NBER WORKING PAPER SERIES

TREATMENT AND SELECTION EFFECTS OF FORMAL WORKPLACE MENTORSHIP PROGRAMS

Jason Sandvik
Richard Saouma
Nathan Seegert
Christopher T. Stanton

Working Paper 29148
http://www.nber.org/papers/w29148

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2021

We thank Emily Beam, Jasmijn Bol, Zoe Cullen, Guido Friebel, Robert Garlick, Jessica Hoel, Mitch Hoffman, Lisa LaViers, John List, Robert Metcalfe, Harish Sujan, Jason Snyder, and seminar participants at Harvard Business School for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Treatment and Selection Effects of Formal Workplace Mentorship Programs
Jason Sandvik, Richard Saouma, Nathan Seegert, and Christopher T. Stanton
NBER Working Paper No. 29148
August 2021
JEL No. J24,L23,L84,M5,M53

ABSTRACT

While formal mentorship programs are ubiquitous, less is known about who gains from receiving mentorship. In this paper, we report the outcome of a Randomized Controlled Trial (RCT) carried out in a US-based inbound sales call center where one branch of the experiment assigned a random subset of new hires to mentors (Broad-Mentoring), whereas a second branch (Selective-Mentoring) gave new hires the opportunity to opt into a mentoring relationship before assigning a random subset to mentors. In the Broad-Mentoring branch, mentored sales agents outperformed non-mentored agents by over 18% in the first six months on the job. Among agents who opt into the program in the Selective-Mentoring branch, those who received mentorship had negligible performance gains. The differences between the two branches indicates that formal mentorship program treatment effects are largest for workers who would otherwise opt out of these programs. Demographic and personality characteristics are relatively weak predictors of selection into the program, suggesting broad-based programs are likely more effective than alternative targeting rules.

Jason Sandvik
Department of Finance
Tulane University
jason.sandvik@gmail.com

Richard Saouma
Eli Broad School of Business
Michigan State University
rs2@msu.edu

Nathan Seegert
University of Utah
Department of Finance
Spencer Fox Eccles Business Bldg. Room 1113
1655 East Campus Center Drive
Salt Lake City, Utah 84112
nathan.seegert@business.utah.edu

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

1 Introduction

The most rapid increases in productivity often come at the beginning of an employee's tenure with a new firm (Shaw and Lazear, 2008). This period usually coincides with extensive on-the-job training, apprenticeships, or mentoring to augment employees' general and firm-specific human capital (Fudenberg and Rayo, 2019). Productivity growth, however, is heterogeneous across employees (Kahn and Lange, 2014), which may explain part of the previously documented, large, persistent productivity differences across people doing similar jobs (Mas and Moretti, 2009; Lazear et al., 2015; Sandvik et al., 2020, 2021). Earlier work has studied the role of management in perpetuating productivity differences across firms and nations (Syverson, 2011; Gibbons and Henderson, 2012; Bloom and Van Reenen, 2007), but additional work is needed to understand how management practices within firms influence intra-firm productivity dispersion.[1] In this paper, we investigate whether formal workplace mentorship influences the productivity and retention of newly hired employees in an environment where workers earn commissions based on their own output.

We define formal workplace mentorship as a structured interaction between a less experienced employee (protégé) and a more experienced coworker (mentor) in which the latter helps the former to (potentially) become more productive. Managers impose structure on formal mentoring relationships—e.g., protégé-mentor pairing, meeting frequency, discussion topics, etc.—which differentiates them from informal mentoring relationships, wherein one individual provides another with professional guidance outside of formal arrangements and without adhering to a set protocol. Formal mentorship programs have long-been of interest both to academics (Payne and Huffman, 2005; Mills and Mullins, 2008; Lyle and Smith, 2014; Porter and Serra, 2020; Ginther et al., 2020) and practitioners, as over 70% of Fortune 500 companies report that they provide their employees with mentorship opportunities (Gutner, 2009).[2] However, the causal evidence in support of such practices remains thin, as espoused by Allen et al. (2017) in a recent overview of the literature on mentoring. Given the prevalence of formal workplace mentorship programs, our first question is

[1] A striking fact from World Management Survey data is that effective human capital management/training is positively correlated with the presence of high-powered incentives within firms, potentially indicating that incentives and training are complementary or that some employees need formal programs to adapt to their jobs because incentives are insufficient. For a discussion of the data, see Bloom et al. (2014).

[2] We use the terms "formal mentorship" and "formal workplace mentorship" interchangeably throughout.

whether they have positive causal effects on worker productivity and retention.

If we find that formal workplace mentorship programs are effective *writ large*, then the next question is how best to allocate mentorship resources. In this paper, we focus on whether employees should have a say in receiving formal workplace mentorship, or whether it should be decided for them. On the one hand, workers may know best whether they stand to benefit from mentorship, so allowing them to self-select into such programs may be optimal. On the other hand, certain (potentially social) frictions may prevent the most needful workers from asking for and receiving help. For example, ambitious new-hires—who would perform well regardless of mentorship—may crowd out weaker (less ambitious) workers. The same outcome occurs if the workers most in need of assistance assign a negative stigma to asking for help (i.e., signalling incompetence, Edmondson and Lei (2014); Chandrasekhar et al. (2016); Bol and Leiby (2018)). These examples highlight the importance of delineating selection effects from treatment effects (Imbens and Wooldridge, 2009; Lee, 2009), which has largely contributed to the lack of precisely identified evidence for mentorship to date.

We use a Randomized Controlled Trial (RCT) with two top-level branches to measure both the treatment *and* selection effects associated with a formal workplace mentorship program involving newly hired employees in a large, US-based, inbound-sales call center. In the first branch, labeled Broad-Mentoring, agents were randomized between two subsequent treatment conditions. In condition 1a, newly hired sales agents were matched to a randomly chosen mentor drawn from a pool of established, non-supervisory sales agents at the firm. Agents in condition 1b were not matched with a mentor. In the second branch, labeled Selective-Mentoring, newly hired agents were allowed to opt in or out of the mentorship program, leading to three subsequent treatment conditions. Before presenting new agents with the choice to opt in or out, management explained that only a random subset of workers who opted in would ultimately be matched to a mentor due to the limited supply of mentors.[3] Agents that opted into the mentorship program were randomly chosen to receive a mentor or not, conditions 2a and 2b, respectively, following the same procedure as conditions 1a and 1b. Condition 2c consisted of agents that opted out of the program (i.e., we did not impose a

[3] The option to opt in or out of mentorship was presented to agents on their first day of training, before they began fielding calls from actual customers.

mentor against their wishes).

