time proportional to $C(\lambda, s)$ helping, where λ is the arrival rate of tasks to be done on projects staffed by subordinates, s is the skill vector of subordinates, and task difficulty d is distributed according to $F(d)$. As a normalization, assume a worker can do a task if $d < s$ and otherwise must request help. More skilled subordinates can do a larger share of tasks, so they are less likely to request help. The production function for manager output can be written as

$$\log(y_{it}) = \alpha_i + \beta_1 C(\lambda, s) + \varepsilon_{it}. \tag{5}$$

Differentiating with respect to a change in worker skill yields

$$\frac{\partial \log(y_{it})}{\partial s_k} = \beta_1 \frac{\partial C(\lambda, s)}{\partial s_k}$$

which is positive if $\beta_1 < 0$ and $\frac{\partial C(\lambda,s)}{\partial s_k} < 0$, which would be expected according to the Garicano framework. In this model, managers become more productive when their connected subordinates increase their skills, but the increase in manager productivity is negatively correlated with communications or connections strength, as these communications signal help requests that take manager time away from other tasks.

The distinguishing feature between **model 1** and **model 2** is whether communications strength changes positively or negatively after a shock to worker skills and how this change in communications affects manager output.

4 Main Results

We first present estimates of the direct changes in productivity due to training. These estimates also account for spillovers to coworkers. We then present results on vertical spillovers from trained workers to managers. We conclude this section with an assessment of the returns to the training program under different scenarios for direct returns and spillovers. Core to this exercise is a metric that translates gains in goal achievement to dollar values of benefits that can be compared to cost. By a revealed preference argument, when the firm's labor demand is elastic, we show how the implied price per goal in the pre-period can translate gains in goal achievement to a monetary value for benefits.

4.1 The Direct Productivity Effects of Training

Figure 2 shows that trained workers increased their goal achievement. Each sub-figure contains a binned scatterplot of yearly average goal achievement in the post-period relative to the pre-period. The figures plot the relationship separately for trained and untrained frontline workers. Two key points from Panel A 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, and ii) there is a positive vertical shift upward for trained workers relative to the untrained. The shift for trained workers is apparent across the support of the pre-period productivity distribution (averaging about 7 percentage points), while gains in percentage terms or percentage points are greater for lower performers in the pre-period (the slope of the line is slightly smaller for trained compared to untrained workers).

It is also apparent from Panel A of Figure 2 that there are several distinct clusters of goal achievement scores. Panel B explores the source of this clustering by netting out division fixed effects, which marginally increases variability along the horizontal axis. Distinct clusters remain after netting out division fixed effects, suggesting that evaluators likely round the sub-components of the goal achievement measures, leading to some bunching in the distribution.

Table 3 contains difference-in-differences estimates confirming the increase in goal achievement when including worker and time fixed effects. 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 2 and 3 add interactions to test for heterogeneity by wage band. In the absence of division fixed effects (Column 2), 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 fixed effects in Column 3, the coefficient of -0.035 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.

The remaining columns present estimates of equation 2 that account for potential spillovers that may violate the SUTVA. In these columns, the connections we include to eventually trained workers (denoted T) in the pre-period are selected via LASSO from a variety of different possible measures of connections. Those measures that survive the LASSO are included in the table. The point estimates remain broadly similar for trained workers. The bottom rows also show our estimates of spillovers to coworkers, which are positive in both columns but are insignificant after we account for division fixed effects. The specification with division fixed effects is our preferred specification because the LASSO selected regressors distinguish between connections with trained wage band 1 and trained wage band 2 workers, but one division had no trained wage band 2 workers, which is captured by the division fixed effects.

We will later return to mechanisms, but for now we note that the increased goal achievement of trained workers does not appear to be driven by increases in motivation or work hours. As a proxy for work hours, we look at absenteeism as measured by days without email activity. At the time of our sample, all email was accessed in the office only, so engagement with email is a proxy for attendance. Table 4 shows that, if anything, absenteeism increased for trained workers despite their increase in goal achievement. A potential explanation for this finding is that the organization compensated trained workers not through additional formal compensation but through relaxed attendance standards or more flexible work hours.

4.2 Spillovers to Managers

Table 5 displays two different reduced form measures of manager exposure to trained workers. In Panel A, the measure is the log number of pre-period emails between a manager and eventually trained workers. The advantage of using log emails is that it closely aligns with the model of manager time use and busyness from Garicano (2000). The disadvantage is that this measure may capture that some managers are simply more central for all communications with workers, which would include trained and untrained workers. Panel B gets around this issue by focusing on the pre-period share of emails with eventually trained workers. This measure is also the one that is directly motivated by the linear-in-means interactions effect model in equation (4).

In both Panels A and B, and across all columns, managers who have stronger pre-period communications connections to eventually trained workers have differentially greater productivity gains in the post-period. In Panel A, average implied effects for the level of goal achievement range from a 1.46 to a 2.11 percentage point increase. This calculation is taken as the predicted effect of the regressors and includes the post-period indicator (from Column 1). The large negative coefficient on the post-period indicator suggests that our model is good only locally (as all managers are somewhat connected to trained workers) and likely would not fit the data well for a manager that was completely unconnected to trained workers. An alternative statistic that does not require extrapolation beyond the range of the data is the interquartile range of the estimated effects. The interquartile range (IQR) of the change in goal achievement due to connections in Panel A is about 4 percentage points.

