Should Workplace Programs be Voluntary or Mandatory? Evidence from a Field Experiment on Mentorship

Jason Sandvik* Richard Saouma † Nathan Seegert‡

Christopher Stanton§

Abstract

There is substantial variation in whether workplace training and mentorship programs are voluntary or mandatory. When programs are voluntary, many workers do not participate. We conducted a natural field experiment on a mentorship program in a sales call center where in one treatment arm, labeled the Mandatory-Condition, all subjects were either randomly assigned a mentor or not. A second treatment arm, labeled the Voluntary-Condition, required subjects to opt into the program before randomization into receiving a mentor. In the Mandatory-Condition, the mentorship treatment raised workers' daily revenue by 17% in their first two months of tenure. In the Voluntary-Condition, those who opted out of the program were substantially less productive than those who opted in, and treatment gains conditional on program participation were negligible. Comparing the conditions indicates that treatment effects are largest for workers who are most likely to opt out of participating in the program. We conclude that workplace programs can raise the productivity of lower performing employees but these workers may require inducements or mandates to participate.

We thank Emily Beam, Jasmijn Bol, Zoe Cullen, Guido Friebel, Robert Garlick, Jessica Hoel, Mitch Hoffman, Lisa LaViers, John List, Michelle Lowry, Bentley MacLeod, Robert Metcalfe, Raffaella Sadun, Harish Sujan, Jason Snyder, Brian Waters, Michael Weisbach, and seminar participants at Harvard Business School, MIT, Michigan Ross School of Business, Indian Institute of Management Ahmedabad, the University of Arizona, the Centre for Economic Policy Research, Econometric Society, the 2021 Financial Management Association Recent Research Meeting, the 2022 Labor and Finance Meeting, the Advances in Field Experiments Meeting, the 2022 Strategy Science Conference, and the 2022 NBER Summer Institute for helpful comments.

*Eller College of Management, University of Arizona

†Eli Broad College of Business, Michigan State University

‡David Eccles School of Business, University of Utah

§Harvard Business School, NBER, and CEPR

1 Introduction

Workplace programs for employee development, such as training, continuing education, and mentorship programs, are thought to increase human capital, productivity, and career advancement (Mincer, 1962; Fudenberg and Rayo, 2019). We conducted a nationally representative survey and found that many employers offer workplace development programs, but they are often not mandatory. When these programs are voluntary, over 20% of workers do not participate. We posit two potential explanations for this disconnect between programs' perceived importance and non-participation. First, workers who opt-out of participation might be the ones that would benefit the least from these programs, which makes voluntary programs an effective screening device. Second, firms may make programs voluntary, and workers may opt out because the individual returns to these programs are unclear. Learning and development program administrators frequently report difficulty in estimating the returns to workplace programs, potentially allowing sub-optimal policies for program participation to persist over time (Bloom et al., 2019).[1] To test these two explanations, we conducted a natural field experiment that allows us to estimate both the efficacy of a workplace program and treatment effect differences when the program is either voluntary or mandatory.

The context for our study is a mentorship program in a U.S.-based inbound sales call center. The workers at this firm answer incoming calls to sell digital products, like television and internet subscriptions. They are strongly incentivized to increase their individual sales, as commissions make up over a third of the median employee's compensation. This setting is well-suited to study workplace program effects because sales agents work independently of each other, we have individual daily sales performance data for all workers, and the firm regularly hires new sales agents in batches/cohorts that train together, allowing the program to be administered under different conditions for similar groups.

We designed the experiment to answer three questions. First, do workplace mentorship programs improve productivity and retention? Second, does a program's efficacy depend on whether it is mandatory or voluntary? Third, are the workers that opt out stronger or weaker? We can experimentally identify these three parameters.

The novelty of our experimental design is the ability to estimate treatment effects and compare them when the identical program is either voluntary or mandatory. This design entailed two levels of randomization: the first is at the new hire training batch/cohort level, and the second is at the agent level within a cohort. We first randomized new hire cohorts into one of two groups, labeled

[1] See, for example, Training Industry Magazine's discussion on the topic here: https://trainingindustry.com/articles/measurement-and-analytics-how-to-identify-the-right-training-kpis-for-your-learning-and-development-programs-spon-eidesign/.

the Mandatory-Condition and the Voluntary-Condition. For cohorts in the Mandatory-Condition, the lower level treatment involved randomly assigning sales agents to have a mentor or not. The Mandatory-Condition label does not imply universal participation or compliance after a mentor was assigned. We refer to agents who were assigned a mentor as "mentored," even though some did not interact with their mentor. For cohorts in the Voluntary-Condition, on the first day of training, the firm's staff asked each new hire to write a private message indicating whether they wanted to participate in a mentorship program that would begin after their two-week training. For those who opted in, the lower level treatment involved randomly assigning agents to have a mentor or not. Agents who opted out of participation were not assigned a mentor. All mentors were randomly drawn from a pool of established, non-supervisory sales agents. Matched mentor-protégé pairs were asked to meet for 30 minutes every week for four weeks and to follow a protocol. This protocol instructed protégés to share written responses to work-related questions with their mentors. Mentors were tasked with talking with their protégés, providing written feedback, and submitting the written responses and feedback to the firm's staff.

We first test whether workplace mentorship programs improve productivity and retention. Despite the prevalence of mentorship programs in many organizations, there is sparse causal evidence on mentorship effects due to non-random selection into participation (Allen et al., 2017).[2] Our intention-to-treat estimates show that agents randomized into receiving mentorship in the Mandatory-Condition generate 17% more daily revenue than non-mentored agents during their first two months on the job. Treated agents' higher productivity arises primarily from improved selling efficiency (e.g., higher revenue-per-call (RPC)), but the treatment also increases schedule adherence (i.e., up-time), increasing agents' availability to take calls. The mechanism behind these productivity gains appears to be skill acquisition, as over 80% of the treatment effects persist through agents' first six months of tenure, well after the program ends (although the estimates are less precise over longer horizons).

Mentorship also improves retention. Treated agents are significantly more likely to remain with the firm through the first 30 days on the job (where attrition rates are traditionally highest). Treatment effects on long-term retention are insignificant. Retention benefits do not explain the productivity treatment effects, as productivity gains remain (i) when accounting for non-random attrition by filling in missing data after separations with the average productivity of replacements and (ii) when using Lee (2009) bounds estimators. After accounting for the increased revenue generation of mentored agents,

[2] Over 70% of Fortune 500 companies report providing employees with mentorship opportunities (Gutner, 2009), and 45% of the respondents in our representative survey have a mentorship program at work. A recent stream of work in economics has addressed identification, albeit these research designs tend to focus on a single type of program administration rather than comparing how voluntary versus mandatory programs may impact treatment effects. See, for example, Lyle and Smith (2014), Porter and Serra (2020), and Ginther et al. (2020).

their improved retention, program administration costs, and the opportunity cost of mentor time, the firm's net present value from randomizing 114 workers into mentorship in the Mandatory-Condition was approximately $578,000 over the agents' first six months of tenure. Taken together, we find that mentorship has a substantial impact on productivity and short-term retention. The mentored agents personally benefit from this increased productivity, as they earn about 8% of their marginal revenue increase in the form of sales commissions.

Second, we test whether the treatment effects of mentorship differ for workers who opt into the program in the Voluntary-Condition. From the earliest studies of selection into programs based on treatment gains (Björklund and Moffitt, 1987), most economists are likely to believe that treatment effects will be largest in the Voluntary-Condition. On the other hand, continuing education and mentorship may disproportionately benefit workers who do not seek out opportunities, in which case individuals who opt into voluntary programs will have smaller than average treatment effects. Miscalibrated beliefs or negative stigma around seeking help are possible explanations for a negative correlation between program participation and treatment gains (Edmondson and Lei, 2014; Chandrasekhar et al., 2016; Bol and Leiby, 2018).

We find that the productivity gains from mentorship are substantially smaller when the program is voluntary, although treatment effects on retention are similar across conditions. Sales revenue and selling efficiency treatment effects for Voluntary-Condition agents who opt into the program are approximately zero. The most likely explanation for the difference in effects between the Mandatory- and Voluntary-Conditions is that the 18% of new hires who opted out of the program before randomization into treatment have the largest treatment benefits from being mentored. We reach this conclusion using both a simple pre-registered estimator and a systems GMM approach to recover heterogeneous treatment effects based on the likelihood of program opt-out.

Finally, we test whether the workers who opt-in are stronger or weaker. Given that we find that workers who opt-in have a larger treatment effect of mentorship, this also tests whether mentorship is a complement or substitute for skills. We find that those who opt-in are about 30% more productive than agents who opt into the program at baseline. This finding suggests that mentorship is a substitute for skills and can increase the productivity of the least productive workers. In our setting, we find that the treatment effect of mentorship is large enough to close the 30% productivity gap between those that opt-in and opt-out.

The productivity selection effects of program participation are not explained by observable characteristics. Demographics (e.g., age, gender, and marital status), work experience (e.g., call-center and sales experience), and Big 5 personality scores (e.g., extroversion, agreeableness) explain little variation in program participation. Non-participants have lower hiring scores from assessments made in their recruiting interviews, but these hiring scores do not predict greater program gains. As a result, program non-participation and treatment effect heterogeneity appear primarily driven by unobservables. We cannot speak directly to whether the source of the unobservable is a mistaken belief about program returns or other factors, as we feared that attempts to elicit agents' beliefs about program efficacy would look unnatural, altering participation patterns relative to workplace programs in other settings. Our nationally representative survey does, however, suggest that non-participation arises from pessimism about program benefits, social frictions that prevent workers from seeking out help from coworkers or bosses, and the cost of time spent in training. In addition, some agents who opted out in our experiment likely underestimated the benefits of the program, as our heterogeneous treatment effects estimator indicates that similar agents who were randomized into treatment in the Mandatory-Condition gained about $15 in additional daily commissions after engaging with the program (the equivalent of an additional 50 minutes at work).

Our results suggest that the firm cannot rely on workers with the largest treatment gains to sort into voluntary programs. Instead, a better approach is either universal mentorship or randomization into a mandatory program if mentorship slots are limited. Because the supply of mentors was fixed in our experiment, we consider how program returns would change if the Voluntary-Condition mentorship slots were instead allocated to agents randomly, without the ability to opt-out of the program. The estimated treatment effects in the Mandatory-Condition suggest that reallocating the 123 Voluntary-Condition mentorship slots to new agents drawn at random would have generated an additional $700,000 in profit for the firm over the agents' first six months of tenure.[3]

We also test for alternative explanations for the differences between the Mandatory- and Voluntary-Conditions. For example, we investigate whether framing the program as voluntary changes subjects' effort or engagement. We analyze meeting completion rates and recorded contents of mentoring sessions and find no evidence in support of these framing effects. As a result, we conclude that the most likely reason for treatment effect differences across the Mandatory- and Voluntary-Conditions is that different agents benefited from similar program features rather than the differing program implementation or engagement across conditions. Similarly, we have checked for violations of the Stable Unit Treatment Value assumption (SUTVA) that could affect the interpretation of our results. For example, spillovers to non-treated agents could confound our estimates. We worked closely with the firm's

[3]It is possible that the firm could improve the allocation of program resources by assigning mentors *after* observing baseline productivity, or potentially combining a baseline observation with questions about participation intent. There are apparent trade-offs associated with delaying mentorship, though, and the optimal policy from the firm's perspective accounts for the expected productivity of replacements and the time to fill positions. It remains an open question whether the Mandatory-Condition can be improved upon, though, from the firm's perspective, it strongly dominates the Voluntary-Condition.

staff to reduce the possibility that non-mentored agents (i) became discouraged after not receiving a mentor or (ii) sought internal mentors on their own. Additional tests confirm these concerns are *not* problematic in our setting. Specifically, we test for SUTVA violations by comparing the performance of non-treated new hires to hold-out groups of agents who were unaware of the mentorship program and to agents hired before the mentorship program began. We find no evidence of discouragement, leakage of the mentorship content/curricula to other agents, or crowd out of mentorship that would have developed organically absent the program.

Our design and results have implications for the evaluation of workplace programs. Selection effects that arise from how programs are administered (i.e., who participates in different types of programs) can drastically change inference. The most important point is that experiments or pilot programs on subjects who opt into participation might not generalize to universal or mandatory programs because participation correlates with treatment effect heterogeneity (List, 2022). Had our experiment only been done on agents who opted in, we would have found zero treatment effects, falsely concluding that the program has limited efficacy. When treatment gains and participation are negatively correlated, randomization among voluntary participants without accounting for selective participation increases type II errors, reducing the likelihood that good programs will be adopted. In contrast, a positive correlation between program gains and participation increases the probability of type I errors, causing benefits to be smaller at scale than observed in studies done on voluntary participants. Because the bias is determined by the covariance between treatment gains and participation, we suggest future researchers vary recruitment procedures to assess how treatment effects vary under different conditions.

Our work speaks to personnel economics studies emphasizing the important alignment between incentives and hiring practices (Friebel et al., 2019; Oyer and Schaefer, 2011). Hiring is often noisy (Hoffman et al., 2017), and incentives alone might be insufficient for *all* workers to invest in new skills through their employers' development programs. Our evidence suggests that mandatory mentorship programs can address what would otherwise appear to be a "hiring mismatch."[4] Other work shows widespread productivity dispersion across workers in a number of settings (e.g., healthcare (Finkelstein et al., 2016; Currie and MacLeod, 2017, 2020; Chan et al., 2022), judges (Coviello et al., 2014), teachers (Chetty et al., 2014), services (Lazear et al., 2015, 2016), and sales (Sandvik et al., 2021)). While most management practices and leadership studies focus on across-firm variation (Bloom and Van Reenen,

[4]The closest related work on training and potential mismatch is likely Hoffman and Burks (2020) on how overconfidence leads to training investment because workers are overconfident in their match. Our results instead indicate that some workers appear to under-invest in seeking out help, even when it is made available to them, which may be addressed with strong leadership, e.g., dictating plans for workers to improve (Lazear et al., 2015; Carter et al., 2019; Hoffman and Tadelis, 2021; Englmaier et al., 2021).

2007; Bloom et al., 2013; Syverson, 2011; Bandiera et al., 2020; Gibbons and Henderson, 2012), our findings show that within-firm variations in the administration of different practices can yield profound consequences for lifting the lower tail of the productivity distribution.[5]

Finally, we provide new estimates of the scope and characteristics of workplace programs, complementing other studies of training and mentoring programs, including those in specific industries (Reif et al., 2020; Jones et al., 2019; Rockoff, 2008; Ginther et al., 2020; Bruhn et al., 2018; Chatterji et al., 2019). The treatment effect heterogeneity in our setting also sheds light on the mechanisms behind productivity spillovers within organizations (Mas and Moretti, 2009; Bandiera et al., 2013; Herbst and Mas, 2015; Carrell et al., 2013; Lazear et al., 2015). The closest related paper, Sandvik et al. (2020), finds substantial productivity dispersion in the cross-section of workers that can be altered with coworker meetings that facilitate knowledge spillovers. The two most important differences between that prior work and this paper are that (1) in the prior paper, there was no analog to a voluntary program, and (2) the experiment in the prior paper was conducted on the cross-section of workers rather than new ones.[6] We find that providing workers with the choice to opt-out of mentoring reduces program's efficacy, whereas making mentorship mandatory can increase the productivity of low performers. Given the large dispersion in productivity documented in our prior work, our results here suggest that some workers may be persistent low performers because they do not seek out assistance or professional development opportunities that would help them close the gap between themselves and higher productivity coworkers.

2 Workplace Programs: Prevalence and Participation

Managers can enhance their employees' human capital through programs such as new hire training, continuing education, and mentorship. The efficacy of these workplace programs is of substantial

[5]Other notable experiments show that small changes can have profound effects. For example, Gosnell et al. (2020) document that practice changes for airline captains have led to out-sized fuel savings. Work by Bandiera et al. (2005) provides evidence on how social preferences interact with incentives, changing the efficacy of relative performance evaluation. At the same time, other experiments have influenced the understanding of when group incentives may work (Bandiera et al., 2013; Friebel et al., 2017).

[6]There are several important differences between this study and Sandvik et al. (2020). First, the present study tests selection into treatment, whereas the prior study had no capacity to detect whether workers would have enrolled in the programs had they been voluntary. Second, the mentorship program presented a clear expectation of who would provide information (mentors) and who would receive it (protégés), whereas Sandvik et al. (2020) randomly paired employees and treated them as equals with no designated hierarchy or roles within the pairings. Finally, the prior study involved greater treatment intensity, and participants in the prior study had greater salience that they were part of an experiment (e.g., relative performance metrics for pairs of agents were publicized). In contrast, the practices analyzed here involved fewer meetings, and subjects were more likely to view the changes as part of the firm's usual operating procedures rather than a temporary change with outside researchers present.

interest to economists (Acemoglu and Pischke, 1998), but little is known about their effectiveness due to limited data availability and a lack of experimental variation. Two important questions about these workplace programs are: how is participation in them determined (e.g., mandatory or voluntary), and which workers participate when programs are voluntary?

We conducted a nationally representative worker survey to provide answers to these questions. We administered the survey through the Lucid platform in June 2022. Respondents were paid between \$1 and \$4 for taking a 7–10 minute survey. The survey began by collecting background information about the employment status of the respondents. Only those who were employed and could pass an attention check continued on to the next set of questions about their experience with workplace programs. Specifically, we asked whether their current employer offers the following programs: (i) mentorship, (ii) training for new hires, and (iii) ongoing training or continuing education. We also asked whether the programs were required/mandatory or optional/voluntary and, if voluntary, the reasons for their participation or lack there of.[7] We display the results from this survey in Table 1.[8]

The survey responses provided three main takeaways: (1) workplace programs are ubiquitous; (2) many are voluntary; and (3) many employees do not participate when a program is voluntary. Specifically, 45% of respondents said their employer offers a mentorship program, 87% said they offer new hire training, and 80% said they offer ongoing training or continuing education. About 59% of the mentorship programs and 43% of the continuing education programs offered are voluntary. Not surprisingly, new hire training is much more likely than the other programs to be mandatory. The last column indicates that non-participation rates in voluntary programs are substantial. For example, 27% and 28% of respondents, respectively, did not participate in their employer's voluntary mentorship program or ongoing training/continuing education program. Even for new hire training, rates of non-participation exceed 20% when training is optional.

Across all voluntary programs, timing issues and doubts about personal program benefits are the most common reasons for non-participation. Forty-seven percent of non-participants in mentorship,

[7]The survey presented respondents with the following prompt: "Consider your current employer. Which of the following programs does your employer offer to you personally? If offered, are you required to participate (required/mandatory) or can you choose to participate or not (optional/voluntary)?" For each program, respondents chose between "Required or Mandatory," "Optional or Voluntary," or "Not offered." For the three core programs—mentorship, new hire training, and continuing education—if a respondent indicated that a program was voluntary, follow-up questions were asked about that person's participation and the reasons for a lack of participation, if applicable.

[8]We also included workplace wellness programs as a validation check on these answers. Sixty-five percent of our respondents indicated that their workplace has a wellness program. This is roughly comparable to numbers cited by Jones et al. (2019) from a 2016 Kaiser Family Foundation report, indicating that 53% of firms with more than 200 employees do biometric screening, 59% assess lifestyle health habits, and 83% have programs that encourage healthy lifestyles.

36% in new hire training, and 42% in ongoing training cite the time or inconvenience of program offerings. The next most common answer for all programs was "Didn't believe these programs would benefit me" (26% for mentorship, 28% for new hire training, and 31% for ongoing training). Other options such as "Didn't plan to stay at the firm, so didn't invest"; "Wanted to avoid interaction with coworkers or bosses"; and "Felt the program would benefit my employer more than it would benefit me" were selected by 8%–13% of the respondents.

These survey responses motivate the need to understand the administration of workplace programs. For example, whether a voluntary versus a mandatory program should be offered depends on the expected participation decisions of employees and on the expected benefits for those who opt in or out. The remainder of the paper studies these questions in the context of a mentorship program in a sales firm.

3 Firm Setting

The mentoring program occurred in an inbound-sales call center from January to December 2019, with data collection on protégé performance continuing after the conclusion of the mentorship program. The firm markets and sells the services of several companies/brands, most of which are television, phone, and internet providers. Sales agents answer incoming calls from potential customers and sell digital services, according to the customers' needs, with a goal of selling premium services packages. Learning the sales process (e.g., how to run credit checks or determine whether callers qualify for regional promotions) and how to up-sell can be challenging for new hires. Getting guidance through mentorship may ease new agents' learning process and skill development.