Once matched with a mentor, all mentor-protégé pairs adhered to a structured four-week mentorship protocol involving a series of brief meetings where protégés were asked to share their written responses to work-related questions with their mentors, and mentors were asked to provide feedback on the responses before submitting the worksheets to the firm's staff. This protocol had the potential to improve worker productivity in at least three ways: (1) through self-reflection via worksheet completion; (2) from the advice shared and support given by the mentor; and (3) by increasing access to future help and support from an experienced sales agent. Taken together, our field experiment setting is well-suited to measure formal mentorship effects for several reasons: the partnering firm features hundreds of agents who autonomously perform the same role (i.e., an agent's production is independent of the production of others); workers field thousands of homogeneous, randomly allocated inbound sales calls from prospective customers; we have access to highly granular, agent-level performance data capturing both extensive margin effort (e.g., up-time) and outcomes (e.g., revenue generation, selling efficiency, and attrition); and we controlled the randomization of agents into the different branches and conditions of the experiment. Section 2 highlights our firm setting, and Section 3 details the experimental design.

We present our results in Section 4, beginning with the main finding from the first branch of our study: agents receiving formal mentorship in the Broad-Mentoring branch generate 19% more daily revenue relative to non-mentored agents during their first two months on the sales floor. The agents' revenue growth comes from two sources (1) increased selling efficiency (higher revenue-per-call (RPC) and revenue-per-hour (RPH)) and (2) increased schedule adherence or up-time, resulting in a greater number of calls answered. This effect is persistent at longer time horizons, with over 90% of these early gains spanning agents' first six months on the job, albeit the estimates become less precise with time. We also find that mentorship improves retention. Specifically, agents receiving formal mentorship are significantly more likely than non-mentored agents to remain with the firm in the first 30 days (where attrition rates are traditionally highest). We do not estimate significant effects of mentorship on long-term retention, however. The benefits to retention from mentorship do not explain the productivity gains previously discussed, as our main findings hold when accounting for non-random attrition using Lee (2009) bounds. Taken together, the first branch of the experiment shows that formal workplace mentorship increased the productivity and retention of newly hired sales agents. These results provide clear evidence of the benefits of mentorship, something that has been elusive in observational data.

We use the second branch, Selective-Mentoring, to see if the effectiveness of mentorship programs can be enhanced with self-selection into the program and to assess how failure to account for self-selection might affect inference. Comparing agents in condition 2a, who opted in and received a mentor, to those in 2b, who opted in and did not receive a mentor, captures the treatment effect of mentoring for agents that opt in. Comparing agents in condition 2b to those in 2c, who opted out of the program, captures the selection effect. In most settings, researchers only observe the difference between conditions 2a and 2c, comparing those that opt in and receive a mentor to those that opt out. This difference, however, conflates the selection and heterogeneous treatment effects. We unpack that difference using condition 2b, composed of agents who opted into the mentorship program but randomly did not receive a mentor.

The Selective-Mentoring branch shows that better agents opt into mentorship but they accrue relatively little benefit from mentorship. Simply put, we find a large selection effect as agents in condition 2b (those who opt in but do not receive a mentor) generate revenue that is over 30% greater than agents in condition 2c (those who opt out). This selection effect suggests that a traditional comparison between mentored and non-mentored agents (i.e., conditions 2a vs. 2c) might vastly *overstate* the effect of mentorship. We also find large treatment effect heterogeneity, which is identified in part by comparing gains from mentorship in the Broad-Mentoring branch to those in the Selective-Mentoring branch. In contrast to the large gains when mentorship is unconditionally, randomly assigned in Broad-Mentoring, agents that opted in and received a mentor in the Selective-Mentoring branch were no more productive than agents that opted in and did not receive a mentor (i.e., conditions 2a vs. 2b). This heterogeneous treatment effect suggests that studies estimating a local treatment effect among agents that opt in would *understate* the effect of mentorship broadly. Using a pre-registered procedure to recover heterogeneous treatment effects, we estimate that the opt-out agents in Selective-Mentoring could have substantially increased their daily productivity had they received a mentor. Together, these findings suggest that the Broad-Mentoring intervention, which did not allow agents to self-select into program participation, generates larger aggregate benefits than the identically resourced Selective-Mentoring intervention.

Why is it then that the agents who opt out of receiving a mentor appear to be those that would benefit the most from mentorship? One might suspect that those who opted out overvalued their capabilities or were perhaps afraid of signalling weakness by opting in (Chandrasekhar et al., 2018). Leveraging the data available to us, we find little evidence that Big 5 personality scores (e.g., degrees of extroversion, agreeableness, conscientiousness, emotional stability, and openness) contribute to opting out, though we do find that low hiring scores—assigned to new hires by interviewers during the hiring process—are a strong predictor of the opt-out decision. That said, complementing hiring scores with demographic factors, personality characteristics, and previous work experience explains less than 7% of the variation in opt-out decisions across the entire Selective-Mentoring sample. This suggests that unobservable characteristics largely drive the decision to forgo mentoring. As a result, schemes that target mentoring based on demographics or other factors are likely to offer small improvements relative to broad-based mentorship programs, whereas schemes that encourage self-selection are likely to reduce, rather than enhance, the treatment effects of mentorship programs.

We have checked a number of different factors that may qualify these results. For example, we worked to limit the effect of spillovers in the design, cautioned the mentoring staff about the possibility of leakage/discouragement, and tested for its effect afterwards. Spillovers across agents could confound our estimates, and, as a result, we worked closely with the firm's internal mentoring staff to reduce the possibility that non-mentored agents: (i) became discouraged after not receiving a mentor or (ii) sought out internal mentors on their own. We test for spillover effects by leveraging hold-out groups of agents who were unaware of the mentorship treatments and data on agents hired before the mentorship program began. We find no evidence that the formal mentorship program in our design crowded out organic mentorship that would have occurred otherwise, nor do we find any evidence of discouragement or leakage of the mentorship content/curricula to other agents. We describe these tests and speak further to the robustness of our results in Section 5.