Columns 3 and 4 introduce managers' sent emails as another connection measure. The coefficients are smaller and become insignificant with the inclusion of division fixed effects in Column 4. Emails received, rather than those sent, appear to best explain changes in manager goal achievement through exposure to trained workers. While this pattern isn't obvious if considering a simple model of connections and complements, in hierarchies models it is inbound requests that determine workload at higher levels of a hierarchy, as problems move upward. Indeed, as we discuss later in the section on mechanisms, a survey of workers indicates that many emails are about seeking out help.

The qualitative patterns are similar in Panel B. These estimates do not appear to have the problem of extrapolating beyond local variation. In Panel B, all of the estimates of the average spillover effect imply a goal achievement gain to managers exceeding 2.4

percentage points. Again the gains load on the share of emails received, rather than those sent. While we estimate positive goal achievement spillovers to managers, these reduced form results could be consistent with several different mechanisms.

One problematic mechanism, that would overstate the gains from the program, would occur if managers actually do not become more productive but instead are perceived to achieve more because their connected workers do. We find very little evidence that this explanation is plausible after we test for co-movement between connected workers' goal achievement and manager goal achievement in the pre-period. Contemporaneous linkage between connected workers' goal achievement and manager goal achievement is minimal, a finding which is inconsistent with this mechanistic view of spillovers to managers. Table 6 displays a variety of estimates from regressions of log manager goal achievement on email-weighted measures of log worker goal achievement. Some specifications also control for workers' log goal achievement outside of the focal week, which isolates transitory deviations from permanent goal achievement. All estimates are small and insignificant, suggesting managers and workers' goals are not mechanically linked. The lack of contemporaneous movement in goal achievement also suggests that the reflection problem (Manski, 1993), which may cause a spurious finding of positive complementarity or peer effects, is not likely to drive our results. It is possible, however, that some complementarity is present but is masked by help requests or other forms of communication that make the underlying relationship difficult to detect in the absence of data on email threads or topics.

4.3 ROI: Benefits Relative to Costs for the Organization

What was the net effect of the program to the organization? To understand whether or not the training program produced positive net returns, we calculate total benefits and costs. Although we do not observe the value of each goal, the fact that the organization was willing to pay workers' salaries allows us to recover an implicit price-per-goal prior to training under the maintained assumption that labor demand is elastic. We use this price to calculate an approximate dollar value to the organization from the increased goal achievement of workers and managers.[10]

For each trained worker or each employee impacted by spillovers, we calculate the change in the monetary value of productivity to the organization as:

$$(GA_{Post} - GA_{Pre}) * \frac{W_{Pre}}{GA_{Pre}} \tag{6}$$

where GA_t is the average goal achievement in year t and W_{Pre} is the total annualized 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. We then sum over all affected workers and net out the fixed and administrative costs of the program. Table 7 presents calculations of program return on investment under a variety of scenarios that alter the assumptions about the persistence of training + spillover gains, the size of the spillovers, and the opportunity cost of the program.

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. In the first scenario, we assume that the program gains last through 6 months post-training and then depreciate completely. We also do not include the opportunity cost of trained workers time in this calculation, so only the fixed administrative costs are weighed against benefits. The ROI is -37% when considering only the direct returns at a 6 month horizon. Adding just a 1 percentage point gain in goal achievement for managers turns the ROI positive at this short horizon. In the second scenario, we assume that the opportunity cost of the program is that trained workers are removed from day-to-day work for the 120 hours when they are in classes. In this case the ROI from direct returns alone are negative even if the gains persist for 18 months. However, again with just a 1 percentage point spillover to managers, the ROI is positive 22% if total gains persist for one year. The remaining rows of the table work through various additional scenarios, including increasing the magnitude of spillovers to managers to 2.2 percentage points (the full time series increase) from the conservative 1 percentage

[10]In a firm or organization with rent-sharing between workers and firms, our approach would likely yield a lower bound on benefits.

point increase assumed earlier.

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.

5 Discussion of Mechanisms

5.1 Communication Patterns

This section begins by exploring changes in communication patterns, helping to provide context for our findings while enabling us to examine mechanism. Figure 3 shows changes in emails between the pre- and post-periods according to sender and recipient type. For each sender, 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 wage band 1 and 2 workers drop dramatically for both trained and untrained workers, but the reduction is larger for trained workers. On the other hand, emails originating from managers and sent to frontline workers are little changed. There are also some differences by wage band. Untrained wage band 1 workers dramatically reduce their emails to managers – but there is a large increase (with a value over 1.0, indicating a doubling of emails) between untrained wage band 1 workers and trained wage band 2 workers. In contrast to untrained wage band 1 workers, untrained wage band 2 workers have much smaller changes in emails with 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.

These striking patterns suggest that the organization re-balanced responsibilities for wage band 2 workers after training, having them take on a helping role (as will be demonstrated in the survey data) for less senior or less educated untrained wage band 1 workers.

This response seems consistent with adding an informal additional layer of management, a la Caliendo et al. (2015), that was made possible by the increase in skills for workers in wage band 2.

As a result of these large changes in communications patterns, however, a more direct test of the Garicano hierarchies model is difficult. This is because the direct test relies on the total emails received by a manager from lower-level workers, but emails drop for nearly all managers because of the diversion of emails to wage band 2 trained workers. The time series decline in emails to managers is sufficiently large that it swamps the cross-sectional first stage variation in pre-period exposure in specifications with division fixed effects.