When hired, sales agents begin a two-week training program, where they learn the sales process through lectures and by listening to other agents' calls. Agents receive training that is specific to the particular sales division (i.e., product and brand) in which they will be placed once training ends. Once agents complete their two-week training, they are allocated to a team and begin responding to sales calls. Teams are typically comprised of 10–15 individuals overseen by a direct sales manager, who is responsible for monitoring performance and troubleshooting issues faced by the agents. Agents eligible for the mentorship program were spread across eight different sales divisions.

This setting has a number of attractive features for studying the efficacy of mentorship. Most importantly, the firm provided us with individual-level performance measures for each sales agent. A sales agent's productivity is independent of his or her coworkers' productivity, as incoming calls are routed to the appropriate division and are allocated to the next available agent within that division (i.e., calls are randomly allocated to agents). The process is designed so that the same agent works with the caller from start to finish. The three focal productivity measures tracked by the firm are total daily revenue, revenue-per-call (RPC), and revenue-per-hour (RPH). These metrics affect each agent's commission pay. Agents generate revenue through each sale they make and, at the end of the week, the total amount of revenue generated is multiplied by an agent's commission rate. The commission rate is a coarse function of the agent's selling efficiency (RPC and RPH), relative to other agents in the same division.[9] Multiplying the agent's weekly revenue and commission rate determines the amount of commission pay that the agent earns that week. Sales agents also earn an hourly wage, which is above the federal minimum wage and increases with tenure, and agents earn occasional bonuses for doing well during temporary promotional periods.

4 Experimental Design

The experiment involves two high-level treatment conditions, assigned at the new hire training class (cohort) level, and sub-treatments within cohort involving the assignment of mentors. Training class cohorts are specific to an office and brand. Cohorts joined the firm on a rolling basis during the experiment; some weeks multiple cohorts entered the firm, whereas in other weeks there were none. We randomly assigned each cohort to either the Mandatory-Condition (probability 40%) or the Voluntary-Condition (probability 60%).[10] Agents in the Voluntary-Condition were given the option to opt in or opt out of mentoring. Those who opted out did not receive a mentor. Agents in the Mandatory-Condition and those in the Voluntary-Condition who opted in were randomly assigned a mentor or not according to the following rule: if the supply of available mentors was greater than 50% of the cohort size, half of the agents would be assigned a mentor; otherwise the available mentors would be assigned at random to those eligible. Pairing of mentors and new hires always occurred at random.

[9] There is mild relative performance evaluation in this setting, and commissions increase at each quintile of selling efficiency. Helping another agent is unlikely to change relative rankings across quintiles, as the probability is small that any two agents are pivotal at the commission rate discontinuity. In addition, each agent has a fixed number of calls audited each week, and the commission rate decreases if the auditors identify conduct violations.

[10] Our pre-registration protocol called for the experiment to run between May 27, 2019, and December 20, 2019. The actual data we employ includes cohorts from a pilot period preceding May 27 that we had not planned to use because the mentoring protocol was slightly different (five weeks of meetings, instead of four) and because the Mandatory- versus Voluntary-Condition assignment was not originally randomized across the firm's offices. However, hiring at the firm was slower than indicated by the original projections we were given, and we could not extend the experiment to the Spring hiring season to make up for the shortfall because of the onset of COVID-19 (there is relatively little hiring from January to March). We detect no significant differences when we test for differences in treatment effects between the pilot cohorts and cohorts arriving after the pre-registered intervention start date. A multivariate test of characteristics differences does not reject the null that observables are the same between the two periods. The lack of across-office randomization of the pilot cohorts into the Mandatory-Condition and Voluntary-Condition does little to affect inference for average treatment effects because our pre-registered strategy calls for treatment effects estimation within each cohort (i.e., cohort fixed effects), absorbing mean differences across offices. Within-cohort randomization was no different in the pilot period and the period after pre-registration. The pre-registration text is documented in Appendix D, where we note instances in which there were minor deviations between the pre-registration and the implementation.

When mentor supply fell below 50% of the number of eligible new hires, the most common reason was because of obligations to mentor other cohorts in the same brand or office over a narrow time window.[11]

4.1 Treatment Allocations

Figure 1 displays the allocation of cohorts and agents to the different groups and conditions of the experiment. There were 603 program-eligible sales agents spread across 53 new hire cohorts.[12] Twenty-two cohorts and their 281 sales agents were allocated to the Mandatory-Condition, whereas the other 31 cohorts, along with their 322 sales agents, were allocated to the Voluntary-Condition. Among the agents in the Mandatory-Condition, 114 were randomized to receive a mentor, and 167 were not (which we label conditions 1a and 1b, respectively). In the Voluntary-Condition, 263 agents chose to opt in, 123 were randomized to receive a mentor, and 140 were not (conditions 2a and 2b, respectively). The remaining 59 agents (18.3%) in the Voluntary-Condition chose to opt out of receiving a mentor (condition 2c).

4.2 Timeline for Administering the Program and Communicating Treatment Allocations

Each cohort followed the same timeline. Prior to starting training, each cohort was allocated to either the Mandatory- or the Voluntary-Condition, and the staff administering the program was made aware of the cohort's assignment. All new hires were asked to complete a new hire survey on the first day of training, which asked questions about personality traits, work styles, and work experiences (specifically, whether they had call center and/or sales work experience). We use these survey responses to identify the characteristics of individuals who opted into versus opted out of mentoring.[13]

[11] In a few cases, supply constraints occurred at the individual level if the assigned mentor did not have an overlapping schedule with the protégé. Computerized randomization into treatment occurred first, followed by randomization to a mentor and then a check for feasibility. The process was not repeated if an assigned match was not feasible due to having no schedule overlap between a mentor and protégé. Infeasible matches are coded as non-treated.

[12] An additional 56 rehired agents—those who had worked at the firm previously—were also program-eligible. We exclude these agents from our main analysis, as their meaningful experience and greater initial productivity upon re-joining the company complicate comparisons between inexperienced new hires. Our results and conclusions do not change when we include these rehired agents into our analysis, but doing so requires us to present a large number of parameters from fully saturated models when we analyze the opt-out decision.

[13] Agents that were not assigned a mentor in the Mandatory-Condition had the lowest new hire survey completion rates among the five treatment arms. We suspect this is due to the within-firm mentoring staff interacting with this group the least, as the agents were never made aware of the mentorship program and, as a result, likely received less encouragement from staff to complete the survey. The staff did extensive follow-ups to encourage survey completion for agents who were assigned

For cohorts in the Voluntary-Condition, the staff described the mentoring program to the newly hired agents and told them that they could either opt in or opt out of participating. The agents were told that a randomly selected subset of those who opted in would receive a mentor, as the supply of mentors was limited, and that an outside research team would help with the randomization to ensure fairness in the assignment.[14] Agents were then asked to write on a piece of paper whether they wanted to opt in or opt out of the mentoring program, making their decision anonymous to their peers.[15]

All of these steps (announcing the program, distributing/completing surveys, collecting opt-in/out decisions from agents, and allocating agents mentors) occurred during the two weeks of formal training. The two weeks of training remained exactly the same for all agents, regardless of their mentoring group allocation. After two weeks of training, new hires graduate to work as regular agents, begin taking customer calls, and have measurable sales productivity metrics.

Agents assigned a mentor learned about their assignment a few days before they completed training. Agents in the Mandatory-Condition who were not assigned a mentor where not informed about the mentorship program by the staff. The firm's workforce management department built into mentors' and protégés' schedules specific times to meet, reducing scheduling conflicts that could prevent meetings. The mentoring relationships lasted for four weeks.

Mentors and protégés met once per week for approximately 30 minutes and completed a worksheet. They were free to discuss any topic, but the worksheet had to be completed for the mentor to receive credit for the meeting. Full documentation of the instructions for mentors can be found in Appendix B. Records of meeting occurrence and completed worksheets were kept by the staff and given to us. Shortly after their fourth and final week of mentorship meetings, protégés were asked to complete a post-mentorship survey about their experience. We use this data in Section 7.1.3 to provide context for whether meetings continued after the formal program and whether agents viewed the experience as beneficial.

---

a mentor and those in the Voluntary-Condition, as the surveys were intended to be used to study personality characteristics associated with program participation.

[14] Staff members were given some latitude to introduce the program with these broad parameters, though they were asked to read the following statement to new workers in the Voluntary-Condition: "We have recently begun a mentorship program to help newly hired sales agents when they begin working on the sales floor. Agents who opt into the program and are chosen by [the research team] will be assigned a mentor. Your mentor will approach you during your first week on the sales floor to initiate the mentoring relationship. The program will run from your first week on the sales floor to your fourth week on the sales floor, and you and your mentor will meet once a week to discuss your progress."

[15] A new hire's decision to opt out was unlikely to be influenced by aversion to the uncertainty around mentor assignment in this setting, as sales agents experience a high level of uncertainty daily, suggesting that these new hires were unlikely to be strongly uncertainty averse.

4.3 Identifying Mentors

The firm's internal mentoring staff sourced mentors by announcing to incumbent sales agents that a mentoring program for new hires would occur and that any agent could volunteer to be a mentor. The staff directly asked some promising candidates to participate. If the staff and sales managers felt an agent was not suited to be a mentor, he or she was excluded from consideration. Mentors were given two main incentives to participate. First, for each pre-scheduled, confirmed meeting they held with their protégé, they received internal currency ("kudos" dollars) worth approximately $10. Second, incumbent sales agents were told that effective mentoring would be a key indicator for promotion. While not formally necessary for promotion to a managerial role, being a mentor helped agents demonstrate their potential aptitude for leading a sales team.

4.4 Balance Across Treatments

Agent characteristics are well balanced across the different treatment groups and conditions of the experiment, as would be expected with successful randomization to treatments. In Panel A of Table 2, we consider the balance across observable characteristics for agents in the Mandatory-Condition, compared to those in the Voluntary-Condition (the top level of randomization). We do not find significant differences between conditions across cohort-level averages of agent age, gender, marital status, hiring score (interviewers' evaluation of the worker's suitability for the position), and referral status. For example, the average agent age in both groups is 23 years old, women make up 44% of the agents in the Mandatory-Condition and 40% of agents in the Voluntary-Condition, and 14%–15% of agents are married in the two groups. The average hiring scores (which have a maximum value of 1) were 0.84 and 0.85, respectively. Agents in the two groups are similarly likely to have been referred to the firm by an existing employee. Formal tests of mean differences never reject the null of equality between the Mandatory-Condition and Voluntary-Condition.[16] Appendix Table A.1 also considers balance based on the productivity metrics of incumbents in the divisions that each cohort in the Mandatory- and Voluntary-Conditions joined. Incumbent agents' average productivity levels did not differ between the Mandatory- and Voluntary-Condition.

Panel B of Table 2 considers the second level of randomization, the allocation of mentors to new hires *within* the Mandatory-Condition or Voluntary-Condition. Columns (1) and (2) show the agent-level average characteristics in the Mandatory-Condition for those who did and did not receive a mentor, respectively. These two groups are similar in age, gender, marital status, hiring scores, and

[16] Although mentors were not designated exclusively to either the Voluntary- or Mandatory-Condition, we check for balance in mentor characteristics across conditions in Table A.2. Mentors of agents in the Mandatory- and Voluntary-Conditions were similar in age, gender, and tenure. The mentors in the Voluntary-Condition were more likely to be married than those in the Mandatory-Condition. Mentors were never informed about which condition their protégés were in.

referral status. Columns (3) and (4) and the associated p-values show that agents assigned mentors and those that were not, conditional on opting into the program in the Voluntary-Condition, are similar across these observable characteristics as well. Column (5) shows that agents who opt out of mentoring in the Voluntary-Condition have worse average hiring scores. Formal tests to compare the demographics of agents across all three Voluntary-Condition cells shows similarity of age, gender, marital status, and referral status.

4.5 Subject Perceptions and Hold Out Cohorts to Test for Violations of the Stable Unit Treatment Value Assumption

Our experimental design is a natural field experiment (Harrison and List, 2004), where the participants never met the researchers, limiting experimenter demand effects. To participants, the mentoring program appeared like a normal work activity. Although participants did know that outside researchers were analyzing their survey and productivity data, the mentorship program would have appeared as a regular part of the firm's onboarding process and participants were not aware that the program, or differences between the Voluntary- and Mandatory-Conditions, were the object of researcher interest.[17] Similarly, Hawthorne effects were not likely to be an issue in our setting, as sales managers constantly monitor the same performance metrics that we study and provide agents with feedback. It is unlikely that behavior was impacted by the knowledge that outside researchers—with whom agents *never* interacted—were tracking their performance.

If agents inquired about their lack of assignment to mentorship, the mentoring staff told new hires that a limited supply of mentors meant that only half (or fewer) of them would receive a mentor and that random allocation, with the help of a team of academics, was the fairest way to distribute mentors. To further reduce any potential feelings of discouragement, those who did not receive a mentor were told that the company provided many other opportunities for new hires to receive help (e.g., from managers, coworkers, and division leaders). The staff reported to the authors on multiple check-in calls that they found no evidence of discouragement among the agents who did not receive a mentor.

Additional variation also allows us to test for discouragement effects in the control group along with other possible violations of the Stable Unit Treatment Value Assumption (SUTVA). There were 288 agents who were hired throughout the year in cohorts ineligible for the mentorship program. These cohorts were ineligible because they joined the firm at times when there was no supply of available

[17]Subjects were asked to provide informed consent when responding to the new hire survey. The survey was framed around understanding employees' preferences, work styles, and personality characteristics, so that university researchers could help the firm better serve its workforce. The consent protocol did not specify that selection into or out of the mentoring program was the key metric being studied, as this decision was elicited by the firm's staff.

mentors. Lack of availability usually occurred because potential mentors already had assignments with other recent hires, but in some cases call volumes relative to available staffing in a given division meant that mentors could not be made available. Agents in these cohorts form hold-out groups that were not aware of the mentorship program. Although these hold-outs were not randomly assigned, they have similar characteristics as program-eligible cohorts because they often joined the firm immediately after another cohort in the same division and office. In Section 6.1, we use these hold-out cohorts to compare the productivity of hold-out new hires with non-treated agents in program-eligible cohorts, showing that there are no violations of the SUTVA.

5 Results

We begin our results presentation with a description of mean differences in productivity by treatment condition and sub-treatment arm. We refer to agents assigned a mentor as "mentored," which we use as shorthand for intention-to-treat. To ease interpretation of effect sizes, throughout the paper, we use logarithms to allow readers to understand percent changes across various measures of sales productivity, but results are similar if we use productivity measures in levels. We concentrate most of the discussion on the natural logarithm of daily sales revenue, as total sales revenue relates directly to profitability for the firm after netting out commissions paid to employees. Total revenue also accounts for the opportunity cost of time spent in the mentoring program.[18]

We first show that mentoring has a positive productivity treatment effect during employees' first two months on the job (the eight weeks immediately after the end of the two-week training period). The comparison of treated and non-treated agents is shown in the two leftmost bars of Figure 2 for agents in the Mandatory-Condition (conditions 1a and 1b, respectively), who were not given the option to opt in or opt out of the program. Mentored and non-mentored agents in the Mandatory-Condition generate average daily log revenues of roughly 6.18 and 6.03, respectively, capturing a difference in productivity of approximately 15% (p-value < 0.01).[19] Figure A.1 shows agents' productivity by treatment cell during months 3–6 of tenure. The point estimates for the long-term effects are roughly similar to the short-term, 1–2 month, effects, but the standard errors are larger. These results suggest that the mandatory program increases productivity.

Second, treatment effects are much smaller for agents who voluntarily opt into the program, as shown in the comparison between 2a and 2b, (the third and fourth bars) of Figure 2. The program had

[18]The primary productivity endpoint in the pre-registration is revenue-per-call (displayed in all tables), but seminar participants encouraged us to use revenue as the main outcome measure because it accounts for how mentorship may detract from taking calls.

[19]Due to mentor supply fluctuations, treatment is not allocated evenly in each cohort. This likely drives the small difference in effect sizes between these bar charts and our regression estimations, which explicitly account for the intention to treat within cohort.

no effect for agents who opted in, as we fail to reject that productivity is the same for treated and non-treated agents who opt in (p-value = 0.331). Said differently, in contrast to the large positive effect of having a mentor in the Mandatory-Condition, we find no effect for those that opt into mentorship in the Voluntary-Condition.

The difference between the estimated treatment effect for all agents in the Mandatory-Condition, compared to the Voluntary-Condition, which comes from only those who select into the program, suggests that those who stood to gain the *most* from mentorship were those who opted out. We will soon formalize this analysis of heterogeneous treatment effects. As an intermediate step, note that our graphical evidence indicates that the strongest agents opted into the program. In the Voluntary-Condition, comparing the sales productivity of those who opted into the program but who were not randomized to receive a mentor with the productivity of those who opted out of the program (conditions 2b versus 2c) provides a measure of selection bias based on productivity levels. Agents who opted out are much less productive, with a performance difference exceeding 20% (p-value $<$ 0.01). Said differently, agents who opt into participating in the voluntary program are much more productive at baseline. Level differences between participants and non-participants do not imply smaller treatment *gains* when programs are complementary with ability. But when programs mainly substitute for ability—by improving outcomes for workers with lower ex ante performance—positive selection based on the level of performance will lead to smaller treatment gains among the workers with higher ex ante performance who opt into the program.

In summary, we find a large, positive average treatment effect on productivity from receiving a mentor in the general population (conditions 1a versus 1b), we find no effect of receiving a mentor for those that opt into the mentorship program (conditions 2a vs. 2b), and we find a positive selection effect in levels, such that agents who opt into the program have higher average daily revenues, even if they do not receive mentors (conditions 2b versus 2c).

These results have profound implications for understanding the returns to mentoring as well as for evaluating research designs intended to estimate returns to workplace programs. Our design highlights a substantial difference in treatment effects estimates based on whether program participation is voluntary or mandatory. These findings suggest that observational estimates of the returns to workplace programs are likely biased, but the direction of bias is unclear. For example, even potential designs that examine randomized programs among those who volunteer may understate effects because of the first-level selection into participation. On the other hand, studies that fail to randomize and instead compare those who select into programs with non-participants might overstate effects given underlying performance differences between those who opt in versus those who opt out.

The following subsections further unpack these results, while describing procedures that we use to infer treatment effect heterogeneity. We ultimately show that heterogeneous treatment effects are largest for those who are least likely to participate in the program. Said another way, agents who chose to opt out likely would have meaningfully benefited the most from mentorship.

5.1 Treatment and Selection Effects of Mentoring

Here we present estimates of the intention-to-treat effects of mentoring across the Mandatory- and Voluntary-Conditions, along with baseline differences for those who opt into and out of the program. To do this, we use a sample of agent-day productivity data for all program-eligible agents in their first two months on the job after they complete training, and we estimate the following model using ordinary least squares:

$$\mathrm{y}_{i,t} = \alpha + \beta_1 \mathrm{Mentored}_i + \beta_2 \mathrm{Mentored}_i \times \mathrm{Voluntary}_i + \beta_3 \mathrm{Voluntary\ Opt\text{-}Out}_i + \gamma_j + \varepsilon_{i,t}. \quad (1)$$

The right-hand side includes *Mentored*$_i$, an indicator that the agent i was randomly assigned to receive a mentor, *Voluntary*$_i$, an indicator for agents in the Voluntary-Condition, and *Voluntary Opt-Out*$_i$, an indicator for agents who chose to opt out of possibly receiving a mentor. The t subscript denotes the calendar date and γ_j is a cohort fixed effect at the unit of randomization to the high-level Voluntary- or Mandatory-Condition. Due to random assignment, the parameter β_1 is the average treatment effect of receiving a mentor. The parameter β_2 captures the difference between the average treatment effect of receiving a mentor and the conditional average treatment effect given selection into the program for those who opt in. Differences in baseline productivity between those who opt into the program and those who opt out are captured by β_3. The base levels for the Voluntary- and Mandatory-Conditions are absorbed by the cohort fixed effects, which also control for differences in productivity that are specific to the time when agents were recruited.[20] The model also contains an idiosyncratic error term, $\varepsilon_{i,t}$, and we estimate standard errors that are clustered by cohort.

The productivity outcome variable, $y_{i,t}$, differs by specification and is one of the following: ln(Revenue), ln(RPC), ln(RPH), or Adherence, where Revenue is daily total sales, RPC is revenue-per-call, RPH is revenue-per-hour, and Adherence captures how closely agents adhere to their pre-set schedule (e.g., having the requisite amount of up-time to take calls, while taking breaks and eating lunch at the correct time).[21]

The results of these estimations are reported in Columns (1)–(4) of Table 3. The point esti-

[20] All of our pre-registered specifications include cohort fixed effects, as we expected that between cohort variation would significantly increase minimum detectable effect sizes. With cohort fixed effects, calendar time and elapsed time since hire are co-linear. In a balanced panel with a short time window around a date, cohort fixed effects absorb time fixed effects. We show in Section 6.3 that our results are robust to the inclusion of date fixed effects or when we exclude cohort fixed effects.