We perform additional tests in Section 6 to understand why the returns of Broad-Mentoring were greater than those of Selective-Mentoring and why the greatest gains to mentorship may accrue to those who opt out of mentorship eligibility. We find evidence that protégés in Broad-Mentoring spent more time with their mentors than did protégés in Selective-Mentoring, but additional tests show that this difference in face-time with one's mentor does not fully explain the greater returns to Broad-Mentoring. We do not find meaningful differences in the worksheet content of agents in the two different mentoring branches. Similarly, if we split Broad-Mentoring protégés into two groups based on their opt-out propensity scores, we find no differences in mentor meeting completion rates or worksheet content. Responses from wrap-surveys completed several weeks after the four-week mentorship protocol ended suggest that one benefit to mentorship was that it provided protégés with a long-term source of both help and support.

Our findings contribute to a variety of literatures. First, our findings help to contextualize several disparate strands of the personnel economics and management practices literatures. The traditional focus in personnel economics has been on incentive provision within firms and the alignment of hiring practices given an incentive scheme (Oyer and Schaefer, 2011). But hiring is often noisy (Hoffman et al., 2017), and incentives that induce the majority of hires to invest in new skills are likely to be insufficient for some workers. As a result, many firms face questions about how to deal with lower performers, either through carefully chosen incentive schemes, curated hiring practices, or the use of mentorship or training to improve productivity. Our results indicate that what might look like a hiring mismatch (the lower tail of new recruits) can often be corrected with interventions like mentoring, and, at least in our context, these interventions can generate positive returns for the firm.[4] We add to recent work on management practices that use experiments within individual firms to determine which interventions work (see, for example, Gosnell et al. (2020) who study different practices among airline captains that complement classical studies on how social preferences interact with incentives (Bandiera et al., 2005, 2009)).

Second, we contribute to decades of research on the efficacy of formal workplace mentorship programs (Kram, 1988), and to industry specific studies, e.g., K–12 education (Rockoff, 2008), higher education (Ginther et al., 2020), and nursing (Mills and Mullins, 2008). Our research design allows us to measure both the treatment *and* selection effects associated with formal workplace mentorship—as opposed to merely controlling for selection concerns. This stands in contrast to

[4]The closest related work on training and potential mismatch is likely Hoffman and Burks (2020), who study how worker overconfidence allows firms to provide training because workers are excessively eager to invest in the job. Our results instead indicate that some workers appear to under-invest in seeking out help when it is available, and targeted interventions in settings where selection into training and the treatment effects from training are negatively correlated may substantially alter the productivity distribution.

the earlier literature which either studies mandatory mentorship programs—common in public and regulated sectors (e.g., nursing and K–12)—or voluntary mentorship programs (e.g., for-profit and higher education), which in our setting corresponds to comparisons between conditions 1a and 1b or between conditions 2a and 2c, respectively. Our findings also complement a body of scholarship emphasizing the importance of leadership (Lazear et al., 2015; Carter et al., 2019; Hoffman and Tadelis, 2021), RCTs involving external mentorship (e.g., consultants or advisors) (Bruhn et al., 2018; Chatterji et al., 2019), and human resource management (Bloom et al., 2015).

Finally, our findings speak to the broader literature on informal workplace learning; e.g., peer-effects (Mas and Moretti, 2009; Bandiera et al., 2013; Herbst and Mas, 2015) and managerial tutelage or "boss-effects" (Lazear et al., 2015). In most contexts, manipulating organizations to capture boss effects is thought to be easier than organizing to capture ephemeral peer effects (Carrell et al., 2013), but recent work has demonstrated that relatively simple—albeit uncommon—interventions can unlock substantial peer learning (Sandvik et al., 2020). We also find large returns associated with a relatively simple and more common managerial intervention, mentoring. However, unlike the extant literature, our research design allows us to concurrently evaluate whether employees themselves are best positioned to identify whether such tutelage would benefit them or not.[5] We find that providing workers with the choice to opt out of mentoring reduces the efficacy of the program. Taken together, these studies corroborate the notion that inducing peer interactions through simple managerial interventions can be a highly effective means for disseminating knowledge, and consequently, increasing performance. Whereas relying only on high-powered incentives and self-selection may inhibit the lower tail of under-performers from reaching their potential.

2 Firm Setting

The mentoring program occurred in an inbound-sales call center from January to December of 2019, with data collection on protégé performance continuing after the conclusion of the formal mentoring relationships. The firm markets and sells the services of several companies, most of which are

[5] An additional novelty with the present study is that we present a bright-line demarcation between who will provide information (mentors) and who will receive it (protégés), whereas Sandvik et al. (2020) randomly paired employees together and treated them as equals, with no designated roles within the pairings.

different television, phone, and internet providers. Sales agents answer incoming calls from potential customers and sell digital services according to the customer's 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 best to up-sell premium packages can be difficult for newly hired sales agents, which makes this setting one where mentorship may help employees acclimate to a new job. All work is individual, so it is possible that mentoring or peer effects matter differently in our setting relative to those featuring team production. Throughout the year, 603 newly hired sales agents across 53 hiring cohorts entered the firm and were eligible for mentorship.

When hired, sales agents begin a two-week training program, where they learn the sales process through lectures and by listening in on 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 calls on the sales floor. 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. Mentorship-eligible agents 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 granular, individual-level performance measures for each sales agent. A sales agent's productivity is independent of their 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), and 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, which is a function of their total revenue generation and their selling efficiency, based on RPC and RPH. An agent generates revenue through each sale they make and, at the end of the week, the total amount of revenue generated is multiplied by their commission rate. The commission rate increases in the agent's selling efficiency (RPC and RPH) relative to other agents in their same division.[6] Multiplying the agent's weekly revenue and

[6]In addition, each agent has a fixed number of calls audited each week, and their commission rate decreases if conduct violations are identified by the auditors.

commission rate determines the amount of commission pay they earn that week. Sales agents also earn an hourly wage, which is above the federal minimum wage and increases with tenure, and agents can earn occasional bonuses for doing well during temporary promotional periods.