A less direct test is possible, however, and supports the hierarchies model. The intuition for that test is that as workers gain skills, they should stop asking managers to help on tasks that they can handle themselves. As a result, the hierarchies model predicts a negative relationship between changes in manager productivity and changes in the share of emails from eventually trained workers. Table 8 presents this test using annual changes in manager productivity. Columns 1 and 2 display regressions of year-over-year changes in manager log goal achievement on changes in the email share from trained workers. The coefficient is sensitive to division fixed effects, but is negative in Column 2 when division fixed effects are included. Including division fixed effects is our preferred specification for OLS regressions, as the post-period email share is endogenous and the source of endogeneity is likely rebalancing of workload at the division level.

To deal with endogeneity directly, Columns 3 and 4 report IV regressions without and with division fixed effects. The instrument for the change in the share of emails with eventually trained workers is the pre-period share of emails with eventually trained workers. The IV coefficients range from -0.62 to -0.54, indicating that managers who had the largest declines in the share of emails with eventually trained workers had the largest increases in goal achievement. The average change in the share of emails with eventually trained workers is -0.05, so the -0.54 coefficient in Column 4 suggests that this channel is responsible for an approximate 3 percent (2 percentage point) increase in aggregate goal achievement for managers. The final columns present the first stage regressions of the change in email share on the pre-period share of emails with eventually trained workers. The first stage effective F-statistics are 23 and 16, implying a maximal bias of 10 and 20 percent, respectively (Olea and Pflueger, 2013).

5.2 Survey Evidence

We also conducted a survey in August of 2020 to improve our understanding of mechanisms. The organization distributed the survey to 63 of the workers trained in 2018 and to 105 untrained workers that were present in 2018.[11]

One of the main concerns with analyzing interactions through email communication is that workers have alternative communication modes that may substitute for emails. Alternatively, email may be complementary to other forms of communication, like face-to-face interaction or phone calls.[12] 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 electronic 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 to contact superiors is to ask for help, with the other responses split evenly between asking for authorization and reporting on progress on tasks. This same figure shows that 85% of surveyed workers think that the main reason a worker would contact those from a lower wage band would be to provide help. 10% think that contact with lower wage band workers is driven by the desire to allocate tasks. Only 5% think that the main reason to

[11] 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 A contains the English version of the survey.

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

contact workers below is related to either monitoring or to organize social events.

It is important to note that the respondents were not aware of the research findings around communication patterns, suggesting these results are independent validation of the interpretation that email patterns show that trained wage band 2 workers became a more important source of help for workers in wage band 1. However, the survey did tell respondents that workers from wage band 2 increased electronic communications with wage band 1 workers, as Figure 4 shows. The survey then asked them to provide what they thought was the main reason to explain such a change. Trained workers reported that there are only two reasons: to provide help (64%) or to respond to requests from wage band 1 workers that ask for help (36%). For untrained workers, these two reasons together represent 85% of their responses. A further 14% of untrained workers thought that the main reason to explain the increase communications from wage band 2 to wage band 1 was either to increase supervision or to ask for help more frequently.

5.2.1 Impressions of Changes Over Time

In the survey, we asked what the main changes (of different characteristics of their workplace) were in 2019 relative to 2018. We asked both trained and untrained workers, so differences in responses across these two groups provide some evidence of the main effect of the training program.

Table 9 shows that trained workers reported much greater improvements in their general skills and knowledge relative to untrained workers, with the exception of goal understanding. That is, trained workers report relative improvements in skills and knowledge, division-specific knowledge and problem recognition, and the ability to sort problems to different divisions. The table transmits a simple message, trainees believe the training program improved their skills.

5.2.2 Alternative Explanations

We also asked survey questions to understand potential alternative explanations. One dimension was changes in monitoring. For example, manager productivity may rise, while emails fall, if trained workers need less supervision, empowering trained workers to take new initiative (see for instance, Kirkman and Rosen (1999) and Mathieu et al. (2006)). Under this explanation, the primary reason for an increase in goal achievement was not because of skill and knowledge increases or the spillovers from trained workers, but rather because the monitoring effort of managers changed. Results from the survey are at odds with explanations around reduced supervision. 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. Another difficulty with the monitoring explanation is that managers do not decrease their outbound communications to frontline workers (see Figure 3).

Another potential effect of the training program is to change the incentives of the trained workers and make them more aware of promotion possibilities inside the organization. Goal achievement might be necessary to enhance the promotion likelihood (although in our data, we see no movement between wage bands). We asked directly whether survey respondents thought that their promotion possibilities increased from 2018 to 2019. Table 9 shows that 9% of workers from both groups, trained and untrained, think that there were more promotion possibilities in the post period. The fact that the percentage is the same across both trained and untrained workers leads us to conclude that the training program did not change perceptions about potential career trajectories. For this organization, promotion from within is rare, making career concerns unlikely to explain our results.

We also asked about changing task composition as a potential explanation for some results (like changes in emails to peers). Table 9 suggests tasks did not change. 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. This similarity suggests there was no differential task assignment of more team-oriented tasks to trained workers.

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 did. 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. We also note that measures of absenteeism actually increase for trained workers.