[21] We use the natural log of one plus Revenue, RPC, or RPH. Adherence is bounded by 0 and 1.

mates on *Mentored* capture the intention-to-treat effect of being assigned a mentor for agents in the Mandatory-Condition. Column (1) shows evidence of a positive and significant treatment effect on ln(Revenue). The intention-to-treat estimate suggests that mentored agents generate 17% more in daily revenue than do their non-mentored peers.[22] In Columns (2)–(4) of Table 3, we show that the positive treatment effect for those in the Mandatory-Condition is apparent across the other performance metrics. The significant effects of *Mentored* on both selling efficiency measures—ln(RPC) and ln(RPH)—and on the main measure of time use—Adherence—suggest that mentored agents both allocate more time to revenue-generating tasks and generate more revenue per unit of time. Table A.3 reports the longer-term treatment effects of mentorship at months 3–6 of tenure. While the longer-term point estimates on *Mentored* have larger standard errors and are not statistically significant at conventional levels, they are at least 80% of the magnitude of the 1–2 month effects, suggesting that mentorship helps newly hired agents get up-to-speed more quickly while also possibly leading to a persistent shift in performance.[23]

To correct for multiple hypothesis testing, we follow Anderson (2008) and report sharpened q-values in brackets. These values are analogous to a p-value after adjusting for the False Discovery Rate (FDR). The q-values are adjusted for tests on all regressors reported in Table 3 across all six columns and for all tests conducted at the 3–6 month horizon in Table A.3. The estimated sharpened q-values are conservative in our case because they do not account for the positive correlation across tests. The q-values indicate that inference regarding our main point estimates is robust in an FDR framework.

The negative coefficients on *Mentored* × *Voluntary* in Columns (1)–(4) of Table 3 indicate that mentorship productivity effects are much smaller for those who opt into the program in the Voluntary-Condition. The penultimate row reports tests of the null that the sum of the coefficients on *Mentored* and *Mentored* × *Voluntary* equals zero. Across the productivity measures in Columns (1)–(4), we can never reject a zero treatment effect of mentorship among those who opt into the program in the Voluntary-Condition.

The point estimates on *Voluntary Opt-Out* capture selection into the program based on productivity levels. This parameter is identified based on differences in performance between those who opt out and those who opt in but do not receive a mentor. The negative, significant estimates in Columns (1)–(3) indicate that agents in the Voluntary-Condition who opted out of the mentorship

[22] When we use revenue levels as the dependent variable, we find that mentorship increases daily revenue generation by $55 per agent ($p$-value 0.08).

[23] To further unpack how productivity differences relate to time use, we report the effect of mentorship on the number of calls answered, the number of hours worked, and the number of calls answered per hour worked in Table A.4.

program performed significantly worse than non-mentored agents who signaled their interest in the program. These differences do not arise purely from differences in time use, as those who opt out have similar levels of schedule adherence to those who opt in. Furthermore, they appear less productive on a per-call basis by approximately 20% (see Column (2)). The stark differences in treatment effects between the Mandatory- and Voluntary-Condition are not driven by the pilot period cohorts, when these conditions were not randomized across the firm's offices.[24]

We estimate the impact of the mentoring program on agent retention in Columns (5) and (6) of Table 3. Call centers have notoriously high levels of attrition (Hoffman et al., 2017), and retention is an important performance metric to evaluate HR executives at the firm. To estimate retention effects, we use data with a single observation per unique hired agent, and we set the dependent variable $Tenure_t$ to equal one for agents who achieve at least t months of tenure at the firm and zero otherwise. The point estimate on *Mentored* in Column (5) indicates that agents in the Mandatory-Condition who were mentored were 10.2 percentage points more likely to achieve one month of tenure, relative to non-mentored agents. The baseline retention rate for non-mentored agents in the Mandatory-Condition (78.4% at one month) is reported in the bottom row of the table. The retention effect in month two decreases to 8.7 percentage points, relative to a baseline rate of 54.5%, and we cannot reject that the coefficient is zero at conventional levels. The point estimate on *Mentored* $\times$ *Voluntary* at the one-month horizon in Column (5) is small and insignificant, suggesting the immediate retention effect of mentorship did not depend on program conditions. The implied retention effect for those in the Voluntary-Condition is approximately zero at the two-month horizon.

These retention effects are visibly present in Kaplan-Meier survival rate estimates, which capture the fraction of agents who remain at each month of tenure. Figure A.2a shows that over 20% of non-mentored agents in the Mandatory-Condition exit the firm within one month of their hire date. Among mentored agents, only about 8% depart. The positive retention effect persists in month two, and then it narrows in months three and four. The cumulative survival rates for mentored and non-mentored agents follow a similar pattern in months five and six post-hire. The survival rates of agents in the Voluntary-Condition are depicted in Figure A.2b. This figure shows that about 17% of non-mentored opt-in agents leave the firm in the first month, whereas less than 3% of mentored opt-in agents do so. By the end of month three, however, retention rates between mentored and non-mentored opt-in agents only differ by a few percentage points, while opt-out agents are about 19 percentage points more likely to have left the firm.

We defer a comprehensive discussion of robustness to Section 6, but it is important to note that

[24]Table A.5 shows that our main results are not significantly different between the pilot period, when the Mandatory- and Voluntary-Conditions were assigned by office, compared to when cohorts were assigned at random.

the estimated productivity treatment effects of mentorship do not arise from the discouragement of agents who are randomized out of receiving a mentor. Nor does the program appear to crowd out mentoring that may have occurred even absent the program. In addition, the estimated effects on productivity are not simply driven by differential retention between mentored and non-mentored agents or by differences in meeting rates or engagement with the program.

5.2 How Do Opt-Out Agents Differ from Opt-In Agents?

The different treatment effects between the Mandatory-Condition and the Voluntary-Condition suggest that opt-out agents stood to gain from mentorship. Before we formally estimate these potential gains, we consider how agents who opt out differ from those who chose to opt into the program. We restrict our sample to only include the 322 agents in the Voluntary-Condition and estimate logistic regressions of an *Opt-Out* indicator on agent characteristics. Column (1) of Table 4, considers *Age*, *Female*, and *Married* as demographic correlates of opting out, none of which differs significantly from zero. It also includes information about the firms' assessment of the agent's suitability for the job through the variable *Hiring Score*, which is the score given to the agent by the recruiter who interviewed them for the job. We are missing hiring score data for 25 agents, so we set their hiring scores to zero and include a dummy variable indicating that they had missing data. Agents with higher hiring scores are more likely to participate in the program. The marginal effect with respect to hiring score is -0.593, which implies that an increase in hiring score of 0.10 (approximately the interquartile range in the sample) is associated with a 6% decrease in the likelihood that the agent opts out of the program.[25]

In Column (2), we control for several other characteristics that may influence the opt-out decision. Specifically, we control for the agent's location (a fixed effect for one office compared to the other), whether the agent was referred by an existing employee, and whether the agent had prior call center experience or sales experience (which we collected from the new hire survey, with dummy variables for missing data). We also include division fixed effects into the model. None of these additional factors have significant predictive power for the decision to opt out. All of these observable characteristics, together, explain less than 12% of the variation in the opt-out decision.

Column (3) adds personality characteristics collected from the initial survey, but is restricted to the 304 agents who took the survey. No personality characteristics load on participation decisions. Column (4) adds back agents who did not take the survey and includes a *Missing Survey* dummy into the model, which is significantly associated with the propensity to opt out. The pseudo R-squared

[25] While hiring scores predict the opt-out decision, we do not estimate a heterogeneous effect of hiring scores on the treatment effect of mentorship. Said another way, in the Mandatory-Condition, both mentored agents with above- and below-median hiring scores outperform non-mentored agents, and there is no differential treatment effect relative to controls with above- and below-median scores.

increases to 0.20 with the addition of the missing survey dummy. The 18 agents who failed to take the survey are very likely to opt out of the program, potentially reflecting a general disengagement with work. These results suggest measured demographics and personality traits do little to explain program participation, while hiring scores, correlates of engagement—like taking surveys—and other unobserved factors can partially explain participation decisions.

We also assess whether observable characteristics explain the low productivity of workers who opt out of the program or whether the opt out decision reveals information beyond the characteristics we see. Our approach in Columns (5)–(7) is to assess whether the coefficient on the *Voluntary Opt-Out* indicator variable in a productivity regression changes when we add controls. Our sample is agents who either opt into the program and are not mentored and those who opt out. Column (5) displays the baseline productivity regression results with controls only for demographics and an agent's referral status. We find that a 33% difference in daily revenue remains for opt-out agents compared to those who opt in. Column (6) adds data from the new hire survey and the *Missing Survey* dummy. The coefficient on *Voluntary Opt-Out* is -0.281 but it is less precisely estimated. Finally, Column (7) reports post-LASSO estimates for the correlates of sales revenue selected by the LASSO. The coefficient on *Voluntary Opt-Out* is approximately -0.4 and only two factors survive the LASSO regularization, *Female* and *High Extroversion*. Differences in observable characteristics thus do little to explain the much lower productivity of agents who opt out of the program. Based on this analysis, unobservables are more important for program participation decisions than the variables observed in standard personnel databases.

5.3 Would Opt-Out Agents Have Benefited from Mentorship?

5.3.1 Estimating Heterogeneous Treatment Effects

We can use the estimated Mandatory-Condition and Voluntary-Condition productivity treatment effects, along with the data on the fraction of Voluntary-Condition agents who opt out of receiving a mentor, to estimate the treatment effect of mentorship among opt-out agents. We pre-registered the following procedure for this purpose. Using productivity measure Y, we define the conditional average treatment effect of mentoring given selection into participation as the difference in expected production between mentored and non-mentored agents conditional on opting in: $ATE|OptIn = \beta_{OptInMentored} = E(Y_{OptInMentored}) - E(Y_{OptIn\sim Mentored})$. We can then express the unconditional average treatment effect of mentorship as the weighted average of heterogeneous effects with shares π: $ATE = E(Y_{MandatoryMentored}) - E(Y_{Mandatory\sim Mentored}) = \beta_{OptInMentored} \times \pi_{OptIn} + \beta_{OptOutMentored} \times \pi_{OptOut}$. Rearranging terms, we get,

$$\beta_{OptOutMentored} = \{ATE - \beta_{OptInMentored} \times \pi_{OptIn}\}/\pi_{OptOut}.$$

We use the estimated treatment effect in the Mandatory-Condition as the estimated ATE, and we use the estimated treatment effect in the Voluntary-Condition as the estimated $ATE|OptIn$.[26] The values of π come from the proportion of agents who opted out in the Voluntary-Condition. We show the estimated treatment effect for opt-out agents in Table 5, where standard errors come from 500 block-bootstrap iterations by cohort. The point estimate of 1.207 in Column (1) of Panel A implies that opt-out agents would have more than doubled their overall revenue generation, on average, had they received mentorship. Based on this analysis, opt-out agents were those who would have benefited the most from receiving mentorship. That is, program participation is negatively correlated with treatment gains.

We also implement a GMM estimator of heterogeneous effects that can account for the role of observables on the opt-out decision. This analysis was not pre-registered but is motivated by the desire to understand whether alternative estimation approaches meaningfully reduce the large estimated effect sizes. Specifically, the GMM estimator builds in additional flexibility for modeling the opt-out probability across cohorts, which may have differences due to sampling variation or to time trends in who is hired at the firm. In particular, we form moment conditions that correspond to two equations. The first equation accounts for the opt-in probability among agents in the Voluntary-Condition based on the moment conditions

$$E[x_i \times (1\{\text{OptIn}\}_i - \exp(x_i\delta)/(1+\exp(x_i\delta))]. \tag{2}$$

Denoting $\pi_{OptIn}(x_i, \delta) = \exp(x_i\delta)/(1+\exp(x_i\delta))$, the second equation moment conditions come from

$$\begin{aligned} E[z_i \times (Y_{it} - \text{Cohort}_i - \beta_{OptOut}(1\{\text{Mandatory}\} \times (1 - \pi_{OptIn}(x_i, \delta)) + 1\{\text{VoluntaryOptOut}\}) \\ -\beta_{OptInMentored}1\{\text{Mentored}\} \times (1\{\text{Mandatory}\} \times \pi_{OptIn}(x_i, \delta) + 1\{\text{VoluntaryOptIn}\}) \\ -\beta_{OptOutMentored}1\{\text{Mandatory}\} \times (1 - \pi_{OptIn}(x_i, \delta)))]. \end{aligned} \tag{3}$$

This is a two-equation system GMM estimator that recovers an interaction between a latent opt-out probability, estimated in Equation (2), and treatment effects, estimated in Equation (3). The procedure iteratively guesses the parameters of the opt-out probability in the Voluntary-Condition as a function of agent and cohort characteristics and applies this function as the latent probability of opting out in the Mandatory-Condition. The agent and cohort characteristics in x_i include *Age*, *Female*, *Hiring Score*, and a time trend that captures when the cohort was hired. The instruments z_i in the second equation are indicators for *Mandatory* $\times$ *Mentored*, *Voluntary Opt-In*, and *Voluntary Opt-In* $\times$ *Mentored*, along with cohort dummies. (In practice, we use the within transformation.)

[26]We include cohort fixed effects in estimating these treatment effects.

We continue to find substantial heterogeneous treatment effect estimates for non-mentored agents in the Mandatory-Condition who were deemed likely to opt out if provided the choice (the rows labeled "Opt-Out Mentored Effect" in Panels B and C). With cohort fixed effects, the estimated treatment effect of mentorship on log revenue for those who would opt out of the program is 0.794 with a standard error (clustered by cohort) of 0.359. The treatment effect is 0.531, and the standard error is 0.290, if we omit cohort fixed effects. Treatment effects for those who are likely to opt into the program, which are estimated across the Voluntary- and Mandatory-Conditions, are not significantly different from zero (the rows labeled "Opt-In Mentored Effect"). Also reported is "Opt-Out Baseline Effect," which capture the average differences in productivity for agents who opt in compared to those who opt out of the program. Across Panels B and C, agents who opt in but who are not randomized to a mentor are found to generate 33%–44% more daily revenue than those who opt out. The remaining columns show similar patterns for other productivity metrics, whereas effects on retention through one month of tenure are similar across agents who opt into the program and those who would have opted out if given the opportunity. At the two-month horizon, we estimate positive but imprecise effects on retention among those who opt out.

5.3.2 Do Effects Differ Because Characteristics of Treated Agents Differ Across the Mandatory- and Voluntary-Conditions?

We use a matching analysis as an alternative way of investigating whether differences between the Mandatory- and Voluntary-Conditions can be explained by selection on observable characteristics. To assess the role of differences in observable characteristics, we re-weight the characteristics of agents who opt into the program in the Voluntary-Condition to match the characteristics of treated agents in the Mandatory-Condition using the entropy balancing procedure proposed by Hainmueller (2012).[27] We estimate our main intention-to-treat analysis on the re-weighted data, excluding agents who opt out of the program, and report the results in Table A.6. We continue to find a positive effect of mentorship on productivity among agents in the Mandatory-Condition, and zero or negative productivity effects of mentorship among agents in the Voluntary-Condition. These results suggest that differences in the observable characteristics of treated agents across conditions (including differences in hiring scores, work experience, and personality traits) have little impact on the differences in mentorship treatment

[27] This procedure is a generalization of propensity score matching that places weights on the data to create ex ante balance in the observable characteristics of treated agents in the Mandatory-Condition and the agents in the Voluntary-Condition who opt in. By matching to the characteristics of treated agents (who are drawn at random), we can include differences in demographics, referral status, hiring scores, work experience, and Big 5 personality factors, as all mentored agents in both the Mandatory- and Voluntary-Conditions completed the new hire survey which elicited personality details, as did all but three non-mentored agents who opted into the program in the Voluntary-Condition.

effects among participants across the Voluntary- and Mandatory-Conditions. Thus, selection on unobservables into program non-participation in the Voluntary-Condition appears responsible for the larger treatment gains in the Mandatory-Condition.

6 Robustness

In this section, we rule out several alternative explanations for our results. First, we examine potential violations of the Stable Unit Treatment Value Assumption (SUTVA) that might change outcomes for non-treated agents compared to their behavior in the absence of any program. We show that our treatment effect estimates are not overstated due to discouragement. We also show that there is little evidence of information leakage to non-mentored agents, which would bias our treatment effect estimates toward zero. The program also does not appear to have crowded out mentorship that would have occurred in its absence. We then discuss a host of tests of the robustness of our results.

6.1 Effects Are Not Driven By Discouragement or Leakage

We first consider the possibility that discouragement from not being assigned a mentor or information leakage to non-mentored agents may have affected our treatment effects estimates. To understand these tests, it is useful to detail how these different channels may have interfered with our estimates. First, agents who did not receive a mentor may have become discouraged, reducing their performance. Discouragement would cause our estimates to overstate the benefits of receiving a mentor because it would negatively affect non-mentored agents in the control group for our estimations. In the Voluntary-Condition, the mentored agents' productivity is no higher than the non-mentored agents' productivity, suggesting discouragement is not at play. Second, agents who did not receive a mentor via random allocation process may have sought out their own mentors, leading to treatment leakage. A different source of leakage may be non-mentored agents querying mentored agents about the information received from mentoring. Any treatment leakage would increase the performance of non-mentored agents. Although the staff implementing the program reported no evidence of leakage in either the Mandatory- or Voluntary-Conditions, non-mentored agents who opted into the Voluntary-Condition might have been most likely to seek help, as the framing of this condition may have made agents more aware of program resources.

We implemented the mentoring program in a way that was meant to limit discouragement and leakage. First, we worked with the company to reduce the chance that non-mentored agents would find non-assignment of a mentor salient. Specifically, we asked the staff to privately notify new hires assigned to receive a mentor about their involvement in the program—reducing the potential for discouragement among non-mentored agents. Second, the firm's staff told agents in the Voluntary-Condition (for whom the program was relatively more salient) that ample opportunities for receiving help were available to those who did not end up receiving a mentor—reducing discouragement and the desire to independently seek out a mentor. We asked the staff to monitor potential discouragement and leakage throughout the study, including any complaints or concerns over not being matched to a mentor. No such feedback was received.

We formally test the net effect of discouragement and leakage by comparing the performance of three groups of agents: (1) new hires who were in program-eligible hiring cohorts; (2) new hires who were in hold-out hiring cohorts during the time of the experiment; and (3) seasoned veteran agents who began working at the firm before the onset of the mentorship program (whose tenure exceeds 18 months). Our approach is to compare the productivity of new hires in program-eligible and hold-out cohorts to veterans. Under the null of no discouragement or leakage, new hires in program-eligible cohorts who are not mentored should have indistinguishable productivity differences (relative to veterans) compared to new hires entering the firm in hold-out cohorts. We use regression adjustments to make conditions comparable between new hires and veterans across program-eligible cohorts and hold-out groups.[28] We estimate the following model using ordinary least squares:

$$\begin{aligned} \ln(\mathrm{Y})_{i,t} = \alpha &+ \beta_1 \text{New Hire}_i + \beta_2 (\text{New Hire} \times \text{Mandatory})_i + \beta_3 (\text{New Hire} \times \text{Voluntary})_i \\ &+ \beta_4 (\text{Mentored} \times \text{Mandatory})_i + \beta_5 (\text{Mentored} \times \text{Voluntary})_i + \zeta_{j,l,t} + \gamma_{n,l} + \varepsilon_{i,t}, \end{aligned} \tag{4}$$

where *New Hire* is an indicator if the agent has tenure of two months or less, *Mandatory* and *Voluntary* are indicators for the Mandatory- and Voluntary-Conditions among program-eligible cohorts, and *Mentored* is an indicator for those assigned a mentor. Our test of discouragement and leakage is the joint test that $\beta_2 = \beta_3 = 0$, indicating that the productivity of new hires relative to veterans in the Mandatory- and Voluntary-Conditions is no different than the new hire to veteran productivity differences in hold-out cohorts in the same brand and office. We include division-by-location-by-date fixed effects, $\zeta_{j,l,t}$, to account for idiosyncratic shocks that can be smoothed with different staffing choices. These fixed effects do not mean that identification of β_2 and β_3 arises from hold-out and program-eligible cohort agents working on the same day.[29] To capture potential differences in new hires across offices for all cohorts, $\gamma_{n,l}$ removes a location-by-new-hire fixed effect. Finally, $\varepsilon_{i,t}$ is an

[28]Part of this adjustment is recognizing that divisions in the firm have different levels of revenue per agent, and our approach requires a comparison of relative revenue differences between new hires and veteran agents. As a result, we restrict to divisions with five or more program-eligible agents and five or more new hires who were not program-eligible.