3 Experimental Design

Cohorts of new agents join the firm on a rolling basis; some weeks two or more cohorts of new hires enter the firm, whereas in other weeks no new cohorts enter the firm. We randomly designated each cohort to either the Broad-Mentoring branch (probability 40%) or the Selective-Mentoring branch (probability 60%).[7] Of those in Broad-Mentoring, approximately half the agents were randomly given a mentor and half were not (conditional on the supply of mentors). Agents in the Selective-Mentoring group were given the option to opt in or opt out of mentoring. Those who opted out did not receive a mentor. Of those who opted in, approximately half were randomly given a mentor and half were not (conditional on the supply of mentors).

Except for randomization into treatment and analyses of survey and productivity data, the program was administered entirely by the firm's on-boarding staff members. To participants, the mentoring program would have appeared like a normal part of the environment, rather than one imposed by outside researchers. Because of this, the experiment is very close to a natural field experiment, based on the taxonomy of Harrison and List (2004). Because participants knew that outside researchers were analyzing their survey and productivity data, they were not entirely unaware of their involvement in an experiment. The main concerns with participants knowing they

[7] 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 Broad- versus Selective-Mentoring branch 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 were unable to extend the experiment to the Spring hiring season to makeup the shortfall because of the onset of COVID-19. When we test for characteristic balance between the pilot cohorts and cohorts arriving after the pre-registered intervention start date, we find no significant differences. While 25 pilot cohorts in the Broad- and Selective-Mentoring branches were not randomized across offices, this does little to affect our estimates because our pre-registered strategy calls for treatment effect estimation using randomization into mentoring within each cohort (i.e., cohort fixed effects). Within-cohort randomization was no different in the pilot period and the period after pre-registration. Our main estimates are similar but less precise if we exclude data from pilot cohorts. The pre-registration text is documented in Appendix F, where we note instances in which there were minor deviations between the pre-registration and the implementation.

are part of an experiment are (1) experimenter demand effects and (2) Hawthorne effects, but these are unlikely to be major concerns in our context. First, protégés were *not* told that the researchers' object of interest was the opt-out decision, and they were not told that some cohorts were given the opportunity to opt out, while others were not. Similarly, mentors were not aware that the opt-out decision was the main focus of the research, nor were they told whether their protégé was in Broad-Mentoring or Selective-Mentoring. As such, neither protégés nor mentors would have been able to ascertain what the experimenters demanded. Second, Hawthorne effects were not likely strong in our setting. Monitoring in this environment is ubiquitous, as the sales managers are constantly monitoring and providing feedback on sale agents' performance. So it is unlikely that agent behavior was impacted by the knowledge that outsider researchers—with whom they *never* interacted—were tracking their performance. Furthermore, the mentoring program did not change the degree of monitoring relative to the day-to-day performance tracking already in place at the firm.[8]

3.1 Identifying Mentors and Administering Allocations

The firm's internal mentoring staff announced to all incumbent sales agents that a mentoring program would be taking place, and they asked for individuals to volunteer to be mentors. In addition, the staff identified sales agents who they and the sales managers agreed would make good mentors. These agents were directly asked if they would like to participate. If the mentoring staff and sales managers felt a particular sales agent was not a good fit 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. This currency could be used to purchase snacks, drinks, accessories, apparel, and other goods. Second, incumbent sales agents were told that effective mentoring would become a key indicator for future promotion prospects. While not formally necessary for promotion into a managerial role, being a mentor helped agents demonstrate

[8]Note that subjects were asked to provide informed consent when responding to an intake survey. The intake 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.

their potential aptitude for managing a sales team down the line.

The timeline of the mentoring program for a given cohort was as follows. Each new cohort began their two-week training on the Monday of training-week one. All new hires were asked to complete a new hire survey on the first day of training, which asked agents questions about their personality traits, work styles, and previous work experiences (specifically, whether they had (a) previous experience working in a call center and/or (b) previous experience working in sales). The responses to these questions were later used to identify the characteristics of individuals who opted into versus opted out of mentoring. During training-week one, cohorts were randomly allocated to either Broad-Mentoring or Selective-Mentoring, and the mentoring staff was informed of this assignment. For cohorts in Broad-Mentoring, half (or fewer, depending on mentor availability) of the new hires were randomly allocated to receive a mentor and the remainder to not receive a mentor. There was a limited supply of mentors, and the mentoring staff told new hires that this was the reason only half (or fewer) of them would receive a mentor and that random allocation by a team of academics was the fairest way to distribute mentors. Those who received a mentor were randomly allocated a particular mentor from the pool of available mentors. To reduce any feelings of discouragement, those who did not receive a mentor were told that the company provides many other opportunities for new hires to receive help while on the sales floor (e.g., from managers, coworkers, and division leaders). The mentoring staff reported to the authors multiple times that they found no evidence of discouragement among the agents who did not receive a mentor, and we provide evidence consistent with this lack of discouragement in Section 5.1.

For cohorts in Selective-Mentoring, the mentoring staff described the mentoring program to the newly hired agents and asked them to either opt in or opt out of the program.[9] The agents were told that a randomly selected subset of those who opted in would receive a mentor.[10] Agents

[9]The staff were given latitude in introducing the program, albeit they were asked to kindly insure that the following statement was explicitly read to the new workers: "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 week 1 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."

[10]It is possible that new hires' decision to opt in or out of the program may have been impacted by potential aversion to uncertainty as to whether they would receive a mentor even if they chose to opt in. We do not think this was likely a major factor in our setting, as sales agents experience a high level of uncertainty on a daily basis, suggesting that these new hires were unlikely to be strongly averse towards uncertainty. In

were then asked to write down 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). Agents who opted out were not given a mentor. Agents who opted in had the chance to receive a mentor. Among agents who opted in, the same procedure to allocate mentors from Broad-Mentoring was used to assign mentors to new hires.

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-week training process remained exactly the same for all agents, regardless of their mentoring branch allocation. The main point of contact between the authors and the firm was the within-firm (internal) mentoring staff, with at least one dedicated employee in each of the firm's two participating locations. The authors did not interface with any of the protégés or mentors, but instead only facilitated the randomization and data collection. After the two weeks of training, sales agents began working on the sales floor. First contact between mentors and protégés was usually initiated by the mentor during the protégés' first week on the sales floor, which coincides with when we first begin tracking their productivity, but sometimes the mentoring staff explicitly facilitated initial introductions. Mentors were incentivized to make sure they met with their protégés, so they were properly motivated to make contact and to conduct scheduled meetings.