5.2.3 Did Changes in Communication Patterns Arise Organically or Were They Encouraged?

The survey also provides context around why trained workers increased communications with peers and 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 the fraction of trained and untrained workers that say that the organization told them to increase communication with peers and decrease it with managers is not statistically different one from each other. As a consequence, the large change in communication patterns from trained workers to other untrained workers in the same layer appears to arise from workers' own initiative rather than organizational mandate.

6 Discussion

6.1 Alternative Uses of Funds Spent on Training

Under a simple illustration that assumes the funds spent on the training program were instead used to hire additional managers, we attempt to calculate under what conditions the training investment would have been preferred to direct changes in the firm hierarchy. If the channel of manager gains is a reduction in demands on their time (busyness), then a simple proportional rule for how one additional manager increases incumbent productivity suggests that the gains will be approximately equal to $\frac{1}{N}$. The funds spent on training could have been used to hire about 1 manager for 12 months. In this case, 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.

6.2 Implications for Other Literature

At least since the Second World War, with the *Training within Industry* program, scholars have focused on studying the effect of training programs and the influence that employees on the top of the hierarchy can have on those on the bottom.[13] 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 wage bands can impact employees at the top of the hierarchy.

A further area for future work would be to consider how to target who gets training and how many workers should optimally be trained. For example, the literature on social network analysis provides tools to consider who might generate the greatest spillovers between coworkers (Bonacich, 1972; Freeman, 1978). Similarly, the economic sociology literature suggests the benefits might be greatest from targeting network brokers (Burt, 1992; Burt and Soda, 2017).[14] This work would help assess how skill changes reverberate

[13] The *Training within Industry* program was a service initiated in WWII that aimed to focus the training programs on those who in turn train other people -supervisors and experienced workers- (Dinero, 2005). There is extensive research on how managers have an effect on lower level employees (Lazear et al., 2015; Bloom et al., 2015, 2020).

[14] Another strand of literature suggests that returns to workplace programs are heterogeneous, so get-

either through professional or social networks, as the latter have been shown to substantially affect firms' internal operations (Bandiera et al., 2010). Extensions may also seek to capture how spillovers leak across organizational boundaries and how training programs that focus on firm rather than division-specific knowledge have an impact in the organization. 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 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 programs. To understand the mechanism behind spillovers, we examine changes in email communications and survey of employees. Both sources are suggestive that spillovers to managers arise reduced needs to assist lower level workers with their own tasks.

These results indicate the importance of considering production hierarchies and or-

ting the targeting rules right may depend on understanding personalized returns as well as spillovers (Sandvik et al., 2021).

ganizational 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.

References

Acemoglu, D. (2005). Politics and Economics in Weak and Strong States. *Journal of Monetary Economics*, 52(7):1199–1226.

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*, Forthcoming.

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.

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

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.

Bonacich, P. (1972). Factoring and Weighting Approaches to Status Scores and Clique Identification. *Journal of Mathematical Sociology*, 2(1):113–120.

Burt, R. S. (1992). *Structural Holes*. Harvard University Press.

Burt, R. S. and Soda, G. (2017). Social Origins of Great Strategies. *Strategy Science*, 2(4):226–233.

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

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.

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.

Dinero, D. (2005). *Training within Industry: The Foundation of Lean*. CRC Press.

Freeman, L. C. (1978). Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1(3):215–239.

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

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.

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.

Kirkman, B. and Rosen, B. (1999). Beyond Self-management: Antecedents and Consequences of Team Empowerment. *Academy of Management Journal*, 42(1):58–74.

Konings, J. and Vanormelingen, S. (2015). The Impact of Training on Productivity and Wages: Firm-level Evidence. *Review of Economics and Statistics*, 97(2):485–497.

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

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

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

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542.

Mathieu, J., Gilson, L., and Ruddy, T. (2006). Empowerment and Team Effectiveness: An empirical Test of an Integrated Model. *Journal of Applied Psychology*, 91(1):97–108.

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.

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.

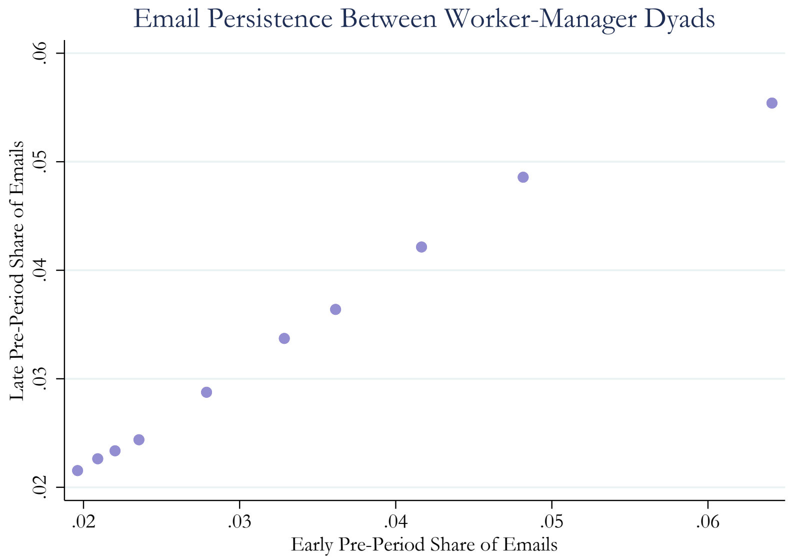
Rasul, I. and Rogger, D. (2018). Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service. *The Economic Journal*, 128(608):413–446.