[29]To see this, assume β_1 hold-out cohort agents only work on day 1 in location 1 while Mandatory-Condition agents only work on day 2 in location 2. Then β_1 is the average difference between new hires in hold-out cohorts and veterans in the same brand who work on day 1 in location 1. $\beta_1 + \beta_2$ is equal to the difference between Mandatory-Condition new hires and veterans who work on day 2 in location 2.

idiosyncratic error term.

The results in Panel A of Table 6 report estimations during the experimental period to test for discouragement and leakage. The negative and statistically significant coefficient on *New Hire* in Column (1) suggests that newly hired agents generate approximately 42% less in daily revenue, relative to veterans. The small and insignificant coefficients on *New Hire* × *Mandatory* and *New Hire* × *Voluntary* suggest that newly hired non-mentored agents in program-eligible cohorts perform like newly hired agents in hold-out cohorts. We thus fail to detect evidence of discouragement or leakage. In addition, we fail to reject equality of β_2 and β_3, indicating that newly hired, non-mentored agents in the Mandatory-Condition performed like newly hired, non-mentored agents in the Voluntary-Condition. The coefficients on *Mentored* × *Mandatory* and the insignificant effect on *Mentored* × *Voluntary* align with our main treatment effects estimates in Section 5.1. Column (2) shows similar results when we use an alternative combination of fixed effects that add flexibility to capture the possibility that newly hired agent performance, relative to veteran agent performance, may vary throughout the year (through the inclusion of new hire-by-date fixed effects) or by division and office. Column (3) shows that our results are robust when controlling for agent demographic characteristics—age, gender, and marital status—which is important, given that randomization of agents into treatments did not occur for veterans and hold-out cohort agents (we do not have data on referral status or hiring scores for many veteran agents). Columns (4)–(6) repeat this exercise while using the natural log of revenue-per-call as the dependent variable. The small, insignificant coefficients on *New Hire* × *Mandatory* and *New Hire* × *Voluntary* in most of the columns further support the notion that discouragement and leakage are not likely driving our estimated mentorship treatment effects.

6.2 Did the Mentoring Program Crowd Out Organic Mentoring?

Our program may have crowded out mentoring that would have occurred in the program's absence. To test this, we assess whether non-mentored agents in program-eligible cohorts perform less well than new agents who joined the firm prior to the program. We continue to use the relative performance difference between new hires and veterans as the basis for comparison. If the mentorship program crowded out organic mentoring then we would expect non-mentored new hires who were program-eligible to perform worse than new hires from prior years. We find no such evidence. Panel B of Table 6 shows the results of estimations that resemble those in Panel A, albeit the comparison group is new hires from before the mentorship program, rather than those from contemporaneous hiring cohorts.[30] (Contemporaneous cohorts are not a good comparison group because they would be subject to the same limited supply of mentors.) The positive point estimates on *New Hire* × *Mandatory* and *New*

[30] We do not include new hire-date fixed effects in these regressions, as there was no overlap in the hiring dates of the program-eligible agents and the agents hired before the mentorship program began.

Hire × *Voluntary* in Columns (1)–(3) suggest that, if anything, new hires at baseline at the time of the program became more productive relative to veterans, suggesting that crowd-out is unlikely to be an issue in our setting. It is not surprising to find little evidence of crowd-out, as the company at baseline had relatively little organic peer-to-peer mentoring (see Sandvik et al. (2020)).

6.3 Robustness to Other Specifications

We consider several additional specifications that highlight the robustness of our main findings. We report these specifications in the coefficient plots in Figure A.3. The baseline coefficients and 95% confidence intervals from Table 3 are labeled "Baseline." We report the coefficients on *Mentored*, *Mentored* + *Mentored* × *Voluntary*, and *Voluntary Opt-Out*. First, we consider specifications that do not include cohort fixed effects (second line). In the third specification, we report estimates where we include date fixed effects, to capture variation in productivity that is idiosyncratic to a particular day on the sales floor. The fourth estimation includes controls for the agent's demographic characteristics: age, gender, and marital status. The fifth estimation includes additional controls for the agent's referral status, hiring score, call center experience, and sales experience. The sixth estimation layers on five more controls for the agent's personality traits: extroversion, agreeableness, conscientiousness, emotional stability, and openness. The seventh estimation removes observations in which agents are no longer working in the division in which they were initially hired. The point estimates and confidence intervals are quite stable across these different specifications, highlighting the robustness of our results. The only exceptions are the noisier estimates on *Mentored* when including personality controls and on *Voluntary Opt-Out* when cohort fixed effects are omitted from the regressions. We note, however, that the personality scores and work experience variables are controls from selected samples, as these variables are only populated for a subset of non-mentored agents in the Mandatory-Condition and opt-out agents. As discussed in Section 5.3.2, reweighting the characteristics of opt-in agents in the Voluntary-Condition, including personality scores, to match those of mentored agents in the Mandatory-Condition does not alter the conclusion that the Voluntary-Condition treatment effects are negligible and that the Mandatory-Condition treatment effects are substantial.

7 Mechanism Analysis, Returns to the Program, and External Validity

This section further explores possible mechanisms that may have contributed to the observed differences in mentorship treatment effects between the Mandatory-Condition and the Voluntary-Condition. We specifically consider differences in mentor meeting completion rates, worksheet content, and anecdotes from post-treatment surveys. We also consider analyses that use an alternative unit of analysis, a mentorship slot rather than the mentored agent. This allows us to estimate returns to the firm and the costs of misallocation in the Voluntary-Condition, and it leads into a discussion of tests indicating that the productivity treatment effects are not simply driven by non-random attrition. Finally, we comment on the external validity of our results.

7.1 Mechanisms

7.1.1 Meeting Completion Rates and Framing Effects

A potential mechanism for differences in treatment effects across conditions is the amount of time protégés spent interacting with their mentors. We tabulate meeting completion rates between mentor-protégé pairs in Table A.7. Of the 114 agents assigned to mentorship in the Mandatory-Condition, four never completed a recorded meeting with a mentor, while nine of the 123 treated agents in the Voluntary-Condition never met with their mentor. Mandatory-Condition protégés completed both more of their scheduled meetings (2.58 versus 2.27) and had a higher meeting completion ratio (83% versus 72%).[31] These differences could arise because the opt-in framing in the Voluntary-Condition may have portrayed the program as optional rather than a job requirement (Hossain and List, 2012; Hong et al., 2015; Englmaier et al., 2017). However, framing effects are unlikely to explain the entire difference in treatment effects between the Mandatory- and the Voluntary-Conditions, as the 72% completion rate in the Voluntary-Condition means that the vast majority of scheduled meetings occurred.

As an additional way of considering the role of meeting completion on the efficacy of mentorship, we perform an instrumental variables estimation of the effects of meetings on productivity across the two conditions. We use the assignment to a mentor indicator, *Mentored*, as an instrument for meeting completion. We report the results in Table 7, estimating separate effects for the Mandatory-Condition and Voluntary-Condition. The results in the Mandatory-Condition columns suggest completing all meetings raised daily revenue for compliers by 20%. Agents in the Mandatory-Condition who complete only half of their meetings have treatment gains of about 10%. The same estimate for those in the Voluntary-Condition indicates that full meeting completion yields a treatment effect of -7%. These estimates indicate that it is not the lower number of meetings driving treatment effect differences. Instead, the effect of a meeting differs between the Mandatory- and Voluntary-Conditions among compliers. As a result, we conclude that the intention-to-treat differences in estimates across the Mandatory- and Voluntary-Conditions are not driven by compliance differences. Two possibilities then explain the different effects for a meeting: i) the identity of compliers differs across the two conditions or ii) the content of meetings differs. Analysis of the records from meetings in the next

[31] While the mentoring protocol called for one meeting per week for four weeks, there were instances in which either a mentor, protégé, or both were absent from work for an extended period of time (e.g., on vacation), reducing the number of possible scheduled meetings from four to three (or fewer, in some cases). As such, the denominator of the meeting completion ratio is occasionally less than four.

section helps to rule out content differences.

7.1.2 Meeting Contents Captured by Worksheets

Examining the recorded worksheets that come from the mentor-protégé meetings suggests that buy-in and meeting contents were roughly similar for Mandatory-Condition and Voluntary-Condition agents. We show this in two ways. First, we consider the amount of content transcribed on each agent's worksheets by counting the total number of words written. While this is an imperfect measure of the quality of the mentor-protégé meetings, it proxies for the agents' level of engagement. In our second approach, which is motivated by the worksheet analysis in Sandvik et al. (2020), we use a bag-of-words to determine how much of a response's content is focused on job-specific skills and knowledge and how much is focused on receiving support or encouragement.[32] Specifically, we tabulate the number of "skill" words an agent uses in their responses, and we do the same thing for the number of "support" words. Words that are not classified as either support words or skill words are categorized as "other," including stop words.

We then compare the worksheet content of Mandatory-Condition agents and Voluntary-Condition agents. We have completed worksheet data for 159 out of the 224 mentored agents, as some worksheets that were turned in to the internal mentoring staff were never returned to us. For each agent, we compute the number of words written on all of their completed worksheets, and we divide this by the number of worksheets received. We do the same thing to create variables for the number of skill words per worksheet, support words per worksheet, and other words per worksheet. We then regress these agent-level word count variables on an indicator for the *Mandatory-Condition*. Table A.8 reports the results. We do not find meaningful differences in the number of total words, skill words, or other words recorded by agents in the two conditions, though Mandatory-Condition agents use about 0.15 more support words than do Voluntary-Condition agents. Taken together, worksheet content suggests similar levels of engagement in meetings between the two treatment conditions. As a result, we conclude that the most likely reason for treatment effect heterogeneity is that different agents benefited from similar program features, rather than the possibility that program implementation or engagement differed by treatment condition. That is, agents who opted out of mentorship would most likely have benefited more from the same types of mentorship that agents who opted in received.

7.1.3 Post-Treatment Impressions and Survey Responses

The results in Table 5 indicate that opt-out agents would have earned about $15 more per day had they received mentorship, equivalent to about 50 minutes of work under their hourly compensation rates, based on their earnings through increased commission pay. To assess whether these likely opt

[32]We list the words in each category in Appendix C, along with multiple example responses.

out agents may have re-assessed the non-participation decision ex-post, we examine whether mentored agents with low hiring scores (the characteristic most predictive of opting out) appear disengaged or dissatisfied with being mentored, as revealed by low meeting completion rates. Figure A.4 shows that Mandatory-Condition treated agents with the lowest hiring scores meet with their mentors frequently and have high compliance with the program, inconsistent with dissatisfaction after experiencing the program for those with the highest opt-out likelihood.

Two weeks after mentors and protégés completed their final meeting, staff asked protégés to complete a post-mentorship survey. The completion rates for this survey were quite low (less than 10%), as the firm did not monitor or provide incentives for completion. The anecdotal survey responses we did collect indicate positive perceptions of the program. Figure A.5 shows that protégés, on average, felt like they and their mentors both benefited from the program and that they did not feel like the program distracted them from their work or advancement. While the relationships between mentors and protégés did not extend beyond the workplace, the average respondent said they continued to seek out help/advice from their mentor and that their mentor continued to teach them skills after the program ended. Respondents also reported that mentorship helped them to learn selling tactics and that the program increased their day-to-day satisfaction at work. This suggests that protégés likely benefited from both an enhanced knowledge of sales techniques and greater social support in the office.

7.2 A Mentorship Slot as the Unit of Analysis and the Role of Attrition for the Productivity Estimates

We also investigate the impact of retention differences on our productivity treatment effects by estimating the treatment effects on a panel based on mentorship slots as the unit of analysis. Specifically, we fill in the productivity of agents who leave the firm with the expected productivity of a replacement for both mentored and non-mentored agents. The total productivity gain to the firm from a mentorship slot is the relative productivity gain of the treated agent while the agent is retained and then the productivity of a replacement going forward. A similar approach is used for the productivity of a non-mentored control.[33] When using the mentorship slot as the unit of analysis and filling in replacement productivity, we find results that largely mirror our main results on the unbalanced panel (see Table A.9).[34] Using the mentorship slot as the unit of analysis, the per-worker benefits

[33]We do this by computing the average productivity of newly hired non-eligible agents in the same location-division-year-quarter as the departed agent. We then re-estimate our main regression models using Equation (1).

[34]We also estimate treatment effect bounds that account for non-random attrition, as proposed by Lee (2009). The key assumption when implementing this approach is that some mentored agents would have left the firm absent mentorship but no mentored agents left the firm *because* they were

to the firm in the Mandatory-Condition remain large and positive, while they are negligible in the Voluntary-Condition.

7.3 Net Present Value of Mandatory Mentorship and the Costs of Misallocation

To estimate returns to the program, we use the mentorship slot as the unit of analysis, account for the productivity of replacements when agents leave the firm, calculate productivity treatment effects through six months of tenure including any replacement agents who join (and are not mentored) after separations, and net out of the administrative and opportunity costs associated with the program for staff and mentors. Across the first six months on the job, the average agent, among the 114 mentored agents in the Mandatory-Condition, earned approximately \$58 more in daily revenue, relative to non-mentored agents in the Mandatory-Condition, which results in \$1,160 more in revenue per agent-month. The firm earns this additional revenue net of an 8% commission rate paid to sales agents. We multiply these monthly net-revenue amounts by six, the number of months, and by 114, the number of mentorship slots. We conservatively assume this additional revenue is realized at the end of the year and discount the future cash flow using a 12.5% discount rate, which gives us a present value of the additional revenue earned by mentored agents equal to approximately \$650,000.

We then subtract the estimated time costs of taking the mentors off the phone (protégés opportunity costs are included in the revenue treatment effects) and administrative costs to calculate the net present value of the mentorship program. Mentors and protégés spent 30 minutes in the mentorship meetings each week. Revenue-per-hour for mentors averaged \$146.25 and they were paid an additional \$10 of "kudos" points for completing each meeting. Together this implies a cost of \$83.13 per meeting. We include the administrative costs of the two internal mentorship staff members who oversaw the program in the two locations, estimated to be approximately \$33,750, (generously) assuming that mentoring administration accounted for 50% of their workload. This leads to a total estimated administrative cost of about \$71,657, and a net present value of the program equal to approximately \$578,000, which may be a lower bound if more productive agents allow sales managers to have larger spans of control (Espinosa and Stanton, 2021).

If the 123 mentorship slots allocated to the Voluntary-Condition had instead been allocated firm-wide to everyone at random under the Mandatory-Condition rules (which would have given mentoring

mentored (a traditional monotonicity assumption). Table A.10 reports upper and lower bounds of the estimated treatment effect in the Mandatory-Condition on productivity in months 1–2 and months 3–6 in Panels A and B, respectively. In Panel A, the lower bound in Column (1) of 0.084 is about half the size of the main effect in Table 3, and the upper bound is over double the main effect. Both upper and lowers bounds are statistically significant, suggesting our estimated treatment effects are largely attributable to the intensive margin of agents becoming more productive.

to those agents who opted out), the implied gains to the firm would have been around $\frac{123}{114} \times \$650,000 = \$700,000$ (as we already allocated all overhead and opportunity costs), and commissions to agents would have increased by about $68,500.

7.4 External validity

As part of the first wave of evidence on voluntary versus mandatory programs, we made decisions to give us high internal validity (List, 2020). Several additional points, including performing our experiment in the field, suggest our results are likely to be externally valid for workers in other frontline or entry level jobs. The most important one is that our representative worker survey shows substantial rates of non-participation in workplace programs, indicating this is a general phenomenon across firms. In addition, our participants are approximately representative of workers in similar occupations. For example, average hourly earnings at the firm were about $21 per hour, while customer service representatives, telemarketers, and miscellaneous sales representatives earned about $23 per hour and $20 per hour, at the national level and in the same state, respectively.[35] The task that agents performed in the mentorship program—reflecting on their work, sharing these thoughts with mentors, and acting on their mentors' advice—was a natural extension of their day-to-day activities. Finally, our intervention was done with features comparable to how the program would look if implemented permanently for the entire firm. To scale to all new hires, the firm would simply need to choose to give mentors more time away from their own tasks to facilitate their serving additional proteges.

8 Conclusion

Many firms use workplace development programs to augment workers' skills and human capital. Evaluation of these programs, however, is often difficult due to selection concerns around who participates. We provide new evidence on the extent and nature of workplace skill development programs, showing that many are voluntary and have substantial rates of non-participation. We then offer a novel evaluation of a mentorship program for new hires by running a natural field experiment that allows us to estimate both the efficacy of the program and treatment effect differences when it is either voluntary or mandatory.

We find that mentorship programs can increase productivity. Specifically, in one group of our experiment, labeled the Mandatory-Condition, individuals in a sales firm who were randomly assigned a mentor had revenues that were 17% higher than agents randomly not assigned a mentor in the first two months on the job. In contrast, new hires in a second group, labeled the Voluntary-Condition, who were given the opportunity to opt-in or out of the program before random mentor assignment,

[35] These figures come from the 2015–2019 5-year American Community Survey for SOC codes 43405, 41904, and 41309. To construct hourly earnings in the ACS data, we divide total individual income by the product of weeks worked last year and usual hours per week.

had very different treatment effects. Agents who opted in and received a mentor did not have higher revenues than agents who also opted in but did not receive a mentor. This finding underscores the practical importance of the potential for selection bias in program recruitment to alter inferences about program efficacy. The direction of the bias in an RCT depends on the correlation between treatment gains and participation, which we find is negative in our setting. This selection effect relates to the wellness program participation in Jones et al. (2019) and the site selection bias identified by Allcott (2015), though the latter finds a positive relationship between selection and treatment effects.

Our design also allows us to conclude that mentoring would most help agents who would opt out if given the opportunity. On-the-job training programs, like this one, may have the largest impact when they are required for all workers rather than delivered to a subset who choose to participate. In our setting, the on-the-job training delivered by mentors is a substitute for ability. Said differently, training that leverages help from coworkers can lift lower-performing workers, but these workers appear to be the *least* likely to seek out the resources for improvement.

There are many reasons why workers may choose to opt out of programs that could improve their productivity. First, they may not want to admit their difficulties, ask for help, or show weakness. Second, they may be pessimistic about program benefits. Finally, low performers may be the least engaged, which could explain why those with the lowest productivity appear least likely to participate in our setting. We find that observable characteristics, such as demographic and personality factors, do not predict self-selection into participation. Taken together, our results suggest that voluntary workplace programs may not reach workers who would benefit the most—even when high-powered incentives are in place to reward productivity—but managers can overcome this obstacle by making programs mandatory.

Implementing an on-the-job randomized control trial in a call center allows us to provide novel and generalizable insights. Some findings, however, may be context-specific, warranting future investigation. For example, we find that the on-the-job training provided by mentors is a substitute for ability. In other contexts, training may help the best workers. Our experimental design can help determine how the returns to training programs vary over the distribution of worker productivity and for those who are more or less likely to engage with workplace programs. An open question is whether in-person interventions will translate to remote or virtual environments (Bojinov et al., 2021).

References

Acemoglu, Daron, Jörn-Steffen Pischke. 1998. Why do firms train? theory and evidence. *The Quarterly journal of economics* **113**(1) 79–119.

Allcott, Hunt. 2015. Site selection bias in program evaluation. *The Quarterly Journal of Economics* **130**(3) 1117–1165.

Allen, Tammy D, Lillian T Eby, Georgia T Chao, Talya N Bauer. 2017. Taking stock of two relational aspects of organizational life: Tracing the history and shaping the future of socialization and mentoring research. *Journal of Applied Psychology* **102**(3) 324.

Anderson, Michael L. 2008. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association* **103**(484) 1481–1495.

Bandiera, Oriana, Iwan Barankay, Imran Rasul. 2005. Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics* **120**(3) 917–962.

Bandiera, Oriana, Iwan Barankay, Imran Rasul. 2013. Team incentives: Evidence from a firm-level experiment. *Journal of the European Economic Association* **11**(5) 1079–1114.

Bandiera, Oriana, Andrea Prat, Stephen Hansen, Raffaella Sadun. 2020. Ceo behavior and firm performance. *Journal of Political Economy* **128**(4) 1325–1369.

Björklund, Anders, Robert Moffitt. 1987. The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics* 42–49.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, John Van Reenen. 2019. What drives differences in management practices? *American Economic Review* **109**(5) 1648–83.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, John Roberts. 2013. Does management matter? evidence from india. *The Quarterly Journal of Economics* **128**(1) 1–51.

Bloom, Nicholas, John Van Reenen. 2007. Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics* **122**(4) 1351–1408.

Bojinov, Iavor, Prithwiraj Choudhury, Jacqueline N Lane. 2021. Virtual watercoolers: A field experiment on virtual synchronous interactions and performance of organizational newcomers. *Harvard Business School Technology & Operations Mgt. Unit Working Paper* (21-125).