3.2 Mentor-Protégé Meeting Protocol

The formal mentoring relationships lasted for four weeks. Each week, the mentoring staff worked with the firm's workforce management department to build into mentors'/protégés' schedules a specific break-time during which they could meet. This reduced the concern that scheduling conflicts would prevent pairs from meeting. Mentors were told the following (full documentation of the instructions can be found in Appendix D):

> You will meet with your protégé at least once a week. Before meeting, your protégé will complete the Protégé Worksheet. If he/she has not completed it, you will kindly

other settings, it may be prudent to solicit opt-in/out decisions before notifying individuals of the chance that they might not be selected even if they opt in.

> help him/her do so. During your meeting, you and your protégé will discuss his/her responses. You should also take this time to do the following:
>
> - Impart knowledge and skill by explaining, giving useful examples, and demonstrating processes, and asking thought-provoking questions.
> - Discuss actions you've taken to become a successful sales agent.
> - Provide him/her with any tips and sales tactics that have helped you overcome customer concerns and up-sell to better services.
> - Practice the designated sales protocol with them and help them gain a strong understanding of the products, services, and bundles available.
>
> After meeting with your protégé, you will deliver the finished worksheet to [the mentoring staff]. [The mentoring staff] will initial and timestamp the worksheet and make a record that you completed your weekly meeting responsibility.

During the meetings between mentors and protégés, the two were free to discuss whatever they wanted, but they did have to complete the worksheets in order for the mentors to receive credit for having held the meeting. Records of meeting occurrence and completed worksheets were kept by the mentoring staff and given to us. Shortly after the final week of the mentoring intervention, protégés were asked to complete a wrap-up survey that asked them questions about their mentorship experience. We use this data to gauge whether pairs continued to meet after the intervention ended and whether agents felt the program was beneficial, as discussed in Section 6.3.3.

3.3 Treatment Allocations

Figure 1 displays the allocation of cohorts and agents to the different branches and conditions of the experiment. There were 603 treatment-eligible sales agents spread across 53 hiring cohorts.[11] Twenty-two cohorts, and their 281 sales agents, were randomly allocated to Broad-Mentoring,

[11] An additional 56 rehired agents—those who had worked at the firm previously—were also mentor-eligible. We exclude these agents from our main analysis because their experience with the company could lead to meaningful differences in their decision to opt out of mentorship and the extent to which they benefit from mentorship. We show in Appendix B that our results are similar when we include these agents.

whereas the other 31 cohorts, along with their 322 sales agents, were randomly allocated to Selective-Mentoring. Among the agents in the Broad-Mentoring branch, 110 received mentorship and 171 did not (conditions 1a and 1b, respectively). Among the agents in Selective-Mentoring, 59 (18.3%) chose to opt out of receiving a mentor (condition 2c). Of the 263 that chose to opt in, 114 received mentorship and 149 did not (conditions 2a and 2b, respectively).[12]

As would be expected with successful randomization to treatments, agent characteristics are well balanced across the different treatment branches and conditions of the experiment. In Panel A of Table 1, we consider the balance across observable characteristics for agents in the Broad-Mentoring cohorts compared to those in the Selective-Mentoring cohorts (the top level of randomization). Comparing the cohort-level averages of agent age, gender, marital status, and hiring score, we find no significant differences between Broad-Mentoring and Selective-Mentoring cohorts. The average agent age in both groups is 23 years old. Women make up 44% of the agents in Broad-Mentoring and 40% of the agents in Selective-Mentoring, 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 for Broad-Mentoring and Selective-Mentoring, respectively. Formal tests of mean differences never reject the null of equality between the Broad- and Selective-Mentoring branches.[13] Panel B performs a similar test, but in this panel we consider agents who were not eligible for mentorship who were working in the divisions that newly hired mentor-eligible agents would join.[14] Non-mentor-eligible agents' average productivity levels in the divisions to which cohorts were assigned did not differ between the Broad-Mentoring and Selective-Mentoring groups. The average agent

[12] Limits on the supply of available mentors prevented us from achieving an exact 50/50 allocation. In addition, there were some newly hired agents who were initially randomly assigned a mentor but who, due to scheduling conflicts, were not able to receive mentorship. These individuals are coded as not having received mentorship. In Section 5.3 we show that our results are robust if we remove these agents from the sample, rather than coding them as non-mentored.

[13] We report similar statistics for the demographics of mentors in Table A.1, which shows that mentors of agents in Broad- and Selective-Mentoring were similar in age, gender, and tenure. The mentors of agents in Selective-Mentoring were more likely to be married than were the mentors of agents in Broad-Mentoring. Mentors were not designated exclusively to either of the mentoring branches. In other words, a mentor's first protégé could have been assigned to Broad-Mentoring, whereas their second protégé could have been assigned to Selective-Mentoring. Mentors were never informed as to whether their protégés were from Broad-Mentoring or Selective-Mentoring cohorts.

[14] We use non-mentor-eligible agents hired at about the same time as the participants in the experiment to capture performance metrics that are most similar to what mentor-eligible agents would have realized in the absence of mentorship. We provide additional discussion about non-mentor-eligible agents in Section 5.1.

who was not mentor-eligible generated \$760–\$794 in daily revenue, between \$46.80 and \$49.27 in revenue-per-call (RPC), between \$115.61 and \$120.21 in revenue-per-hour (RPH), while taking taking approximately 17 calls each day and spending a little over 6.5 hours at work.[15] In addition, adherence to schedule (i.e., *Adherence*, a measure of uptime and overlap of time on/off with planned breaks) does not differ between groups, nor does the conversion rate (an indicator of making any sale on a particular call). Across all productivity measures, difference-in-means tests fail to reject equality between groups.