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

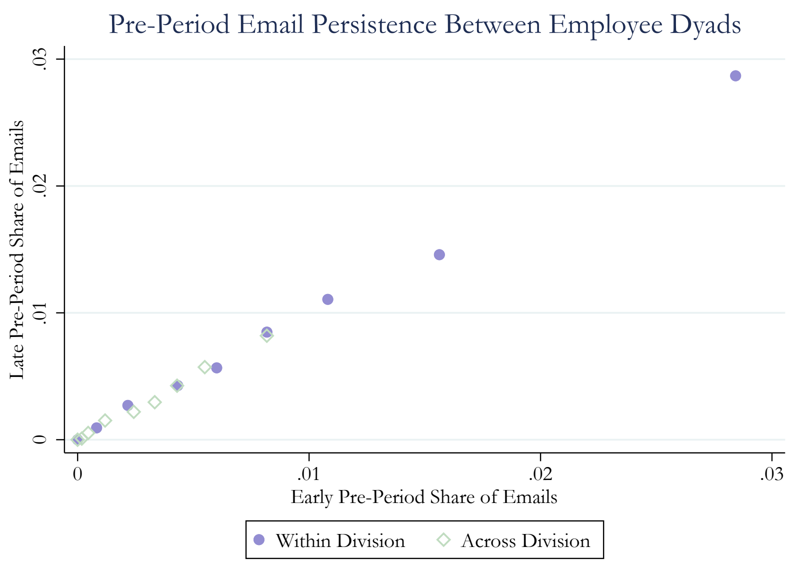
Sandvik, J., Saouma, R., Seegert, N., and Stanton, C. T. (2021). Treatment and selection effects of formal workplace mentorship programs. Technical report, National Bureau of Economic Research.

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



(a) Emails from Workers to Managers

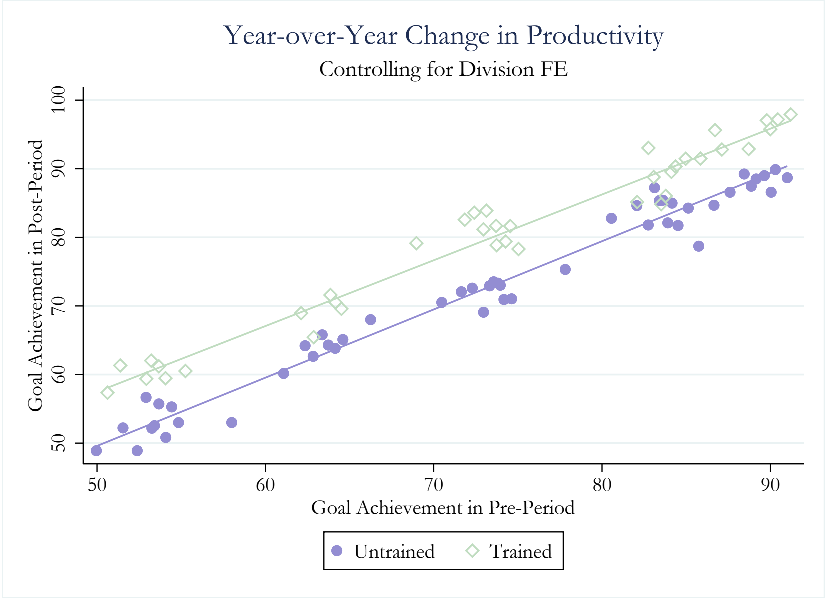


(b) Emails from Workers to Workers

Figure 1: 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 or worker-worker dyads 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. For worker-to-worker dyads, we distinguish between email persistence to workers within and outside of the division. Workers do not email managers outside of their own division.

(a) Raw Goal Achievement



(b) Net of Division Fixed Effects

Figure 2: Goal Achievement in the Pre- and Post-Period for Trained and Untrained Workers

Note: This figure displays pre-period individual goal achievement and post-period individual goal achievement for frontline workers. The unit of observation is worker-by-year. The top figure is raw goal achievement, whereas the bottom figure partials out Division fixed effects.