Bol, Jasmijn C, Justin Leiby. 2018. Subjectivity in professionals' incentive systems: Differences between promotion-and performance-based assessments. *Contemporary Accounting Research* **35**(1) 31–57.

Bruhn, Miriam, Dean Karlan, Antoinette Schoar. 2018. The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in mexico. *Journal of Political Economy* **126**(2) 635–687.

Carrell, Scott E., Bruce I. Sacerdote, James E. West. 2013. From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica* **81**(3) 855–882.

Carter, Susan Payne, Whitney Dudley, David S Lyle, John Z Smith. 2019. Who's the boss? the effect of strong leadership on employee turnover. *Journal of Economic Behavior & Organization* **159** 323–343.

Chan, David C, Matthew Gentzkow, Chuan Yu. 2022. Selection with variation in diagnostic skill: Evidence from radiologists. *The Quarterly Journal of Economics* **137**(2) 729–783.

Chandrasekhar, Arun G., Benjamin Golub, He Yang. 2016. Signaling, stigma, and silence in social learning. *Working Paper* .

Chatterji, Aaron, Solène Delecourt, Sharique Hasan, Rembrand Koning. 2019. When does advice impact startup performance? *Strategic Management Journal* **40**(3) 331–356.

Chetty, Raj, John N Friedman, Jonah E Rockoff. 2014. Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American economic review* **104**(9) 2633–79.

Coviello, Decio, Andrea Ichino, Nicola Persico. 2014. Time allocation and task juggling. *American Economic Review* **104**(2) 609–23.

Currie, Janet, W Bentley MacLeod. 2017. Diagnosing expertise: Human capital, decision making, and performance among physicians. *Journal of labor economics* **35**(1) 1–43.

Currie, Janet M, W Bentley MacLeod. 2020. Understanding doctor decision making: The case of depression treatment. *Econometrica* **88**(3) 847–878.

Edmondson, Amy C., Zhike Lei. 2014. Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annual Review of Organizational Psychology and Organizational Behavior* **1**(1) 23–43. doi:10.1146/annurev-orgpsych-031413-091305.

Englmaier, Florian, Stefan Grimm, Dominik Grothe, David Schindler, Simeon Schudy. 2021. The value of leadership: Evidence from a large-scale field experiment .

Englmaier, Florian, Andreas Roider, Uwe Sunde. 2017. The role of communication of performance schemes: Evidence from a field experiment. *Management Science* **63**(12) 4061–4080.

Espinosa, Miguel, Christopher Stanton. 2021. Worker skills and organizational spillovers: Evidence from linked training and communications data. Tech. rep., Harvard Business School.

Finkelstein, Amy, Matthew Gentzkow, Heidi Williams. 2016. Sources of geographic variation in health care: Evidence from patient migration. *The quarterly journal of economics* **131**(4) 1681–1726.

Friebel, Guido, Matthias Heinz, Miriam Krueger, Nikolay Zubanov. 2017. Team incentives and performance: Evidence from a retail chain. *American Economic Review* **107**(8) 2168–2203.

Friebel, Guido, Michael Kosfeld, Gerd Thielmann. 2019. Trust the police? self-selection of motivated agents into the german police force. *American Economic Journal: Microeconomics* **11**(4) 59–78.

Fudenberg, Drew, Luis Rayo. 2019. Training and effort dynamics in apprenticeship. *American Economic Review* **109**(11) 3780–3812.

Gibbons, Robert, Rebecca Henderson. 2012. *What do managers do?: Exploring persistent performance differences among seemingly similar enterprises*. Harvard Business School.

Ginther, Donna K, Janet M Currie, Francine D Blau, Rachel TA Croson. 2020. Can mentoring help female assistant professors in economics? an evaluation by randomized trial. *AEA Papers and Proceedings*, vol. 110. 205–09.

Gosnell, Greer K, John A List, Robert D Metcalfe. 2020. The impact of management practices on employee productivity: A field experiment with airline captains. *Journal of Political Economy* **128**(4) 1195–1233.

Gutner, Toddi. 2009. Finding anchors in the storm: Mentors. *The Wall Street Journal* .

Hainmueller, Jens. 2012. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis* **20**(1) 25–46.

Harrison, Glenn W, John A List. 2004. Field experiments. *Journal of Economic literature* **42**(4) 1009–1055.

Herbst, Daniel, Alexandre Mas. 2015. Peer effects on worker output in the laboratory generalize to the field. *Science* **350**(6260) 545–549.

Hoffman, Mitchell, Stephen V Burks. 2020. Worker overconfidence: Field evidence and implications for employee turnover and firm profits. *Quantitative Economics* **11**(1) 315–348.

Hoffman, Mitchell, Lisa B. Kahn, Danielle Li. 2017. Discretion in hiring. *The Quarterly Journal of Economics* **133**(2) 765–800.

Hoffman, Mitchell, Steven Tadelis. 2021. People management skills, employee attrition, and manager rewards: An empirical analysis. *Journal of Political Economy* **129**(1) 000–000.

Hong, Fuhai, Tanjim Hossain, John A List. 2015. Framing manipulations in contests: a natural field experiment. *Journal of Economic Behavior & Organization* **118** 372–382.

Hossain, Tanjim, John A List. 2012. The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science* **58**(12) 2151–2167.

Jones, Damon, David Molitor, Julian Reif. 2019. What do workplace wellness programs do? evidence from the illinois workplace wellness study. *The Quarterly Journal of Economics* **134**(4) 1747–1791.

Lazear, Edward P, Kathryn L Shaw, Christopher Stanton. 2016. Making do with less: working harder during recessions. *Journal of Labor Economics* **34**(S1) S333–S360.

Lazear, Edward P, Kathryn L Shaw, Christopher T Stanton. 2015. The value of bosses. *Journal of Labor Economics* **33**(4) 823–861.

Lee, David S. 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies* **76**(3) 1071–1102.

List, John A. 2020. Non est disputandum de generalizability? a glimpse into the external validity trial. Tech. rep., National Bureau of Economic Research.

List, John A. 2022. *The voltage effect: How to make good ideas great and great ideas scale*. Currency.

Lyle, David S, John Z Smith. 2014. The effect of high-performing mentors on junior officer promotion in the us army. *Journal of Labor Economics* **32**(2) 229–258.

Mas, Alexandre, Enrico Moretti. 2009. Peers at work. *American Economic Review* **99**(1) 112–45.

Mincer, Jacob. 1962. On-the-job training: Costs, returns, and some implications. *Journal of political Economy* **70**(5, Part 2) 50–79.

Oyer, Paul, Scott Schaefer. 2011. Personnel economics: Hiring and incentives. *Handbook of Labor Economics* **4** 1769–1823.

Porter, Catherine, Danila Serra. 2020. Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics* **12**(3) 226–54.

Reif, Julian, David Chan, Damon Jones, Laura Payne, David Molitor. 2020. Effects of a workplace wellness program on employee health, health beliefs, and medical use: a randomized clinical trial. *JAMA internal medicine* **180**(7) 952–960.

Rockoff, Jonah E. 2008. Does mentoring reduce turnover and improve skills of new employees? evidence from teachers in new york city. Tech. rep., National Bureau of Economic Research.

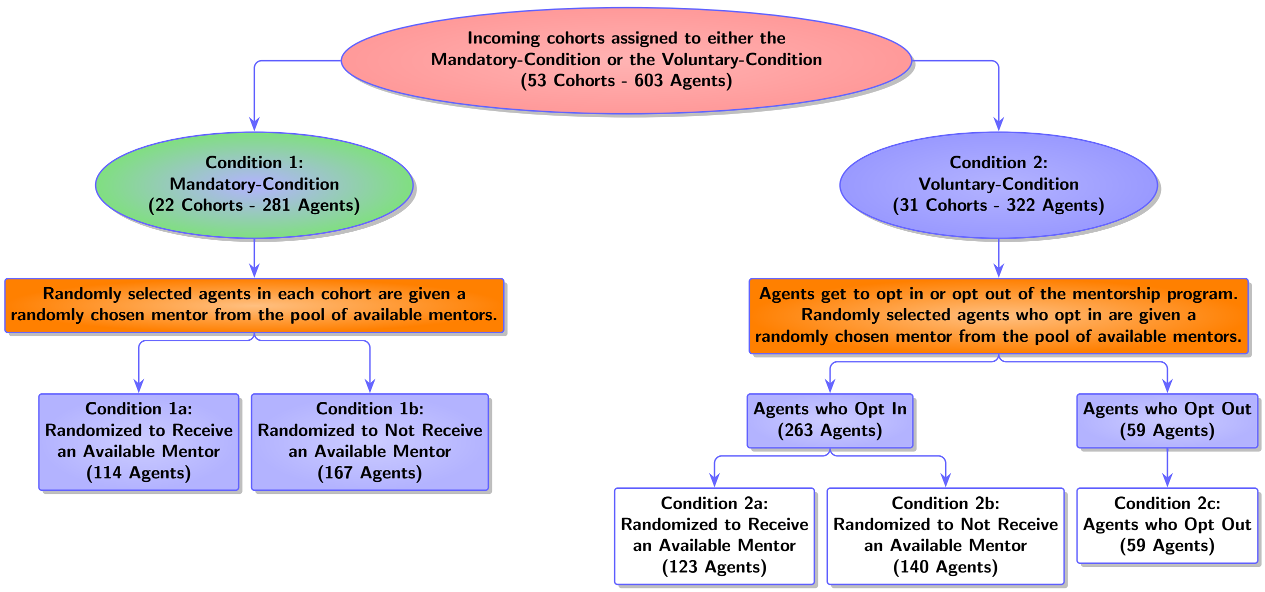
Sandvik, Jason, Richard Saouma, Nathan Seegert, Christopher Stanton. 2021. Employee responses to compensation changes: Evidence from a sales firm. *Management Science* .

Sandvik, Jason J, Richard E Saouma, Nathan T Seegert, Christopher T Stanton. 2020. Workplace knowledge flows. *The Quarterly Journal of Economics* **135**(3) 1635–1680.

Syverson, Chad. 2011. What determines productivity? *Journal of Economic literature* **49**(2) 326–65.

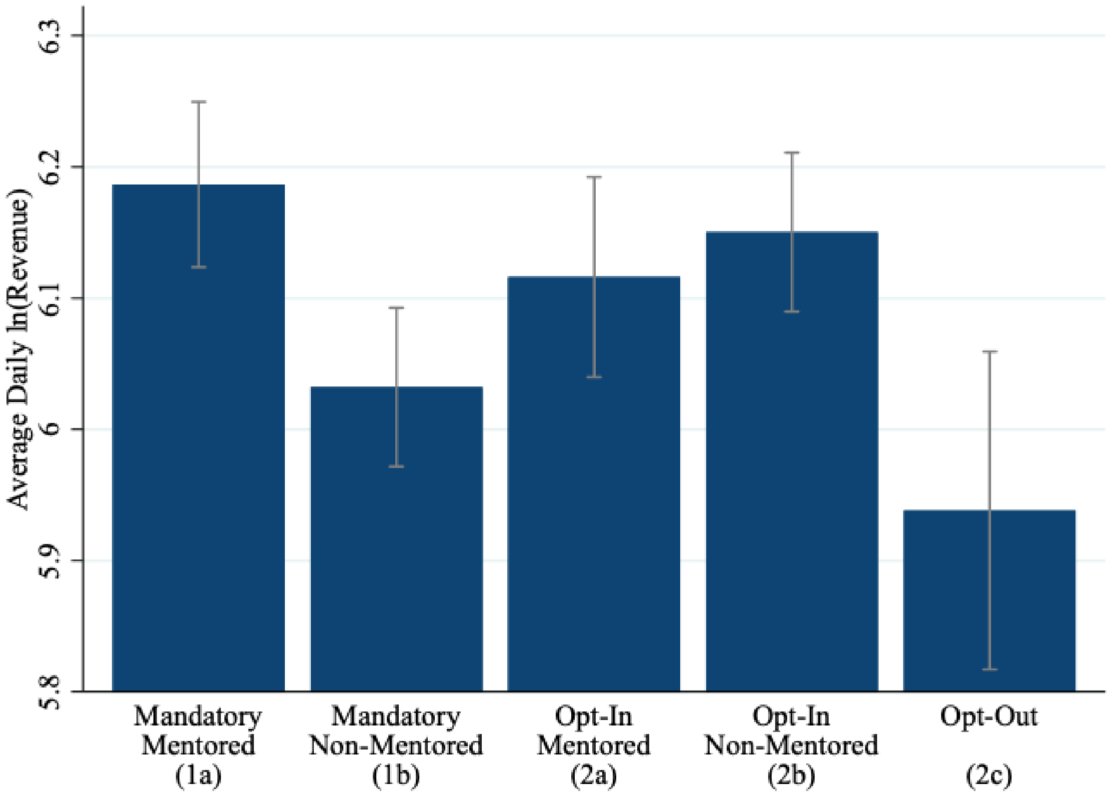
Figures and Tables

Figure 1: Allocation of Cohorts and Agents to Treatment Conditions



Notes. This figure displays the allocation of the 53 mentor-eligible cohorts to either the Mandatory-Condition or the Voluntary-Condition, our first level of variation. It then shows the allocation of the 603 mentor-eligible agents within these cohorts into different treatment conditions, our second level of variation.

Figure 2: Effect of Mentoring on Productivity



Notes. This figure plots the average daily natural log revenue (and 95% confidence intervals) for agents' first two months on the job after they complete training, split by treatment condition. Before we aggregate revenue amounts, we net out cohort fixed effects and then we add in the average productivity level across all agents that did not receive a mentor as a baseline. The p-values from difference-in-means tests that compare various treatment conditions are as follows. (1a) = (1b) p-value < 0.001; (2a) = (2b) p-value $= 0.331$; (2b) = (2c) p-value < 0.001; (1a)−(1b) = (2a)−(2b) p-value < 0.001. A similar bar chart is displayed in Figure A.1 to capture agents' productivity during months 3–6 on the job.

Table 1: Survey Data on Characteristics of Workplace Programs and Participation in Voluntary Programs

Program	Offered?	Voluntary if Offered?	Non-Participation if Voluntary?
Formal Mentorship	0.45	0.59	0.27
	(0.01)	(0.01)	(0.02)
New Hire Training	0.87	0.22	0.21
	(0.01)	(0.01)	(0.02)
Ongoing Training or Cont. Ed.	0.80	0.43	0.28
	(0.01)	(0.01)	(0.01)
N = 3,191			

Notes. This table displays means and standard deviations (in parentheses) for the prevalence and nature of a variety of workplace programs taken from responses from a nationally representative online survey conducted through the Lucid platform in June 2022. The survey was restricted to respondents currently employed by others. Respondents were asked about whether their employer offers a program and whether it is voluntary or mandatory with the question: "Consider your current employer. Which of the following programs does your employer offer to you personally? If offered, are you required to participate (required/mandatory) or can you choose to participate or not (optional/voluntary)?" For each program, respondents chose between "Required or Mandatory," "Optional or Voluntary," or "Not offered." For the three core programs—mentorship, new hire training, and continuing education—if a respondent indicated that a program was voluntary, follow-up questions were asked about their participation and the reasons for their lack of participation, if applicable.

Table 2: Balance in Agent Demographics

Panel A: Cohort-Level Balance in Agent Characteristics

	Mandatory-Condition	Voluntary-Condition	p-value
	(1)	(2)	(2)−(1)
Age (yrs.)			
Mean	22.89	23.11	0.762
Std Dev.	(2.44)	(2.68)	
Woman			
Mean	0.44	0.40	0.338
Std Dev.	(0.14)	(0.18)	
Married			
Mean	0.14	0.15	0.787
Std Dev.	(0.09)	(0.17)	
Hiring Score			
Mean	0.84	0.85	0.200
Std Dev.	(0.04)	(0.04)	
Referral			
Mean	0.55	0.59	0.447
Std Dev.	(0.18)	(0.23)	
N Cohorts	22	31	

Panel B: Agent-Level Balance in Agent Characteristics

	Mandatory-Condition			Voluntary-Condition			
	Mentored	Non-Mentored	p-value	Mentored	Non-Mentored	p-value	Opted-Out
	(1)	(2)	(2)−(1)	(3)	(4)	(4)−(3)	(5)
Age (yrs.)							
Mean	22.41	23.82	0.128	22.56	22.79	0.773	23.30
Std Dev.	(4.39)	(9.17)		(5.79)	(6.70)		(9.08)
Woman							
Mean	0.46	0.41	0.477	0.45	0.42	0.676	0.34
Std Dev.	(0.50)	(0.49)		(0.50)	(0.50)		(0.48)
Married							
Mean	0.11	0.15	0.281	0.14	0.18	0.375	0.15
Std Dev.	(0.31)	(0.36)		(0.35)	(0.38)		(0.36)
Hiring Score							
Mean	0.82	0.84	0.144	0.85	0.85	0.997	0.83
Std Dev.	(0.09)	(0.08)		(0.08)	(0.07)		(0.09)
Referral							
Mean	0.60	0.53	0.251	0.55	0.60	0.442	0.58
Std Dev.	(0.49)	(0.50)		(0.50)	(0.49)		(0.50)
Number of Agents	114	167		123	140		59

Notes. In Panel A, we average agent characteristics to the cohort-level, then take averages across cohorts. In Panel B, we take averages across agents within each treatment condition. We report standard deviations in parentheses, and we report p-values from difference in means tests when comparing cohorts in different treatment conditions, in Panel A, and when comparing agents who do and do not receive mentors, in Panel B.

Table 3: Treatment and Selection Effects of Mentoring on Productivity and Retention

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	Tenure_1	Tenure_2
	(1)	(2)	(3)	(4)	(5)	(6)
Mentored	0.170**	0.100**	0.104**	0.018*	0.102***	0.087
standard errors	(0.066)	(0.038)	(0.050)	(0.009)	(0.037)	(0.071)
sharpened q-value	[0.060]	[0.060]	[0.116]	[0.119]	[0.060]	[0.359]
Mentored × Voluntary	-0.214**	-0.150**	-0.146*	-0.014	-0.024	-0.096
standard errors	(0.095)	(0.060)	(0.080)	(0.011)	(0.053)	(0.089)
sharpened q-value	[0.086]	[0.060]	[0.159]	[0.345]	[0.723]	[0.403]
Voluntary Opt-Out	-0.323***	-0.193***	-0.215**	-0.008	-0.066	-0.171**
standard errors	(0.103)	(0.058)	(0.081)	(0.013)	(0.060)	(0.084)
sharpened q-value	[0.052]	[0.052]	[0.060]	[0.718]	[0.403]	[0.116]
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.029	0.046	0.028	0.117	0.073	0.070
Observations	15,137	15,137	15,137	15,137	603	603
p-value: Mentored + Mentored × Voluntary	0.522	0.280	0.502	0.414	0.043	0.864
Baseline Retention					0.784	0.545

Notes. The sample used in Columns (1)–(4) is composed of agent-day productivity data for all mentor-eligible agents with post-training productivity data. The data covers agents' productivity on their first two months on the job after they complete training. *Mentored* equals one for agents who were randomized to received an available mentor, and zero otherwise, *Voluntary* equals one for agents in the Voluntary-Condition, and zero otherwise, and *Voluntary Opt-Out* equals one for agents who chose to opt-out of possibly receiving a mentor, and zero otherwise. Columns (5)–(6) use data with a single observation per unique hired agent to capture retention effects. $Tenure_t$ equals one for agents who achieve at least t months of tenure at the firm, and zero otherwise. We estimate ordinary least squares regressions in all columns. All specifications include cohort fixed effects. Standard errors are clustered by cohort and are reported in parentheses. Sharpened q-values are presented in brackets, following Anderson (2008). The penultimate row reports the p-values from post-estimation tests that the sum of the coefficients on *Mentored* and *Mentored* × *Voluntary* equals zero. The bottom row reports the baseline retention estimates, measured as the fraction of non-mentored agents in the Mandatory-Condition who achieve one month (two months) of tenure in Column (5) (Column (6)). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Long-term productivity results are displayed in Table A.3.