Panel C of Table 1 considers the second level of randomization used in our experimental design, the random allocation of mentors to new hires *within* the Broad-Mentoring and Selective-Mentoring branches. Columns (1) and (2) show the agent-level average characteristics of agents in Broad-Mentoring who did and did not receive a mentor, respectively. Difference-in-means tests show that these two groups of agents are similar in age, gender, marital status, and in their hiring scores. Columns (3) and (4) and the associated p-values show that mentored and non-mentored agents who opted in among the Selective-Mentoring branch are similar across all four observable characteristics. Column (5) shows that agents who opt out of mentoring in the Selective-Mentoring branch have worse average hiring scores and are less likely to be women. Formal tests to compare the demographics of agents across all three Selective-Mentoring conditions show that all three are similar in age and marital status. Taken together, the summary statistics in Table 1 suggest that agent characteristics are, ex ante, well balanced across the different treatment branches and conditions of our experimental design.

4 Results

The first question we consider is whether mentoring has a positive effect on the productivity of employees during their first months of tenure. To answer this question, we compare the performance of mentored and non-mentored agents in Broad-Mentoring cohorts (conditions 1a and 1b, respectively), who were not given the option to opt in or opt out of the program. The comparison between conditions 1a and 1b is shown in the two leftmost bars of Figure 2a, which report agents' average

[15]The *Hours* measure that we use throughout this study captures the amount of time agents are at work and not on unpaid breaks (i.e., total time "clocked-in").

daily revenue in their first two months on the sales floor. Mentored and non-mentored agents in the Broad-Mentoring group generate average daily revenues of roughly \$782 and \$722, respectively, a \$60 (or 8.3%) gap in productivity. Figure 2b uses the natural logarithm of revenue on the y-axis and shows that mentored agents outperform non-mentored agents by approximately 0.18 log points (roughly 18%). We can reject that productivity is the same between conditions 1a and 1b at the 1% level, as reported in the figure note. Both figures indicate that Broad-Mentoring raises productivity, though our preferred productivity measure is the natural logarithm of revenue because it provides less weight to outliers and provides an intuitive interpretation as the approximate percent change.[16]

The second question we consider is whether mentoring has a similar productivity effect when new hires can select into the program. To answer this question, we compare the performance of agents who opted in and then did or did not receive a mentor via randomization (conditions 2a and 2b, respectively). The comparison between 2a and 2b, shown in the third and fourth bars of Figure 2b, shows that mentorship had no effect for agents who opted into the program. The difference between conditions 2a and 2b is small, and we fail to reject that productivity is the same with a p-value of 0.452 in the log specification. Said differently, in contrast to the large positive effect of having a mentor in the Broad-Mentoring group, we find no effect for those that opt into mentorship in the Selective-Mentoring group.

The difference between the Broad-Mentoring effect and the Selective-Mentoring effect suggests that those who stood to gain the *most* from mentorship were likely those who opted out. This motivates our third question: what type of agents opt out? To answer this question, we compare agents who opted in and did not receive a mentor to agents who opted out (conditions 2b and 2c, respectively). The comparison of 2b and 2c, in Figure 2b, shows that agents who opted in and did not receive a mentor have higher performance than agents who opted out, suggesting that opt-out agents were, on average, worse performers. The difference between conditions 2b and 2c is large and we can reject that productivity is the same at the 1% level. Said differently, we find a large positive selection effect on the decision to opt into mentorship.[17]

[16] We display similar bar charts in Figure A.1 to show agents' productivity during months 3–6 on the sales floor. These long-term effects are similar to the short-term, 1–2 month, effects, albeit they are less precise.

[17] We discuss the characteristics of agents who chose to opt out in Section 4.4. In general, demographic

In summary, we find a large, positive productivity treatment effect of receiving a mentor in the general population (conditions 1a vs. 1b), find no effect of receiving a mentor for those that opt into the mentorship program (conditions 2a vs. 2b), and find a positive selection effect, such that agents that select into our mentorship program have higher average daily revenues even if they do not receive a mentor (conditions 2b vs. 2c). These results have profound implications for understanding the returns to mentoring programs as well as for evaluating research designs intended to estimate returns to mentoring. Our design highlights a negative correlation between voluntary enrollment (participation) in mentorship programs and the gains from those programs. These findings suggest that observational estimates of the gains to mentoring programs are likely biased, but the direction of bias is unclear. For example, even potential designs that look at randomized mentoring among those who volunteer for a program may understate the effect of mentorship because of the first-level selection into participation. On the other hand, studies that fail to randomize and only look at those who select into mentorship might otherwise conclude that the effects of mentorship are substantial given underlying performance differences between those who opt in versus those who opt out. Our design allows us to separate treatment from selection, showing heterogeneous treatment effects that are largest for those who are least likely to participate. Said another way, agents who chose to opt out likely would have meaningfully benefited from mentorship. We defer discussion of why these agents opted out in the face of large potential gains to Section 4.4. We now formalize these insights with statistical tests and analyze additional measures of performance that provide context for these gains.

4.1 Does Broad-Mentoring Affect Productivity and Retention?

4.1.1 Sales Agent Productivity with Broad-Mentoring

Figure 2 shows a meaningful treatment effect of mentoring on the daily revenue production of agents in Broad-Mentoring. In this section, we show that this treatment effect extends to a host of other measures of agent productivity. To do this, we use a sample of agent-day productivity data for agents in Broad-Mentoring (those in conditions 1a and 1b) and estimate the following model

information, hiring scores, previous work experience, and personality traits explain very little of the variation in the opt-out decision.

using ordinary least squares:

$$\mathrm{y}_{i,t} = \alpha + \beta_1 \mathrm{Mentored}_i + \sum_{j=1}^{22} \mathrm{Cohort}_j + \varepsilon_{i,t}, \tag{1}$$

where $Mentored_i$ equals one if agent i receives mentorship, and zero otherwise, and t denotes the date of the observed productivity. All models include cohort fixed effects, $Cohort_j$, which control for differences in productivity that are specific to cohorts, and an idiosyncratic error term, $\varepsilon_{i,t}$.[18] 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 RPC is revenue-per-call, RPH is revenue-per-hour, and Adherence captures how closely agents adhere to their pre-set schedule (e.g., taking breaks and eating lunch at the correct time).[19]

In Table 2, we report the treatment effects of Broad-Mentoring on productivity across two different time periods: months 1–2 of an agent's tenure in Panel A and months 3–6 of an agent's tenure in Panel B. Column (1) of Panel A shows evidence of a positive and significant effect of mentorship on ln(Revenue). The point estimate suggests that mentored agents generate 19% more in daily revenue than do their non-mentored peers.[20] Panel B shows that over 90% of this effect persists in the 3–6 month period, but the point estimate is less precise. The lack of precision comes in part due to a smaller sample size, as there are fewer observations in months 3 through 6 due to turnover, even though the period in Panel B contains 2 additional months of data. Combining the insights from Panels A and B suggests that mentorship helps newly hired agents get up-to-speed more quickly, while also possibly leading to a level shift in performance that persists over time.