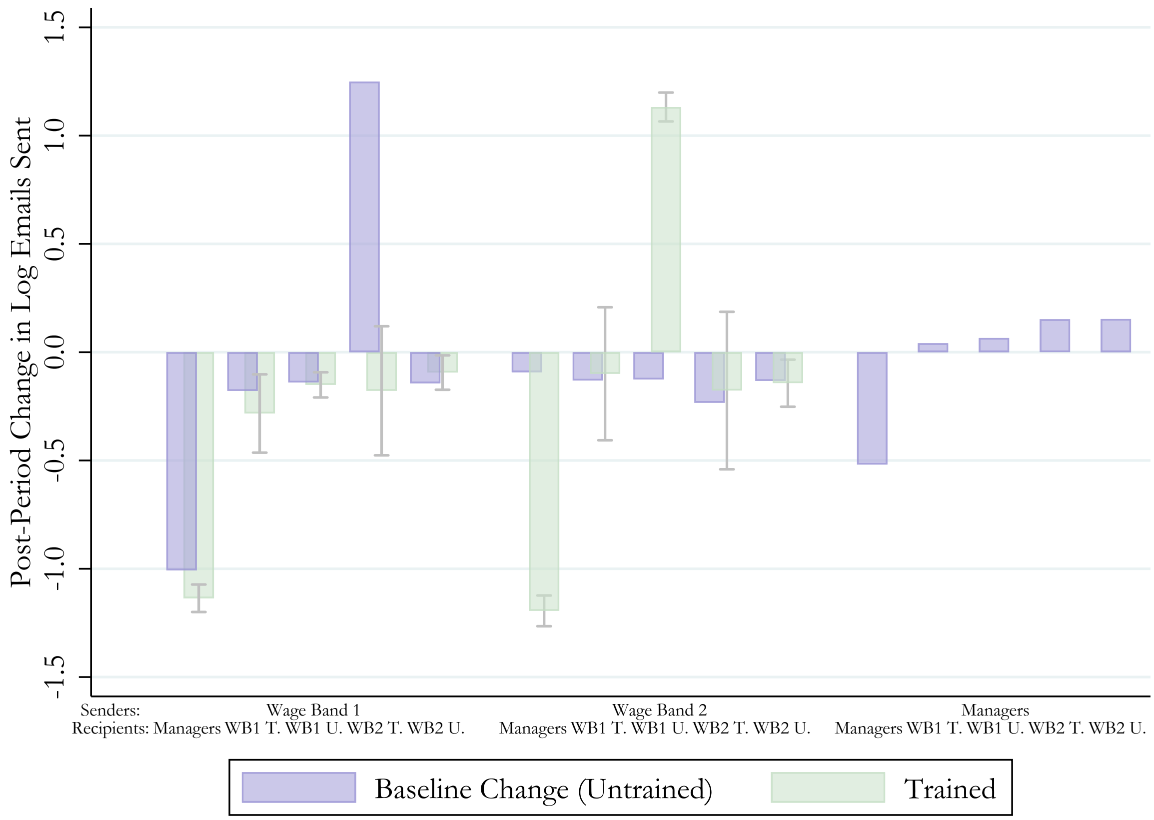


Figure 3: Changes in log Emails Sent Within Division Between the Pre- and Post-Period by Sender/Recipient Type

Note: This figure displays the average change in log emails sent at baseline for untrained workers and for trained workers. The figure splits by the sender and recipient type, with recipient type further broken down by wage band (WB1, WB2) and training status (T, U). This yields 5 types of recipients and senders: managers, trained wage band 1 and 2 workers, and untrained wage band 1 and 2 workers. The baseline change is computed as the difference in log emails sent in 2019 and log emails sent in 2018. The "Trained" change comes from the baseline change plus the coefficient on Treated x Post estimated from a difference-in-differences regression of log weekly emails, fit by recipient group, with fixed effects for workers and time. Standard errors are clustered at the sender level.