Table 4: Determinants of Opting Out and The Relationship Between Opting Out, Productivity, and Characteristics

Dep. Variable	= 1 if Opted Out				ln(Revenue)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Voluntary Opt-Out					-0.333*	-0.281	-0.395**
					(0.195)	(0.199)	(0.173)
Age	0.015	0.005	0.017	-0.009	-0.013	-0.013	
	(0.022)	(0.024)	(0.028)	(0.036)	(0.013)	(0.012)	
Female	-0.413	-0.300	0.048	-0.007	-0.358*	-0.386**	-0.348**
	(0.285)	(0.316)	(0.333)	(0.340)	(0.176)	(0.181)	(0.150)
Married	-0.082	-0.083	-0.068	0.152	0.112	0.130	
	(0.357)	(0.375)	(0.419)	(0.377)	(0.219)	(0.248)	
Hiring Score	-4.060**	-5.652***	-5.883***	-5.680***	2.677**	2.935**	
	(1.704)	(2.054)	(2.118)	(1.983)	(1.173)	(1.127)	
Location		-0.127	0.065	-0.080			
		(0.513)	(0.450)	(0.467)			
Referral		0.115	0.258	0.204	-0.215	-0.161	
		(0.397)	(0.455)	(0.412)	(0.143)	(0.148)	
Call Center Exp.		0.693	0.655	0.689		0.235	
		(0.565)	(0.537)	(0.515)		(0.169)	
Sales Experience		-0.063	-0.032	-0.016		-0.042	
		(0.553)	(0.546)	(0.545)		(0.180)	
High Extroversion			-0.268	-0.244		0.193	0.243
			(0.385)	(0.371)		(0.176)	(0.145)
High Agreeableness			-0.529	-0.542		-0.217	
			(0.385)	(0.390)		(0.134)	
High Conscientiousness			-0.480	-0.451		-0.109	
			(0.483)	(0.472)		(0.151)	
High Emotional Stability			0.445	0.439		-0.412**	
			(0.384)	(0.379)		(0.166)	
High Openness			0.036	0.037		0.249	
			(0.433)	(0.437)		(0.167)	
Missing Survey				2.372***		-0.554*	
				(0.702)		(0.308)	
Division Fixed Effects		✓	✓	✓			
Cohort Fixed Effects					✓	✓	✓
Post-LASSO OLS							✓
(Pse.) R-Square	0.024	0.112	0.091	0.203	0.075	0.085	0.070
Observations	322	322	304	322	5,525	5,525	5,525

Notes. The sample in Columns (1)–(4) is restricted to the 322 agents in the Voluntary-Condition. The dependent variable is an indicator that equals one if the agent chose to opt out, and zero otherwise. We run logistic regressions of this indicator on different potential predictors of the choice to opt out. We split personality scores on the sample median to capture whether an agent's personality score is high or low along a particular dimension. For instances in which an agent did not complete the new hire survey or answer a particular question, we include a *Missing Survey* dummy variable in Column (4). In Columns (5)–(7), we regress ln(Revenue) on the opt-out decision of agents in the Voluntary-Condition and various control variables. In Column (7), we only include the control variables that survive the LASSO estimation. Standard errors are clustered by cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Estimated Treatment Effect of Mentoring Among Opt-Out Agents

Panel A: Pre-Registered Estimates of Opt-Out Treatment Effect (Months 1–2)

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	Tenure_1	Tenure_2
	(1)	(2)	(3)	(4)	(5)	(6)
Opt-Out Mentored Effect	1.207**	0.614*	0.725	0.130**	0.153	0.496
	(0.554)	(0.342)	(0.439)	(0.063)	(0.288)	(0.468)

Panel B: GMM Estimation With Cohort FE (Months 1–2)

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	Tenure_1	Tenure_2
	(1)	(2)	(3)	(4)	(5)	(6)
Opt-Out Baseline Effect	-0.331***	-0.200***	-0.217**	-0.007	-0.053	-0.155
	(0.116)	(0.069)	(0.093)	(0.011)	(0.067)	(0.103)
Opt-In Mentored Effect	-0.082	-0.073	-0.076	0.005	0.056	-0.020
	(0.072)	(0.048)	(0.066)	(0.006)	(0.036)	(0.056)
Opt-Out Mentored Effect	0.794**	0.543**	0.561*	0.057	0.063	0.195
	(0.359)	(0.234)	(0.294)	(0.039)	(0.151)	(0.276)
Opt-In Likelihood			0.729			

Panel C: GMM Estimation Without Cohort FE (Months 1–2)

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	Tenure_1	Tenure_2
	(1)	(2)	(3)	(4)	(5)	(6)
Opt-Out Baseline Effect	-0.436**	-0.181	-0.376**	-0.090***	-0.072	-0.210
	(0.177)	(0.142)	(0.150)	(0.027)	(0.087)	(0.136)
Opt-In Mentored Effect	-0.066	-0.057	-0.073	0.003	0.076*	-0.003
	(0.074)	(0.057)	(0.065)	(0.007)	(0.039)	(0.050)
Opt-Out Mentored Effect	0.531*	0.281	0.232	-0.019	-0.022	0.153
	(0.290)	(0.221)	(0.228)	(0.036)	(0.134)	(0.205)
Opt-In Likelihood			0.729			

Notes. The results in Panel A show estimates of the treatment effect among agents who opt out of mentorship. To estimate standard errors, we block-bootstrap by cohort (N = 53) over the whole procedure, with 500 bootstrap replications for each column. Panels B and C report two-step GMM estimates as described in the text. The *Opt-Out Baseline Effect* is the difference in productivity among untreated agents who participate in the program and those who opt out. The *Opt-In Mentored Effect* is the effect of mentoring for agents who opt in relative to their baseline effect. The *Opt-Out Mentored Effect* is the effect of mentoring in the mandatory program for those who would likely not have participated in the voluntary program. We include (exclude) cohort fixed effects in the Panel B (Panel C) estimations of Equation (3). Standard errors for the GMM estimation are clustered by cohort. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Testing for Discouragement, Leakage, and Crowd-Out

Panel A: Comparing New Hires and Veterans in Mentor-Eligible and Non-Eligible Cohorts

	Ln(Revenue): Months 1–2			Ln(RPC): Months 1–2		
	(1)	(2)	(3)	(4)	(5)	(6)
New Hire	-0.422***			-0.483***		
	(0.108)			(0.072)		
New Hire × Mandatory	0.027	0.076	0.059	-0.010	0.026	0.018
	(0.086)	(0.093)	(0.093)	(0.051)	(0.051)	(0.052)
New Hire × Voluntary	0.079	0.098	0.086	0.083*	0.082	0.073
	(0.085)	(0.095)	(0.097)	(0.050)	(0.055)	(0.054)
Mentored × Mandatory	0.208***	0.213***	0.210***	0.106***	0.115***	0.111**
	(0.066)	(0.066)	(0.069)	(0.041)	(0.041)	(0.044)
Mentored × Voluntary	-0.040	-0.042	-0.025	-0.042	-0.048	-0.037
	(0.086)	(0.087)	(0.087)	(0.054)	(0.056)	(0.055)
Division-Location-Date FE	✓	✓	✓	✓	✓	✓
Location-New Hire FE	✓			✓		
Division-Location-New Hire FE		✓	✓		✓	✓
New Hire-Date FE		✓	✓		✓	✓
Demographic Controls			✓			✓
Adj. R-Square	0.084	0.089	0.094	0.112	0.117	0.120
Observations	41,867	41,858	41,858	41,867	41,858	41,858
$\text{New}_M = 0$, $\text{New}_V = 0$	0.642	0.550	0.654	0.138	0.314	0.378
New_M - $\text{New}_V = 0$	0.568	0.811	0.775	0.074	0.290	0.302
New_M - New_V + $\text{Men}_M = 0$	0.072	0.042	0.059	0.799	0.285	0.318

Panel B: Comparing New Hires and Veterans in Mentor-Eligible and Pre-Experimental Cohorts

	Ln(Revenue): Months 1–2			Ln(RPC): Months 1–2		
	(1)	(2)	(3)	(4)	(5)	(6)
New Hire	-0.503***			-0.366***		
	(0.057)			(0.039)		
New Hire × Mandatory	0.211**	0.217**	0.220**	-0.037	-0.023	-0.017
	(0.092)	(0.092)	(0.093)	(0.051)	(0.053)	(0.054)
New Hire × Voluntary	0.218**	0.246***	0.246***	0.021	0.037	0.040
	(0.085)	(0.090)	(0.090)	(0.050)	(0.054)	(0.053)
Mentored × Mandatory	0.237***	0.228***	0.227***	0.130***	0.128***	0.126***
	(0.075)	(0.075)	(0.076)	(0.044)	(0.045)	(0.047)
Mentored × Voluntary	-0.030	-0.029	-0.020	-0.034	-0.039	-0.033
	(0.071)	(0.071)	(0.070)	(0.043)	(0.043)	(0.043)
Division-Location-Date FE	✓	✓	✓	✓	✓	✓
Location-New Hire FE	✓			✓		
Division-Location-New Hire FE		✓	✓		✓	✓
Demographic Controls			✓			✓
Adj. R-Square	0.160	0.161	0.162	0.165	0.165	0.166
Observations	73,359	73,359	73,359	73,359	73,359	73,359
$\text{New}_M = 0$, $\text{New}_V = 0$	0.023	0.017	0.016	0.456	0.423	0.459
New_M - $\text{New}_V = 0$	0.929	0.735	0.753	0.214	0.193	0.223
New_M - New_V + $\text{Men}_M = 0$	0.007	0.018	0.019	0.153	0.171	0.164

Notes. Panel A reports tests of the net effect of discouragement and leakage by comparing the performance of three groups of agents: (1) new hires who were in mentor-eligible hiring cohorts; (2) new hires who were not in mentor-eligible hiring cohorts during the time of the experiment; and (3) seasoned veterans who began working at the firm before the onset of the mentorship program. The dependent variable is ln(Revenue) in Columns (1)–(3) and ln(RPC) in Column (4)–(6). In the bottom three rows, New_M stands for new hire in the Mandatory-Condition, New_V stands for new hire in the Voluntary-Condition, and Men_M stands for those who were mentored in the Mandatory-Condition. The estimations in Panel B are analogous to those in Panel A, with the exception that instead of using new hires from non-mentor-eligible cohorts as a control group, we use new hires who began working at the firm before the mentorship program began. This allows us to test for crowd-out effects. All specifications include division-by-location-by-date fixed effects. Columns (1) and (4) also include location-by-new-hire fixed effects, whereas Columns (2), (3), (5), and (6) include division-by-location-by-new-hire fixed effects. Columns (3) and (6) control for agent age, gender, and marital status. Standard errors are clustered by hiring cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 7: Instrumental Variables Estimates of Meetings with Mentors Across Different Conditions

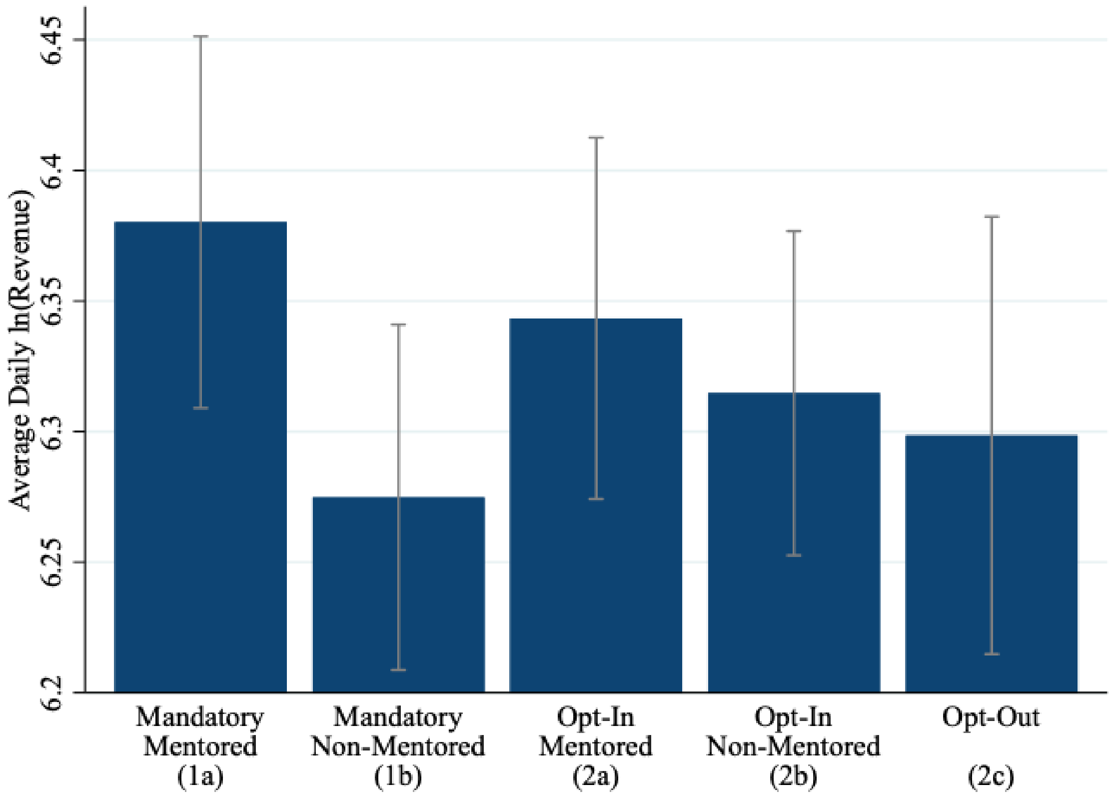
	Mandatory-Condition				Voluntary-Condition			
	First Stage (1)	IV (2)	First Stage (3)	IV (4)	First Stage (5)	IV (6)	First Stage (7)	IV (8)
Mentored	0.836***		2.842***		0.746***		2.377***	
	(0.046)		(0.206)		(0.061)		(0.172)	
Meeting Completion Ratio		0.203**				-0.073		
		(0.080)				(0.091)		
Number Recorded Meetings				0.060**				-0.023
				(0.023)				(0.029)
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓
Cragg-Donald Wald F	50,058		19,338		32,493		17,991	
Centered R-Square	0.843	0.002	0.718	0.005	0.734	0.000	0.666	-0.000
Observations	6,744	6,744	6,744	6,744	7,035	7,035	7,035	7,035

Notes. In this table we present IV regressions of the natural log of daily revenue on measures of mentor-protégé meeting completion using *Mentored* assignment as an instrumental variable. Columns (1)–(4) consider only agents in the Mandatory-Condition, and Columns (5)–(8) consider only agents in the Voluntary-Condition who did not choose to opt out of the mentorship program. The dependent variable in first stage regression Columns (1) and (5) is *Meeting Completion Ratio*, the fraction of possible (i.e., scheduled) mentor-protégé meetings that the protégé completed. The dependent variable in first stage regression Columns (3) and (7) is *Number Recorded Meetings*, the number of mentor-protégé meetings that the protégé completed. In Columns (2), (4), (6), and (8), we present IV regressions of ln(Revenue) as a function of the completion ratio or the number of completed meetings, using *Mentored* as an instrument. Agents who opt out in the Voluntary-Condition are not included in this sample. All specifications include cohort fixed effects. Standard errors are clustered by cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

ONLINE APPENDIX MATERIALS

A Appendix Figures and Tables

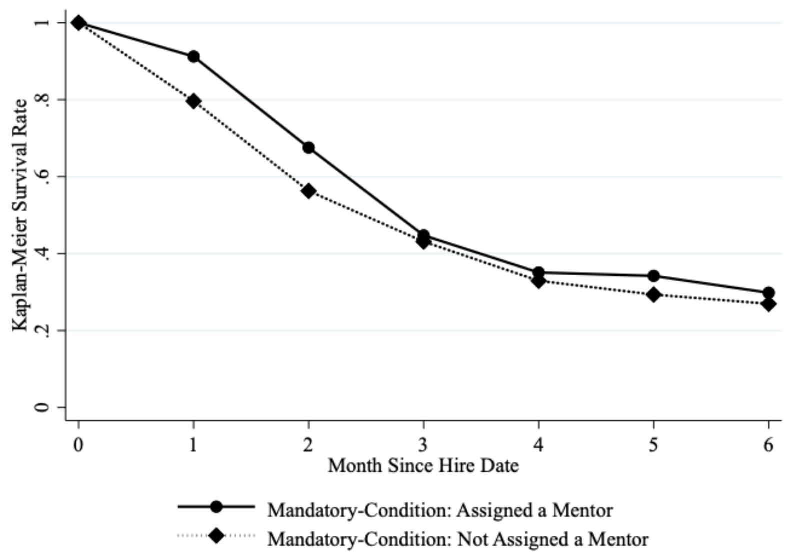
Figure A.1: Effect of Mentoring on Productivity (Months 3–6 on Sales Floor)



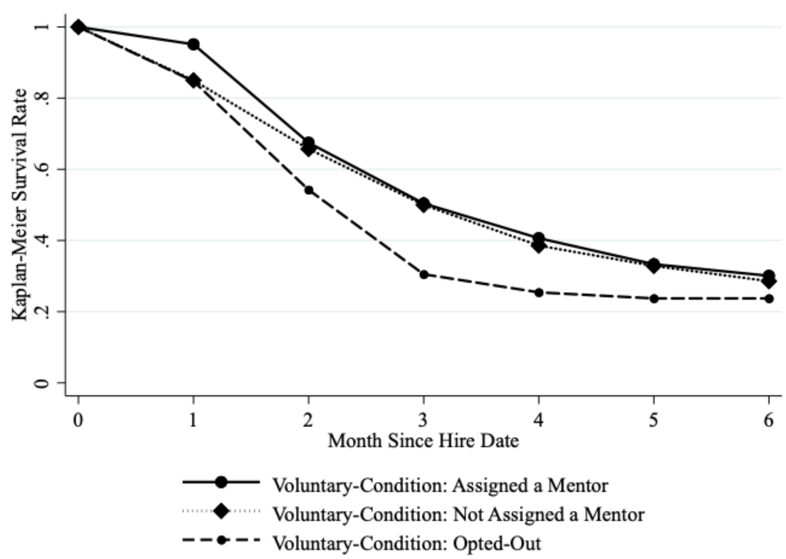
Notes. This figure plots the average daily natural log revenue (and 95% confidence intervals) for agents' third to sixth months on the job after they complete training, split by treatment condition. Before we aggregate revenue amounts, we net out cohort fixed effects and then we add in the average productivity level across all agents that did not receive a mentor as a baseline.