In Columns (2)–(4) of Table 2, we show that the positive treatment effect of Broad-Mentoring on productivity is apparent across the other performance metrics captured by the firm. The signif-

[18] All of our pre-registered specifications include cohort fixed effects, as we expected ex-ante that between cohort variation would significantly reduce detectable effects because cohorts face different demand conditions based on the products they sell and the timing of their hiring. Note that with cohort fixed effects, calendar time and elapsed time since hire are co-linear, so cohort fixed effects absorb time fixed effects when looking at fixed windows of time after the agent's hiring date. We show in Section 5.3 that our results are robust to the inclusion of date or day-of-week fixed effects. In addition, we show that our results are robust when we exclude cohort fixed effects.

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

[20] When we use revenue as the dependent variable, we find that mentorship increases daily revenue generation by over \$65. This effect size is similar to that presented in Figure 2a.

icant effects of *Mentored* on both selling efficiency measures—ln(RPC) and ln(RPH)—and on the main measure of labor supply—Adherence—suggest that mentored agents are both allocating more time to revenue-generating tasks and are generating more revenue per unit of time. Comparing coefficients between Panels A and B indicates that at least 70% of the 1–2 month treatment effects on ln(RPC) and ln(RPH) persist in the later months. These estimates, which standardize revenue generation based on the number of calls taken and hours worked, suggest that mentoring improves protégés' sales ability. The treatment effect on Adherence, however, decreases to nearly 25% of the 1–2 month baseline. This is likely due to the fact that the firm's policy requires agents to maintain an adherence level above 80% and provides no direct rewards for having incrementally higher adherence levels.

We defer a comprehensive discussion of robustness to Section 5, but it is important to note that these results do not arise from the discouragement of agents who are randomized out of receiving a mentor. Nor does the program appear to crowd out organic mentoring that may have occurred in the absence of a formal program. In addition, the estimated effects on productivity are not simply driven by differential retention between mentored and non-mentored agents.

4.1.2 Sales Agent Retention with Broad-Mentoring

We also consider the effect of mentorship on sales agent retention. Call centers have notoriously high levels of attrition (Hoffman et al., 2017). Executives at the study firm talk about the importance of new hires hitting multiple important milestones with tenure. The most important is the initial transition from training to the sales floor, where new hires begin working independently. Many new hires experience a "deer in the headlights" moment when they begin fielding calls from customers full-time, as they are not yet fully comfortable navigating the sales process and have little experience responding to the questions and objections of customers. While new hires have nearby coworkers and a sales manager who can potentially assist them, the additional guidance and support of a mentor may help agents persevere through this transition period. We thus expected mentorship to have a positive effect on the retention of newly hired agents, especially during their first month of employment.

We examine retention by reducing the sample to include only a single observation per agent. We then create indicator variables for whether an agent achieves a certain tenure threshold. Specifically, we create $Tenure_t$, which equals one for agents who achieve at least t months of tenure at the firm, and zero otherwise. We then estimate Equation (1) with $Tenure_t$ as the dependent variable. In Column (5) of Panel A in Table 2, we show that mentored agents are 11 percentage points more likely to stay with the firm than non-mentored agents at the one month threshold. This effect appears to persist to the two-month threshold, as shown in Column (6), but the effect is not precisely estimated. In Panel B, we report a much smaller and statistically insignificant effect of mentorship on retention at the three- and four-month thresholds.

These retention effects are visibly present in Figure 3a. This figure displays Kaplan-Meier survival rate estimates, which take agents' hire dates as a starting point and then display the fraction of agents who remain at the firm at various points in event time, relative to their hire dates. This measure allows us to depict retention over time for mentored and non-mentored agents separately, illustrating if and where their retention patterns diverge. In Figure 3a, we see that over 20% of non-mentored agents in the Broad-Mentoring cohorts exit the firm within one month of their hire date. Among mentored agents, however, only about 8% depart in their first month. This gap in retention persists in month two, and then the gap shrinks in months three and four—consistent with the results in Columns (5) and (6) of Panel B of Table 2. The survival rates for both groups of agents then follow a similar pattern, with a flattening of the curve in months five and six at around 23%–29% retention. Taken together, the results in Table 2 provide experimentally-identified evidence that a formal workplace mentorship program, when broadly applied across newly hired sales agents, significantly increases productivity and early stage retention.

4.2 Does Selective-Mentoring Affect Productivity and Retention?

4.2.1 Sales Agent Productivity with Selective-Mentoring

Next we ask whether mentoring has a similar effect on productivity for agents that opt into the mentoring program. To answer this question we compare the performance of agents who opt in and received a mentor with those who opt in but did not receive a mentor due to our randomization.

We report estimates using the following model:

$$y_{i,t} = \alpha + \beta_1 \text{Mentored}_i + \beta_2 \text{Selective Opt-In}_i + \sum_{j=1}^{31} \text{Cohort}_j + \varepsilon_{i,t}, \tag{2}$$

which is the same as Equation (1), except it includes *Selective Opt-In*$_i$, which equals one for agents in Selective-Mentoring who opt in, and zero otherwise. For these estimations we use only individuals in the Selective-Mentoring group, which consists of three conditions: Selective-Mentoring, opted in and mentored (condition 2a); Selective-Mentoring, opted in and non-mentored (condition 2b); and Selective-Mentoring, opted out (condition 2c). The excluded category in Equation (2) is thus agents who opt out.

Table 3 reports results from these estimations, again with Panel A reporting productivity in months 1–2 and Panel B reporting productivity in months 3–6. The point estimates on *Mentored* are indistinguishable from zero in Columns (1)–(4) of Panel A of Table 3. This suggests that mentorship had no discernible effect on the 1–2 month productivity of newly hired agents who signaled their desire to receive a mentor. This finding is consistent with the results in Figure 2, which show that the mentored opt-in agents do not generate more revenue than do non-mentored opt-in agents in the first months on the sales floor.