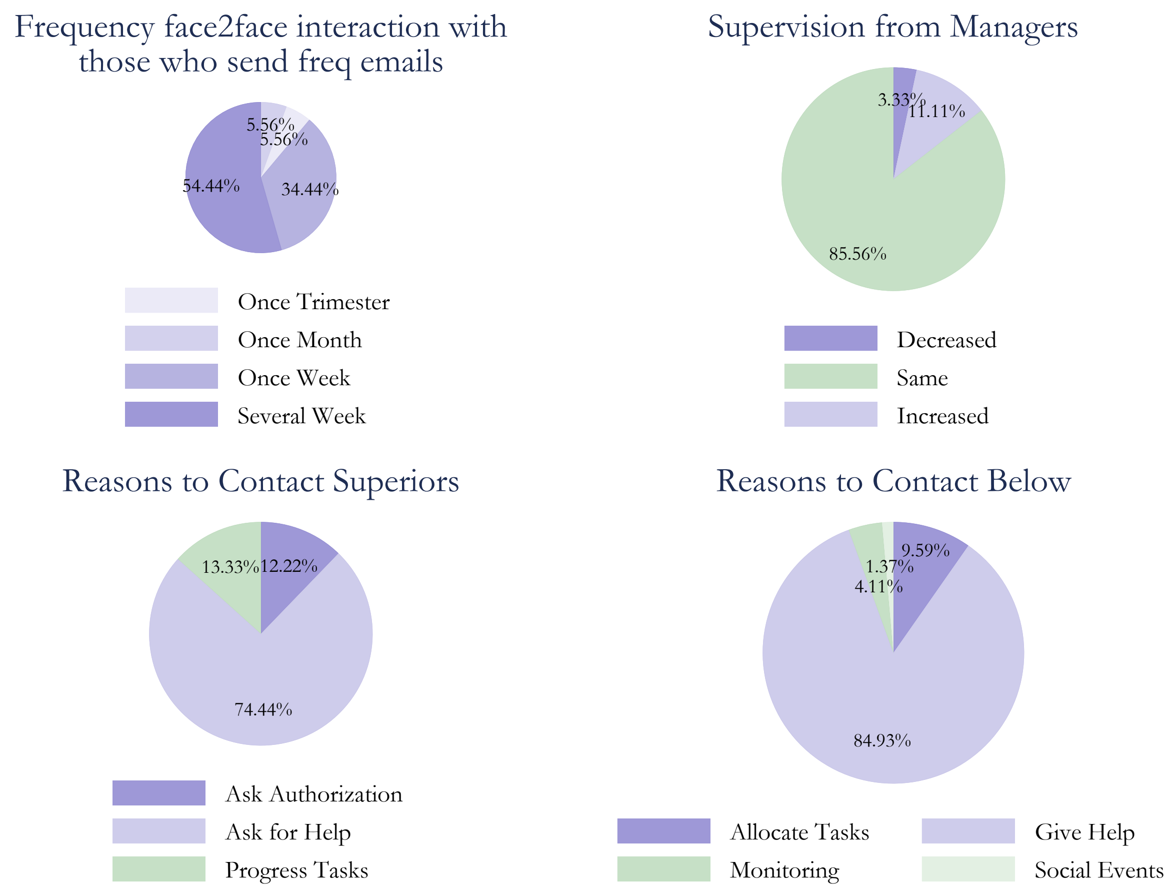


Figure 4: Distribution of Survey Responses to Questions Regarding the Mechanism

Note: This figure displays answers to an ex-post survey designed to understand the environment and mechanisms behind results. From top-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 electronically contacted workers from a higher wage band (choose only one option)." 4. "What was the main reason you electronically contacted workers from lower wage bands (choose only one option)."

	(1) W. Band 1 Workers	(2) W. Band 2 Workers	(3) Managers	(4) Untrained Workers	(5) Trained Workers	(6) Difference (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 (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 (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	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	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	

Table 1: Descriptive Statistics and Balance on Observable Characteristics

This table displays descriptive statistics for workers' observable characteristics in Wage Band 1, Wage Band 2, and Wage Bands 3-5 (Managers). The table also provides evidence of balance on observable characteristics between trained and untrained 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.

	(1)	(2)	(3)	(4)	(5)
	Untrained Workers		Trained Workers		Managers
	Within	Across	Within	Across	Within
Pre-Period Emails from Untrained Workers	4,920	13,242	4,907	13,121	12,016
	(2,411)	(2,203)	(2,352)	(2,368)	(6,095)
Pre-Period Emails from Eventually Trained Workers	674	1,796	631	1,829	1,670
	(385)	(418)	(354)	(468)	(893)
Share of Pre-Period Emails from Eventually Trained Workers	0.118	0.119	0.111	0.122	0.121
	(0.034)	(0.017)	(0.027)	(0.018)	(0.020)
Post-Period Emails from Untrained Workers	3,615	29,289	6,013	9,716	5,624
	(1,768)	(6,308)	(5,331)	(1,742)	(1,961)
Post-Period Emails from Trained Workers	817	2,525	468	1,314	428
	(465)	(868)	(301)	(367)	(264)
Share of Post-Period Emails from Trained Workers	0.180	0.080	0.088	0.119	0.067
	(0.075)	(0.026)	(0.047)	(0.023)	(0.024)

Table 2: Summary Statistics about Email Communications

Note: This table displays pre and post-period emails received by each recipient type in the columns. Email origins are divided between eventually trained and untrained workers and whether the email occurs within division (odd numbered columns) or across divisions (even numbered columns).

	(1)	(2)	(3)	(4)	(5)
Trained × Post	0.105***	0.108***	0.115***	0.100***	0.078***
	(0.006)	(0.008)	(0.009)	(0.022)	(0.028)
Wage Band 2 × Trained × Post		-0.008	-0.035**		
		(0.011)	(0.016)		
Wage Band 2 × Post		0.015**	0.008		
		(0.007)	(0.007)		
Untrained x Post x T Email Share				-1.007***	-1.116***
				(0.194)	(0.201)
Untrained x Post x T WB2 Email Share				0.881***	0.862***
				(0.272)	(0.296)
Untrained x Post x log T WB1 Emails				0.004	0.008
				(0.005)	(0.005)
Untrained x Post x log T WB2 Emails				0.014***	0.008
				(0.003)	(0.005)
Avg. Horizontal Spillover				.061	.034
Spillover Std. Error				(0.015)	(0.032)
N	13327	13327	13327	13327	13327
R^2	.903	.903	.911	.913	.914
Division-Time FE:	No	No	Yes	No	Yes

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

Table 3: Regressions of Log Goal Achievement on Training and Coworker Exposure Controls

Note: The dependent variable is log goal achievement. Measures of email exposure to eventually trained workers are computed from received emails in the pre-training period. We select relevant regressors via LASSO from a set of candidates including shares and log email levels from trained workers from the same and different divisions. Only the within division measures survive the LASSO. All models include worker and time fixed effects, while columns 3 and 5 include time-by-division fixed effects. Estimates of the average horizontal spillover take the average of the predicted value for untrained workers in the post period and standard errors are computed with 300 block bootstrap replications. Standard errors are clustered by worker.

VARIABLES	(1) absent	(2) absent	(3) absent	(4) absent
Trained × Post	0.047***	0.047***	0.011***	0.011***
	(0.004)	(0.004)	(0.001)	(0.001)
Observations	101,525	101,525	101,525	101,525
R-squared	0.852	0.853	0.880	0.881
Mean DV	.21	.21	.21	.21
Worker FE	×	×	✓	✓
Date FE	✓	×	✓	×
Division x Date FE	×	✓	×	✓
Sundays	Excluded	Excluded	Excluded	Excluded

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

Table 4: Effects of the Training Program on Absenteeism.

Note: Differences in differences regressions similar to those for log goal achievement. The dependent variable is daily absenteeism, inclusive of Saturdays. Absenteeism is calculated from the email data (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 day. All models include worker and date fixed effects. The sample is all frontline workers. Standard errors are clustered by worker.