Figure A.2: Effect of Mentoring on Retention

(a) Agents in the Mandatory-Condition

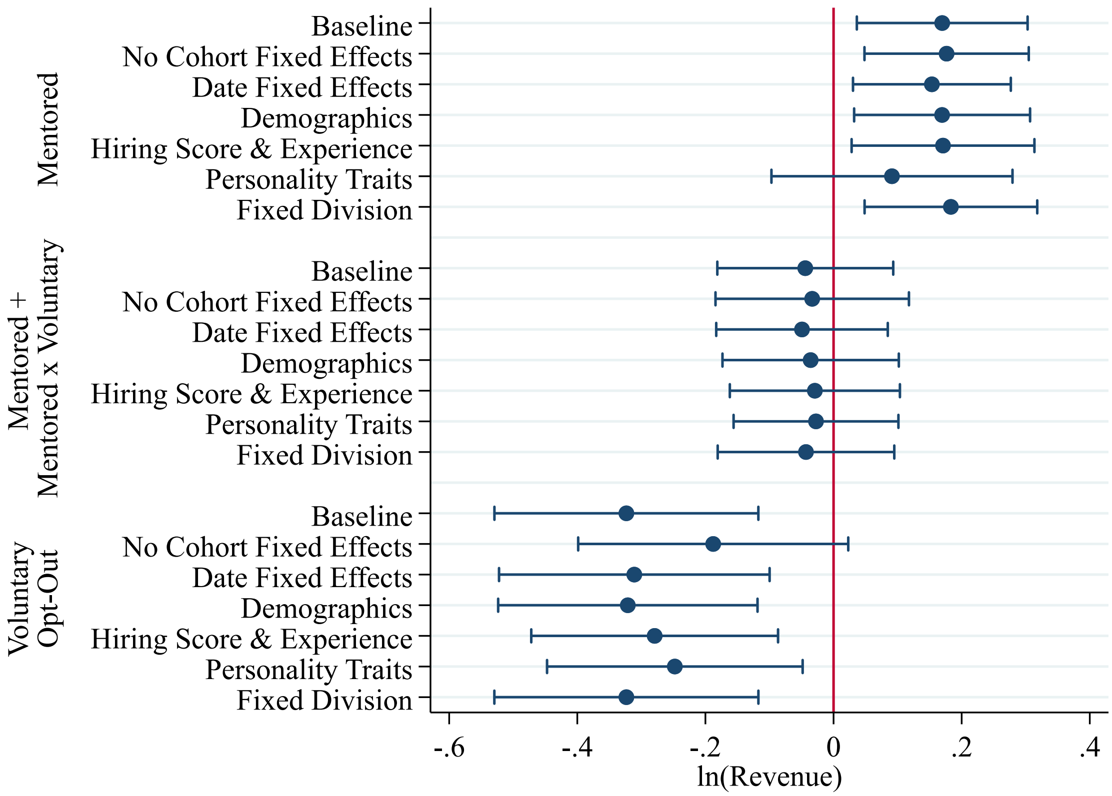


(b) Agents in the Voluntary-Condition



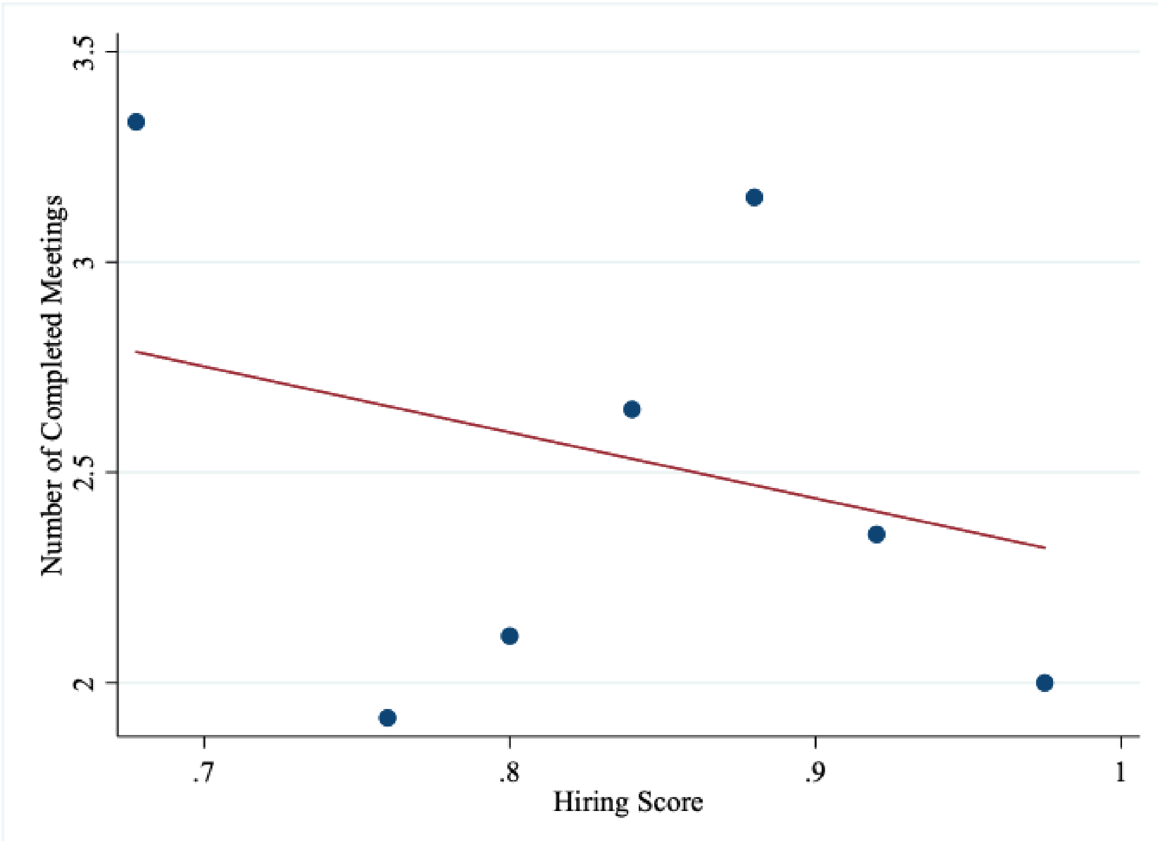
Notes. Figure (a) plots Kaplan-Meier survival rates over time for agents in the Mandatory-Condition, and Figure (b) considers those in the Voluntary-Condition. The survival rate estimator considers a starting point, in our case an agent's hire date, and then, from that time, displays the fraction of agents that remain at the firm.

Figure A.3: Robustness of the Treatment and Selection Effects of Mentoring on Productivity



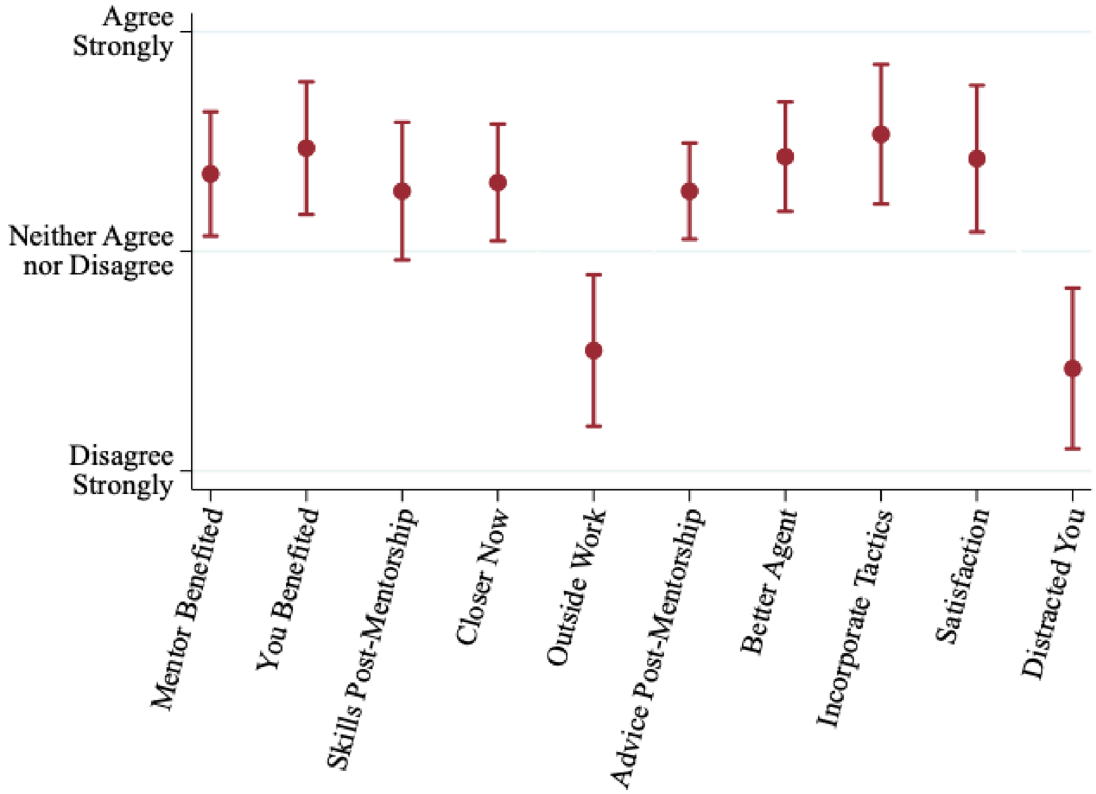
Notes. This figure plots the regression coefficients (and 95% confidence intervals) on *Mentored*, the sum of *Mentored* and *Mentored* × *Voluntary*, and *Voluntary Opt-Out* from Equation (1). We use ln(Revenue) as the dependent variable. The "Baseline" estimation replicates the result from Column (1) of Table 3. The second estimation excludes cohort fixed effects. The third includes date fixed effects. The fourth estimation includes controls for the agent's demographic characteristics: age, gender, and marital status. The fifth estimation includes additional controls for the agent's referral status, hiring score, previous call center experience, and previous sales experience. The sixth estimation layers on five more controls for the agent's personality traits: extroversion, agreeableness, conscientiousness, emotional stability, and openness. Note that these controls are frequently missing for non-mentored agents in the Mandatory-Condition, suggesting these are not good controls for estimating the average treatment effect because the data come from a selected sample. The seventh estimation removes observations in which agents are no longer working in the division in which they were initially hired.

Figure A.4: Relation Between Hiring Score and Meeting Completion Numbers



Notes. This figure is a binned scattreplot of the relationship between agents' hiring scores and the number of mentor-protégé meetings completed by mentored agents in the Mandatory-Condition.

Figure A.5: Responses to Post-Mentorship Survey



Notes. This figure plots the average values (and 95% confidence intervals) for responses to the post-mentorship survey questions. All responses were made on a scale from -3 to 3, with -3 indicating "Disagree Strongly," 0 indicating "Neither Agree nor Disagree," and 3 indicating "Agree Strongly." The statements, from left to right, are as follows: "Your mentor benefited from the mentoring relationship"; "You benefited from the mentoring relationship"; "Since your formal meetings have ended, your mentor has continued to teach you skills to help you make more sales"; "You and your mentor are closer now than you were during the mentor program"; "Since your formal meetings have ended, you and your mentor have spent time together outside of the office"; "Since your formal meetings have ended, you have reached out to your mentor for help/advice"; "You have become a better sales agent as the result of being mentored"; "Being mentored helped you incorporate important selling tactics into your day-to-day work"; "Having a mentor increased your day-to-day satisfaction at work"; "Being mentored distracted you from reaching your potential each week." Seventeen protégés completed the post-mentorship survey.

Table A.1: Balance in Division Performance

	Mandatory-Condition	Voluntary-Condition	p-value
	(1)	(2)	(2)−(1)
Revenue			
Mean	760.27	793.95	0.386
Std Dev.	(112.95)	(153.25)	
RPC			
Mean	46.80	49.27	0.361
Std Dev.	(8.42)	(10.40)	
RPH			
Mean	115.61	120.21	0.387
Std Dev.	(13.69)	(21.83)	
Calls			
Mean	17.18	17.18	0.996
Std Dev.	(0.98)	(1.32)	
Hours			
Mean	6.57	6.59	0.839
Std Dev.	(0.39)	(0.39)	
Adherence			
Mean	0.82	0.84	0.283
Std Dev.	(0.04)	(0.04)	
Conversion			
Mean	0.23	0.22	0.375
Std Dev.	(0.03)	(0.03)	
Number of Cohorts	22	31	

Notes. In this table, we take average productivity measures of agents who were not mentorship eligible within each sales division. Cohorts are assigned to a particular sales division, so the tests estimate the balance in brand-level productivity measures between cohorts in the Mandatory-Condition versus those in the Voluntary-Condition. We report standard deviations in parentheses, and we report p-values from difference in means tests to compare values across the different treatment conditions.

Table A.2: Balance in Mentor Demographics

	Mandatory-Condition	Voluntary-Condition	p-value
	(1)	(2)	(2)−(1)
Mentor Age (yrs.)			
Mean	24.06	23.68	0.516
Std Dev.	(4.66)	(4.44)	
Mentor Woman			
Mean	0.28	0.22	0.278
Std Dev.	(0.45)	(0.42)	
Mentor Married			
Mean	0.08	0.20	0.010
Std Dev.	(0.27)	(0.40)	
Mentor Tenure			
Mean	1.32	1.21	0.390
Std Dev.	(0.83)	(1.07)	
Number of Protégés	114	123	

Notes. In this table we report average characteristics of the agents who mentored protégés in the Mandatory-Condition in Column (1) and of the agents who mentored protégés in the Voluntary-Condition in Column (2). Mentors were not designated exclusively to either of the mentoring conditions. In other words, a mentor's first protégé could have been assigned to the Mandatory-Condition, whereas their second protégé could have been assigned to the Voluntary-Condition. Mentors were never informed as to whether their protégés were in the Mandatory-Condition or the Voluntary-Condition. We report standard deviations in parentheses, and we report p-values from difference in means tests to compare values across the different treatment conditions.

Table A.3: Long-Term Treatment and Selection Effects of Mentoring

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	Tenure_3	Tenure_4
	(1)	(2)	(3)	(4)	(5)	(6)
Mentored	0.162	0.083	0.107	0.005	0.009	0.031
standard errors	(0.106)	(0.064)	(0.094)	(0.007)	(0.070)	(0.071)
sharpened q-value	[0.240]	[0.345]	[0.403]	[0.718]	[0.761]	[0.723]
Mentored × Voluntary	-0.126	-0.103	-0.084	0.004	-0.013	-0.019
standard errors	(0.135)	(0.084)	(0.114)	(0.010)	(0.095)	(0.097)
sharpened q-value	[0.463]	[0.359]	[0.638]	[0.723]	[0.761]	[0.761]
Voluntary Opt-Out	-0.043	0.013	0.023	-0.012	-0.252***	-0.212**
standard errors	(0.101)	(0.042)	(0.080)	(0.013)	(0.090)	(0.081)
sharpened q-value	[0.723]	[0.761]	[0.761]	[0.463]	[0.060]	[0.060]
Cohort FE	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.049	0.058	0.058	0.083	0.007	0.003
Observations	13,075	13,075	13,075	13,075	603	603

Notes. The sample used in Columns (1)–(4) is composed of agent-day productivity data for all mentor-eligible agents with post-training productivity data. The data covers agents' productivity on their third to sixth months on the job after they complete training. *Mentored* equals one for agents who were randomized to received an available mentor, and zero otherwise, *Voluntary* equals one for agents in the Voluntary-Condition, and zero otherwise, and *Voluntary Opt-Out* equals one for agents who chose to opt out of possibly receiving a mentor, and zero otherwise. Columns (5)–(6) use data with a single observation per unique hired agent to capture retention effects. $Tenure_t$ equals one for agents who achieve at least t months of tenure at the firm, and zero otherwise. We estimate ordinary least squares regressions in all columns. All specifications include cohort fixed effects. Standard errors are clustered by cohort and are reported in parentheses. Sharpened q-values are presented in brackets, following Anderson (2008). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.4: Treatment and Selection Effects of Mentoring on Calls and Hours

Panel A: Calls and Hours in Months 1–2

	ln(Calls)	ln(Hours)	ln(Calls/Hour)
	(1)	(2)	(3)
Mentored	0.051	0.050***	0.001
	(0.033)	(0.018)	(0.025)
Mentored × Voluntary	-0.037	-0.050**	0.013
	(0.043)	(0.022)	(0.037)
Voluntary Opt-Out	-0.074	-0.075**	0.001
	(0.047)	(0.036)	(0.029)
Cohort FE	✓	✓	✓
Adj. R-Square	0.124	0.082	0.177
Observations	15,137	15,137	15,137

Panel B: Calls and Hours in Months 3–6

	ln(Calls)	ln(Hours)	ln(Calls/Hour)
	(1)	(2)	(3)
Mentored	0.057	0.040*	0.018
	(0.046)	(0.023)	(0.037)
Mentored × Voluntary	-0.015	-0.041	0.026
	(0.062)	(0.043)	(0.052)
Voluntary Opt-Out	-0.050	-0.068	0.017
	(0.072)	(0.059)	(0.055)
Cohort FE	✓	✓	✓
Adj. R-Square	0.108	0.063	0.200
Observations	13,075	13,075	13,075

Notes. The sample used in Columns (1)–(3) is composed of agent-day productivity data for all mentor-eligible agents with post-training productivity data. *Mentored* equals one for agents who received mentorship, and zero otherwise, *Voluntary* equals one for agents in the Voluntary-Condition, and zero otherwise, *Voluntary Opt-Out* equals one for agents who chose to opt out of possibly receiving a mentor, and zero otherwise, We estimate ordinary least squares regressions in all columns. All specifications include cohort fixed effects. Standard errors are clustered by cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.5: Treatment and Selection Effects of Mentoring (Pilot vs. Post-Pilot Cohorts)

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	$Tenure_1$	$Tenure_2$
	(1)	(2)	(3)	(4)	(5)	(6)
Mentored	0.197***	0.107***	0.121**	0.024**	0.121***	0.150*
	(0.069)	(0.034)	(0.055)	(0.012)	(0.029)	(0.086)
Mentored × Post	-0.096	-0.023	-0.059	-0.019	-0.063	-0.206
	(0.172)	(0.109)	(0.124)	(0.018)	(0.105)	(0.129)
Mentored × Voluntary	-0.256*	-0.172*	-0.174	-0.017	-0.061	-0.218**
	(0.144)	(0.100)	(0.128)	(0.017)	(0.076)	(0.106)
Mentored × Voluntary × Post	0.121	0.046	0.076	0.015	0.093	0.289*
	(0.228)	(0.152)	(0.184)	(0.023)	(0.134)	(0.161)
Voluntary Opt-Out	-0.270**	-0.206***	-0.228***	-0.011	0.013	-0.328***
	(0.109)	(0.061)	(0.077)	(0.025)	(0.029)	(0.075)
Voluntary Opt-Out × Post	-0.126	0.023	0.023	0.008	-0.170	0.303**
	(0.229)	(0.131)	(0.178)	(0.028)	(0.113)	(0.143)
Cohort FE	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.029	0.046	0.028	0.118	0.073	0.074
Observations	15,137	15,137	15,137	15,137	603	603
No Differential Post Effects	0.864	0.992	0.969	0.569	0.455	0.074

Notes. The sample used in Columns (1)–(4) is composed of agent-day productivity data for all mentor-eligible agents with post-training productivity data. *Mentored* equals one for agents who were assigned a mentor, and zero otherwise, *Voluntary* equals one for agents in the Voluntary-Condition, and zero otherwise, *Voluntary Opt-Out* equals one for agents who chose to opt out of possibly receiving a mentor, and zero otherwise, and *Post* equals one for cohorts that entered the firm on or after May 27th (the post-pilot cohorts), and zero otherwise. We estimate ordinary least squares regressions in all columns. Columns (5)–(6) use data with a single observation per unique hired agent to capture retention effects. $Tenure_t$ equals one for agents who achieve at least t months of tenure at the firm, and zero otherwise. The bottom row reports p-values from post-estimation tests that the coefficients on *Mentored* × *Post*, *Mentored* × *Voluntary* × *Post*, and *Voluntary Opt-Out* × *Post* are jointly equal to zero. All specifications include cohort fixed effects. Standard errors are clustered by cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.6: Treatment Effects of Mentoring in the Voluntary-Condition With Covariate Rebalancing to Match the Mandatory-Condition Treated Sample

Panel A: Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	Tenure_1	Tenure_2
	(1)	(2)	(3)	(4)	(5)	(6)
Mentored	0.170**	0.100**	0.104**	0.018*	0.102***	0.087
	(0.066)	(0.038)	(0.050)	(0.009)	(0.037)	(0.071)
Mentored × Voluntary	-0.332*	-0.247**	-0.208	-0.009	0.098	0.094
	(0.168)	(0.111)	(0.143)	(0.020)	(0.109)	(0.122)
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.032	0.046	0.030	0.083	0.115	0.096
Observations	13,779	13,779	13,779	13,779	544	544

Panel B: Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	Tenure_3	Tenure_4
	(1)	(2)	(3)	(4)	(5)	(6)
Mentored	0.162	0.083	0.107	0.005	0.009	0.031
	(0.106)	(0.064)	(0.094)	(0.007)	(0.070)	(0.071)
Mentored × Voluntary	-0.165	-0.107	-0.156	0.015	0.199	0.116
	(0.147)	(0.094)	(0.128)	(0.015)	(0.158)	(0.141)
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.065	0.078	0.062	0.103	-0.003	0.015
Observations	12,270	12,270	12,270	12,270	544	544

Notes. This table reports weighted regressions such that the distribution of agents who opt into the program in the Voluntary-Condition matches the distribution of characteristics among mentored agents in the Mandatory-Condition. Agents who opt out are not included in the sample. Matched characteristics include demographics, Big 5 personality factors, and work experience. There is significant missing data on personality scores for non-mentored agents in the Mandatory-Condition, but nearly all Voluntary-Condition agents who opt in have personality data. Because treatment assignment is random, matching to the distribution of treated agents (where there is no missing data) holds constant the distribution of observables across the Voluntary- and Mandatory-Conditions. We use entropy balancing up to the third moment of covariate distribution for the reweighting procedure. We then regress productivity or retention variables on a *Mentored* indicator and an interaction of *Mentored x Voluntary* for the agents in the Voluntary-Condition. Dependent variables are defined in Table 3. All specifications include cohort fixed effects. Standard errors are clustered by cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.7: Meeting Completion Rates

	Mandatory-Condition	Voluntary-Condition	p-value
	(1)	(2)	
Number of Agents	114	123	
At Least One Recorded Meeting	110	114	
No Recorded Meeting	4	9	
Number Recorded Meetings (avg.)	2.58	2.27	0.073
	(1.43)	(1.22)	
Meeting Completion Ratio (avg.)	0.83	0.72	0.009
	(0.29)	(0.34)	

Notes. In this table we report the mentor meeting completion details of protégés in the Mandatory-Condition and the Voluntary-Condition. *No Recorded Meeting* indicates that there is no record that the mentor-protégé pair ever met with one another. The *Meeting Completion Ratio* measure is based on the number of possible meetings the mentor-protégé pair could have had. While the mentoring protocol called for one meeting per week for four weeks, there were instances in which either a mentor or protégé or both were absent from work for an extended period of time (e.g., on vacation), reducing the number of possible scheduled meetings from four to three (or fewer, in some cases). As such, the denominator of the meeting completion ratio is occasionally less than four.