	(1)	(2)	(3)	(4)
Panel A: Log Pre-Period Emails with Eventually Trained Workers				
Post x log Pre- Emails Received from Trained	0.063***	0.063***	0.083***	0.066**
	(0.006)	(0.006)	(0.011)	(0.033)
Post x log Pre- Emails Sent to Trained			-0.021***	-0.005
			(0.007)	(0.009)
Post	-0.430***			
	(0.043)			
Mean percentage point Δ GA	2.11	2.11	1.83	1.46
P75-P25 percentage point Δ GA	4.18	4.18	4.27	4.11
N	3276	3276	3276	3276
R^2	.95	.951	.951	.953
Panel B: Share of Pre-Period Emails with Eventually Trained Workers				
Post x Pre-Share of Emails Received from Trained	0.294**	0.292**	0.299*	0.666**
	(0.124)	(0.124)	(0.179)	(0.323)
Post x Pre-Share of Emails Sent to Trained			-0.006	-0.028
			(0.119)	(0.095)
Post	-0.005			
	(0.014)			
Mean percentage point Δ GA	2.13	2.11	2.12	5.06
P75-P25 percentage point Δ GA	.769	.764	.743	1.72
N	3276	3276	3276	3276
R^2	.943	.943	.943	.953
Time FE or Post-Indicator:	Post	Time	Time	Time
Division-Time FE:	No	No	No	Yes

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

Table 5: Effects of Pre-Period Manager 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 and sent in the pre-period between managers and 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 by manager. All models include manager and time fixed effects, while column 4 includes 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.

	(1)	(2)	(3)	(4)	(5)	(6)
Email Share Weighted Worker Log GA	0.001	0.040	-0.000	-0.003	-0.001	-0.003
	(0.004)	(0.029)	(0.001)	(0.004)	(0.001)	(0.005)
Weighted Worker Leave Out Mean Log GA					-0.080	-0.053
					(0.053)	(0.319)
N	1569	1569	1569	1569	1569	1569
R^2	1.1e-03	.072	.956	.957	.956	.957
Manager FE:	No	No	Yes	Yes	Yes	Yes
Division-Time FE:	No	Yes	No	Yes	No	Yes

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

Note: The dependent variable is weekly log goal achievement for managers. To construct the regressors, we take the email-share weighted average of connected workers' log goal achievement to measure concurrent movement of manager and worker goals. Measure that use the leave out mean control for the weighted average of workers' goal achievement in other weeks. Email shares are constructed for the entire period and are time invariant. All models include time fixed effects. Standard errors are clustered by manager.

Gains Horizon	Manager Spillover (Pct Points)	Opportunity Cost	Direct Benefit (USD)	Vertical Spillover Benefit (USD)	ROI From Direct Benefit	ROI From Direct + Spillovers
Months 1-6 Post Training	1	0	49,565	32,000	102.26%	232.84%
1 Year Post Training	1	0	99,130	64,001	304.51%	565.68%
18 Months Post Training	1	0	148,696	96,001	506.77%	898.52%
Months 1-6 Post Training	1	55,098	49,565	32,000	-37.74%	2.46%
1 Year Post Training	1	55,098	99,130	64,001	24.53%	104.93%
18 Months Post Training	1	55,098	148,696	96,001	86.79%	207.39%
Months 1-6 Post Training	2.2	0	49,565	70,401	102.26%	389.54%
1 Year Post Training	2.2	0	99,130	140,802	304.51%	879.07%
18 Months Post Training	2.2	0	148,696	211,203	506.77%	1368.61%
Months 1-6 Post Training	2.2	55,098	49,565	70,401	-37.74%	50.70%
1 Year Post Training	2.2	55,098	99,130	140,802	24.53%	201.41%
18 Months Post Training	2.2	55,098	148,696	211,203	86.79%	352.11%

Table 7: Return on Investment Under Different Scenarios

This table displays different scenarios for calculating program ROI. The first row assumes a gains horizon of 6 months, meaning that the estimated boost in goal achievement in the post-period data lasts through the first 6 months post-training and then depreciates to 0. The second and third scenarios assume a 1 year and 18 month gains horizon. These horizons are repeated for different scenarios. We vary the size of the vertical spillover to managers, from 1 percentage point to 2.2 percentage points and we vary the opportunity cost of the program from 0 to 15 days of trainees' wages. The benefits columns translate changes in goal achievement to dollar values using equation (6). Direct benefits are based on the 6.5 percentage point increase in goal achievement in Table 1. ROI calculations in each column include a $24,500 overhead cost of the program.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS		IV		First Stage	
Change in Share of Emails with Trained Workers	0.213*	-0.599***	-0.623**	-0.542*		
	(0.118)	(0.150)	(0.315)	(0.288)		
Eventually Trained Pre-Period Email Share					-0.472***	-1.140***
					(0.099)	(0.273)
N	129	129	129	129	129	129
R^2	.012	.673	.	.673	.154	.428
Division-Time FE:	No	Yes	No	Yes	No	Yes

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

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 share of emails from (eventually) trained workers. IV regressions instrument the change with the pre-period share 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 Tasks	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 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 semester of 2018?:
   (a) Yes___.
   (b) No___.
   (c) DK/NA___.[15]
3. Remember your work environment in 2018 and 2019. Consider all the people 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___.

[15] 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 that explains why you electronically contacted 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 for which you electronically contacted 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.