Table A.8: Differences in Worksheet Content

	Total Words per Worksheet	Skill Words per Worksheet	Support Words per Worksheet	Other Words per Worksheet
	(1)	(2)	(3)	(4)
Mandatory-Condition	-0.023	0.253	0.145*	-0.422
	(2.686)	(0.311)	(0.079)	(2.584)
Adj. R-Square	-0.006	-0.002	0.013	-0.006
Observations	159	159	159	159
Mean DV	47.36	4.27	0.448	42.65

Notes. This table considers differences in worksheet content between protégés in the Mandatory-Condition and those in the Voluntary-Condition. For each worksheet, we identify the fraction of words in the responses that relate to job-specific skills or knowledge (*Skill*), those that relate to receiving support, encouragement, and friendship (*Support*), and those that are neither related to skill nor support (*Other*), which include stop words. These become the dependent variables in our regression specifications of worksheet content on mentorship type. Robust standard errors are reported in parentheses. The mean of the dependent variable is listed below the observation count line. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.9: Treatment Effect Estimates When a Mentorship Slot is the Unit of Analysis

Panel A: Months 1–2 with Imputed Replacement Productivity

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence
	(1)	(2)	(3)	(4)
Mentored	0.205**	0.115**	0.130*	0.026*
	(0.095)	(0.054)	(0.068)	(0.014)
Mentored × Voluntary	-0.197	-0.131*	-0.135	-0.011
	(0.125)	(0.073)	(0.095)	(0.016)
Voluntary Opt-Out	-0.255***	-0.142***	-0.171***	-0.010
	(0.078)	(0.042)	(0.058)	(0.014)
Cohort Fixed Effects	✓	✓	✓	✓
Adj. R-Square	0.036	0.030	0.035	0.109
Observations	22,840	22,840	22,840	22,840

Panel B: Months 3–6 with Imputed Replacement Productivity

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence
	(1)	(2)	(3)	(4)
Mentored	0.169*	0.089*	0.113	0.020
	(0.098)	(0.053)	(0.071)	(0.012)
Mentored × Voluntary	-0.077	-0.032	-0.031	-0.016
	(0.118)	(0.063)	(0.085)	(0.013)
Voluntary Opt-Out	-0.069	-0.013	-0.011	-0.005
	(0.060)	(0.029)	(0.044)	(0.006)
Cohort Fixed Effects	✓	✓	✓	
Adj. R-Square	0.061	0.059	0.071	0.083
Observations	39,022	39,022	39,022	39,022

Notes. The results in this table show estimates of the treatment effects of mentorship when a slot (i.e., an occupied position as a sales agent within the firm) is the unit of analysis regardless of the agent's tenure. For this analysis, we form a balanced panel of agents made up of the observed productivity of those who remain at the firm and imputed productivity of a replacement for agents who separate before the indicated time horizon. In Panel A, for mentor-eligible agents who leave the firm before the two-month mark, we extend the time series of their productivity provision to two months and replace their post-termination productivity values with the average productivity of a newly hired replacement agent. We do this by computing the average productivity of newly hired non-mentor eligible agents in the same location-division-year-quarter as the departed agent. In Panel B, we use the same productivity replacement procedure to extend the sample out to six months. We then re-estimate our main intention-to-treat regression models using Equation (1). Standard errors are clustered by cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.10: Lee Bounds Estimates of the Effect of Mentoring on Productivity

Panel A: Mandatory-Condition, Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence
	(1)	(2)	(3)	(4)
Mentored_{lower}	0.084**	-0.002	-0.002	0.006*
	(0.040)	(0.028)	(0.031)	(0.003)
Mentored_{upper}	0.464***	0.273***	0.299***	0.037***
	(0.035)	(0.028)	(0.029)	(0.003)
Observations	7,819	7,819	7,819	7,819

Panel B: Mandatory-Condition, Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence
	(1)	(2)	(3)	(4)
Mentored_{lower}	0.129***	0.040	0.048*	-0.006**
	(0.036)	(0.026)	(0.028)	(0.003)
Mentored_{upper}	0.355***	0.214***	0.229***	0.010***
	(0.034)	(0.025)	(0.027)	(0.002)
Observations	6,592	6,592	6,592	6,592

Panel C: Voluntary-Condition, Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence
	(1)	(2)	(3)	(4)
Mentored_{lower}	-0.045	-0.048	-0.053	0.002
	(0.058)	(0.035)	(0.042)	(0.006)
Mentored_{upper}	-0.030	-0.036	-0.042	0.004
	(0.038)	(0.028)	(0.030)	(0.003)
Observations	7,756	7,756	7,756	7,756

Panel D: Voluntary-Condition, Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence
	(1)	(2)	(3)	(4)
Mentored_{lower}	0.023	-0.037	-0.004	-0.005*
	(0.060)	(0.036)	(0.044)	(0.003)
Mentored_{upper}	0.190***	0.073***	0.123***	0.006***
	(0.037)	(0.026)	(0.029)	(0.002)
Observations	6,872	6,872	6,872	6,872

Notes. Panels A and B use agent-day productivity data for agents in the Mandatory-Condition. Panels C and D use agent-day productivity data for agents in the Voluntary-Condition, excluding those agents who opt out of the program. We estimate treatment effect bounds that account for non-random attrition as proposed by Lee (2009). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

B Documentation of Instructions to Mentors and Example Worksheet for Structuring Conversations

Mentor Instructions

What is a Mentor?

In *The Odyssey*, Odysseus prepared to fight in the Trojan War. Before leaving home to fight in the war, he asked his trustworthy friend, named Mentor, to train and educate his son, Telemachus. Similarly, mentors today are meant to train and educate their protégés. Management at has chosen you to be a mentor---a source of further skill development---for newly hired sales agents. You have been selected specifically because you've demonstrated a willingness to teach other sales agents and help them become a successful and productive sales agent.

The responsibility to mentor a newly hired sales agent should not be taken lightly. Management strongly believes new agents will benefit from the additional training and the insider knowledge received as a result of being mentored by a talented, more seasoned agent. Because of this, has devoted significant resources to give mentors and protégés the best opportunity to spend productive time together, so please take your mentorship responsibilities seriously.

What will You Do as a Mentor?

As a mentor, you will do the following:

1. You will meet with your protégé at **least once a week**.
    a. Before meeting, your protégé will complete the Protégé Worksheet.
        i. If he/she has not completed it, you will kindly help him/her do so.
    b. During your meeting, you and your protégé will discuss his/her responses. You should also take this time to do the following:
        i. Impart knowledge and skill by explaining, giving useful examples, and demonstrating processes, and asking thought-provoking questions.
        ii. Discuss actions you've taken to become a successful sales agent.
        iii. Provide him/her with any tips and sales tactics that help you overcome customer concerns and that help you up-sell to better services.
        iv. Practice the designated sales protocol with them and help them gain a strong understand of the products, services, and bundles available.

2. After meeting with your protégé, you will deliver the finished worksheet to .
    a. will initial and timestamp the worksheet and make a record that you completed your weekly meeting responsibility.

3. Every two weeks, you will be asked to complete an on-line survey.
    a. These questions are meant to gauge the progress of your protégé and the overall benefit of the mentoring relationship.
    b. Please answer these questions honestly, as they are not meant to punish but, instead, to help assess the effectiveness of the mentorship program.

Protégé Worksheet (Week 1)

Protégé: ____
Mentor: _____________________________________ Number of times mentor has reached out: _____
Date: ____

Weekly Self-Reflection:

What are **your expectations** regarding your sales ability? Does your mentor know this?

What may **prevent you from** having a successful first week? Does your mentor know this?

Think of the **MOST** successful call you had recently. What made it **successful**?

Think of the **LEAST** successful call you had recently. What made it **unsuccessful**?

Weekly Goal:
What **ONE goal** are you setting for yourself for this coming week?

What will you do to **reach this goal**? Have you told your mentor about this goal? ___________
____

For Mentors to Respond:
How have you, as a mentor, been a source of **skill development** for your protégé? What have you done so far to help him/her succeed on the sales floor here at ?

Protégé's Initials	Mentor's Initials	Intern's Initials & Timestamp

C Worksheet Response Examples

Panel A: Think of the most successful call you had recently. What made it successful?

Skill	· I pitched TV really well · Having different examples of pitches from my coach to fall back on
Support	· I was confident and tried to connect · The person I spoke with was very nice

Panel B: Think of the least successful call you had recently. What made it unsuccessful?

Skill	· Customer didn't want to pay the deposit, [I] didn't rebuttal · Not doing call flow, not caring, not enough discover
Support	· Not being confident in my ability to rebuttal · The person was rude and wanted me fired

Panel C: What will you do to reach this goal? Have you told your mentor about this goal?

Skill	· I'll follow the call flow · I will create better pitches · Be better with the triple play, use what [the] mentor told [me] · My mentor is going to help me pitch DTV by giving me her tips on what helped her · Practice on every unserviceable call · Try upsell technique
Support	· [Goal to achieve] 1500 a day, build confidence in it · Be more positive · Stay positive · Stay in communication with [my coach] · Motivations - self discipline · Check in with my coach and be confident

Panel D: Words Associated with Sales Skills and Knowledge

Adherence, Conversion, Customer, Direct, Dish, Double, DPI, DTV, Internet, Knowledge, Phone, Pitch, Price, Pricing, Process, Revenue, RPC, RPH, Sale, Security, Sell, Skill, Sold, System, Television, Triple, TV

Panel E: Words Associated with Receiving Support

Annoy, Breath, Confidence, Confident, Cool, Encourage, Encouraging, Friend, Introduce, Kind, Laugh, Mean, Motivate, Motivation, Nice, Patience, Patient, Positive, Rude, Social, Support, Welcome, Welcoming

D AEA Pre-Registration Text

Here we replicate the AEA pre-registration text. Differences between the AEA pre-registration and our actual implementation are denoted in footnotes.

D.1 Abstract

Mentoring is increasingly encouraged in workplaces, and a number of firms have implemented formal programs. While a growing body of research suggests that mentoring relationships benefit those being mentored (protégés), there is scant evidence to delineate whether these favorable outcomes are driven by the mentoring experience on average, by the self-selection of protégés into mentoring who anticipate having the largest gains (selection based on gains), or by the self-selection of protégés who would have performed well in the absence of mentoring (selection based on levels). We use a field experiment to evaluate a workplace mentoring program inside a large sales organization.

Experienced employees opt-in as mentors, and new hires are slated as potential protégés. The project objective is to study the mentoring consequences across protégés who actively elect to be formally mentored relative to those who are randomly allocated a mentor. We estimate treatment effects on sales productivity and turnover for those who select into mentoring and for those who opt out.

D.2 Intervention(s)

We analyze the effectiveness of a workplace mentoring program where employees opt-into mentoring or are randomly assigned a mentor. More details are provided in the design field.

D.2.1 Intervention Start Date

2019-05-27

D.2.2 Intervention End Date

2019-12-20

D.3 Primary Outcomes (end points)

Log revenue-per-call (RPC), an indicator for worker turnover, log completed tenure, the firm's internal adherence to schedule measure (e.g. time spent working whilst at work), and the firm's internal engagement metrics (online surveys asking for willingness to recommend employment at firm, comfort with leadership, etc.).

D.3.1 Primary Outcomes (explanation)

Agent's weekly RPC is a measure of sales productivity that removes demand variation outside of the worker's control. RPC is the primary productivity measure used by the firm, combining both agent's firm-specific knowledge and their individual effort.[36] Worker turnover measures whether the interventions changed the agents' propensity to leave the firm. Log of completed tenure is a different measure of retention that

[36] RPC was the primary endpoint based on our experience analyzing the productivity of veteran agents within the firm (Sandvik et al., 2020), but total revenue picks up different margins of adjustment for new agents, which is why we report both metrics, along with revenue-per-hour (RPH).

has been used in the prior literature and the attendance measure provides an adjacent measure of agent effort. Finally, engagement measures are hypothesized to be forward looking measures of productivity.

D.4 Experimental Design Details

Seasoned sales agents are invited to apply as internal mentors to incoming recruits (the firm "qualifies" mentors as having sufficient sales experience). New mentorship opportunities are periodically announced, and prior mentors are permitted to re-enter the mentor pool. The firm communicates that serving as a mentor is a useful first step to being considered for a managerial position. New mentors complete a survey asking them about their personality, interests, work preferences, and values. Mentors are randomly assigned with probability 50% to receive a set of instructions emphasizing that mentoring is about teaching protégés how to do the job. The remaining mentors receive instructions emphasizing that mentoring is about providing protégés support. Sales agents are hired in batches (cohorts). Newly hired sales agents complete two weeks of training, primarily in a classroom or listening in on other agents' sales calls. New agents then complete the same personality and preference survey that mentors take. At the end of their two-week training, each cohort of agents is eligible for randomization into a mentoring treatment arm. Any mentoring relationship commences as soon as the agent completes their training.

The randomization procedure is as follows:

D.4.1 Cohort Level Randomization

The initial level of randomization is cohorts of new hires (potential protégés). Each cohort (a group of new hires who are joining the firm at the same time, are in the same training group, and will be working in the same sales division and office location) will be randomized into one of two conditions: Mandatory-Condition or Voluntary-Condition. 40% of the cohorts will be in the Mandatory-Condition group and 60% of the cohorts will be in the Voluntary-Condition group.

D.4.2 Within Cohort Randomization

For cohorts in Mandatory-Condition, new hires will receive a mentor with probability 50%. This will be communicated privately between sales floor staff and the individual workers. Agents in the Mandatory-Condition who do not receive a mentor will not receive communication regarding the program. For cohorts in Voluntary-Condition, sales floor staff verbally explain the firm's mentorship program, answer questions, and provide each agent a confidential ballot where they can decide whether or not to enter a lottery which randomly determines whether the agent is allocated a randomly assigned mentor, or no mentor at all. Of the agents who enter the lottery, approximately 50% will be assigned a mentor. Agents who choose not to be mentored will never be assigned a mentor.

D.5 Compliance Tracking

The firm's training staff will track whether mentors and protégés meet. This tracking will be aided by worksheets. Upon completion of the worksheets, the firm will reward "kudos" points that can be accumulated to purchase items from the company store. As mentioned earlier, mentors may participate more than once, however they will never have more than one protégé at a time.[37] Eligible protégés and mentors will each take an electronic survey at the end of the formal program. The survey for protégés will ask about the protégé's initial excitement when told about the mentoring program, their perceived engagement with their mentor, and an estimate of the effectiveness of mentoring. This question will be phrased as: "What was your average RPC last week? What do you think your average RPC would have been had you not been working with a mentor?" The survey for mentors will ask about the protégé's enthusiasm for the mentorship program and an estimate of the mentor's perceived treatment effect on the protégé. This question will be phrased as: "If your protégé had not received mentoring, his/her RPC would have been [40% lower — slider — 40% higher].[38] Note that numbers greater than zero mean that mentoring was not effective for improving protégé performance. Please be candid, as your responses will not be shared with management."

D.6 Edit June 4, 2019

To assess the potential for spillovers, we have revisited the design in consultation with the company such that there will be "hold out" cohorts for one division-office who never receive mentoring. Any cohorts/individuals who are switching brands also will be held-out. Work-from-Home cohorts will also present a possible "hold out" group for comparison and all cohorts in a smaller third office (which no longer exist, but for whom historical data is available) were "hold out" cohorts who knew nothing about mentoring. A "sentiment survey" will be administered to all agents in their 5th week on the sales floor.[39] This will be one week after mentored agents finish hiring. We will gather information on their feelings towards the onbaording process and ask questions, common in the literature, to solicit their sentiment towards the firm, their perceptions of their ability, their enthusiasm about the job, etc. We will use this survey to test for spillovers based on survey responses.

D.7 Randomization Method

Randomization done by computer. Participants will be informed if randomized in.

[37] As the program progressed, the internal mentoring staff felt that many of the mentors could effectively mentor multiple protégés as once. As a result, we adjusted the protocol such that it was possible for a single individual to mentor multiple new hires concurrently, but mentor-protégé pairs always met individually, meaning the protocol was the exact same from the point of view of the protégé.

[38] The post-mentorship survey completion rates of mentors and protégés were very poor, so we do not have meaningful data for this question. Anecdotally, the average responses of both sets of individuals suggests that protégés' RPC would have been lower in the absence of mentorship, but the inference is not precise.

[39] We were not able to administer this survey. The firm had several of its own survey initiatives occurring simultaneously, so additional surveys connected to the mentorship program were not conducted due to the concern of "survey fatigue" among the sales agents.

D.8 Randomization Unit

Clustered randomization of cohorts in a first level, with individual randomization within the cohort. See design details.

D.9 Was the treatment clustered?

Yes

D.10 Sample size: planned number of clusters

The exact sample size is stochastic and depends on the firm's actual hiring. We have 46 planned clusters.

D.11 Sample size: planned number of observations

In one office, the firm has projected 269 new hires in 22 cohorts. There are 350 new hires in 24 cohorts projected in the second office.

D.12 Sample size (or number of clusters) by treatment arms

Please see design field.

D.13 Minimum detectable effect size for main outcomes

Using pre-intervention data to estimate the intra-class correlation coefficient and residual variation, the minimum detectable effect size for log RPC between those randomized into and out of mentoring is 0.07 (accounting for sample design and clustering).

D.14 Analysis Plan

The Treatment Effect of Mentoring on those who opt in is:

$$\beta_{OptInMentor} = mean(Y_{OptInMentor}) - mean(Y_{OptInNoMentor}).$$

We will estimate this mean difference using a regression of Y on an indicator for receiving a mentor along with cohort fixed effects and indicators for the type of instructions mentors receive.[40] The sample will be the workers in the voluntary treatment cohorts who opt into mentoring.

The Treatment Effect of Mentoring on those who opt out can then be derived by writing the average gain from mentoring in the population as:

$$mean(Y_{RandomMentor}) - mean(Y_{NoMentor}) = \beta_{OptInMentor}\pi_{OptIn} + \beta_{OptOutMentor}\pi_{OptOut}.$$

The β parameters are the heterogeneous treatment effects and the π are the population fraction who opt in and opt out. This yields:

$$\beta_{OptOutMentor} = [mean(Y_{RandomMentor}) - mean(Y_{NoMentor}) - \beta_{OptInMentor}\pi_{OptIn}]/\pi_{OptOut},$$

[40] Mentors received instructions that either put more emphasis on the supportive nature of the program or the skills-building nature of the program. We detect no differences between instruction type. Because of this and for brevity, we omit this indicator from the models in our heterogeneous treatment effects tests.

where the difference in means is net of cohort fixed effects and indicators for mentoring instruction type. The population average treatment effect (ATE) of mentoring can be estimated from a regression of Y on a dummy for receiving a mentor and cohort fixed effects in cohorts that have (entirely) randomly assigned mentoring. This yields:

$\beta_{OptOutMentor} = [ATE - \beta_{OptInMentor}\pi_{OptIn}]/\pi_{OptOut}$.

Inference for $\beta_{OptOutMentor}$ will come from block bootstrapping the statistic. Selection bias will be measured among voluntary treatment cohorts as the regression analogue of:

$mean(Y_{OptInNoMentor} - mean(Y_{OptOut})$,

where the means are net of cohort fixed effects. This procedure allows us to estimate sales productivity differences among protégés who opt into mentoring and those who do not. We use the sample of agents in the voluntary cohorts who did not receive a mentor. We regress Y on an indicator that the agent opted into mentoring along with cohort fixed effects and their mentor instruction-type fixed effects. Other regressions will look at opt-in as a function of early sales and demographic characteristics (gender, age, office location) and past experience (prior sales or call center experience).

We plan to validate these estimates using the electronic survey responses collected after the protégé graduates from the formal mentoring program, approximately 4 weeks following the initial onboarding instruction (e.g. how to use the systems, enroll for benefits, etc.).[41] We will compare average perceived gains from mentors and protégés to the actual estimated treatment effects across different assignment conditions. We will then assess whether the effectiveness of the mentoring pair differs based on characteristics of the mentor and protégé. We will regress protégé sales on fully saturated interactions of demographic characteristics for the mentor-protégé pair (old/young based on coarse buckets; gender) as well as similarity in survey responses on the intake survey.[42]

Finally, to assess whether mentoring detracts from–or improves sales–for the mentor, we will regress mentor log RPC and other sales measures on indicators demarking whether the mentor is eligible to mentor but has not yet done so, whether they have previously mentored in the program, or whether they are actively mentoring a protégé. This regression will include mentor fixed effects and mentor tenure.[43]

[41] As mentioned earlier, we were not able to administer this survey.

[42] This is a very high-dimensional exercise, and the most interesting potential differences (like females mentoring females) ended up having very small cell sizes, reducing statistical power for these tests.

[43] Tests that compare the characteristics of mentors and protégés, and those that look at the impact of mentorship on mentor productivity, are likely to be discussed in a separate article.