The revenue differences displayed in Figure 2 also show that agents who opted out had worse productivity in their first months on the sales floor than did agents who opted in and did not receive a mentor. The point estimates in Panel A of Table 3 on *Selective Opt-In* allow a comparison between those who opt in and those who opt out of the program. The coefficients in Columns (1)–(3) are positive and statistically significant, supporting the view that agents who opt in are more productive than those who opt out in the absence of mentoring. The estimate in Column (1) implies that agents who opt in generated over 31% more revenue each day than did agents who opt out. Hence, the average opt-out agent is less productive than the average opt-in agent in months one and two. This is an estimate of selection bias in the absence of mentorship, but it does not say anything about heterogeneous effects of mentorship. Treatment effect heterogeneity is critical for understanding whether different allocation rules can lead to gains from the design of mentorship programs. We further discuss differences between opt-out and opt-in agents in Section 4.4.

4.2.2 Sales Agent Retention with Selective-Mentoring

Mentorship in the Selective-Mentoring group does have a positive effect on the one-month retention of newly hired agents. In Column (5), we report that agents that opted in and received a mentor are nearly 12 percentage points more likely to achieve one month of tenure than are non-mentored opt-in agents. This effect size is similar to the 11 percentage point estimate in the Broad-Mentoring group. This retention effect can be seen visually in Figure 3b, which shows the Kaplan-Meier survival rate estimator for the three different treatment conditions of Selective-Mentoring. This figure shows that about 17% of non-mentored opt-in agents in the Selective-Mentoring cohorts 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, the overall retention rates between mentored and non-mentored opt-in agents only differ by 4 percentage points. While increased retention is an important outcome for the firm, the reduced turnover costs are unlikely to offset the loss in revenue associated with pulling mentors away from calls. Simply put, our estimates suggest that the retention benefits associated with Selective-Mentoring do not warrant assigning a mentor to every agent within the Selective-Mentoring group. We further discuss the net present value of the mentoring program in Section 6.4.

4.3 Differential Effects of the Mentoring Programs

The results so far indicate that mentorship has a significant, positive effect on productivity in the Broad-Mentoring group and a more negligible effect in the Selective-Mentoring group. To determine whether the effect of mentorship is significantly different between the two groups, we include an interaction term between *Mentored* and *Broad*, a binary indicator equal to one for agents in the Broad-Mentoring treatment, and zero otherwise, into Equation (1). We use a sample of agent-day productivity data of all mentor-eligible agents, albeit we exclude agents who opt out. We then estimate the following model using ordinary least squares:

$$\mathrm{y}_{i,t} = \alpha + \beta_1 \mathrm{Mentored}_i + \beta_2 (\mathrm{Mentored} \times \mathrm{Broad})_i + \sum_{j=1}^{53} \mathrm{Cohort}_j + \varepsilon_{i,t}, \tag{3}$$

where *(Mentored* $\times$ *Broad)*$_i$ equals one for agents who were mentored and in a Broad-Mentoring cohort, and zero otherwise. The inclusion of cohort fixed effects subsumes the *Broad* indicator, which is why we do not separately include *Broad* in the model.

We report the results of this test for differential effects in Table 4. Columns (1), (2), and (3) of Panel A show that the treatment effect of Broad-Mentoring on ln(Revenue), ln(RPC), and ln(RPH), respectively, was significantly larger than that of Selective-Mentoring. Using level revenue as the dependent variable implies that the treatment effect of Broad-Mentoring was greater than that of Selective-Mentoring by over \$108 worth of daily revenue. We do not find significantly different treatment effects between the two mentorship groups at the 3–6 month horizon, nor do we find a differential effect on retention. However, the differential treatment effect on productivity at the 1–2 month horizon suggests that mentoring programs that do not give workers the choice to opt out are significantly more effective at improving productivity than those that make mentorship optional. This finding also implies that the greatest benefits to a formal mentorship program may be realized by those who opt out of receiving mentorship when given the choice.[21]

4.4 How do Opt-out Agents Differ from Opt-In Agents?

The differential treatment effects of Broad-Mentoring and Selective-Mentoring suggest that opt-out agents stood to gain from mentorship. Before we formally estimate these potential gains, we consider how this subset of agents differs from those who chose to opt into the mentorship program. In Section 4.2, we showed that opt-out agents were less productive, on average, than opt-in agents. To understand why opt-out agents have lower early-stage productivity than opt-in agents, we use data on the observable characteristics of the sales agents to estimate the determinants of opting out. We restrict our sample to only include the 322 agents in the Selective-Mentoring cohorts, and then we estimate the following logistic model:

$$\text{Opt-Out}_i = \alpha + \beta_1 \text{Age}_i + \beta_2 \text{Female}_i + \beta_3 \text{Married}_i + \varepsilon_i. \tag{4}$$

[21] 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 Tables A.2–A.4. These results suggest that mentorship had a positive impact on the number of calls taken and the number of hours work of protégés in Broad-Mentoring, whereas mentored agents in Selective-Mentoring did not realize similar effects.

where *Opt-Out* equals one if the agent opts out of receiving a mentor, and zero otherwise. *Age* captures the age of the agent at the time they were hired. *Female* equals one if the agent is a woman, and zero otherwise. *Married* equals one if the agent is married, and zero otherwise, and ε_i is an idiosyncratic error term. We display the result of this estimation in Column (1) of Table 5, which indicates that neither age, gender, nor marital status are significant predictors of the opt-out decision.

In Column (2) we include *Hiring Score* into the logistic regression, 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 equal to zero and include a dummy variable into the model that indicates that they had missing data. The results in Column (2) show that *Hiring Score* is a significant predictor of *not* opting out. Said another way, agents with low hiring scores are significantly more likely to opt out of mentorship when given the choice.

In Column (3), we control for several other observable characteristics that may influence an agent's opt-out decision. Specifically, we control for their location, whether they were referred by an existing agent, and their personality traits (which we gathered via the new hire survey). One of the two call center offices is located near two large universities, which results in a labor force made up of many college students, some of whom are peers at school, which potentially facilitate informal peer-to-peer learning opportunities that may crowd out the need for formal mentoring. Similarly, being referred to the company by an existing employee suggests that a newly hired agent may already have access to informal mentorship opportunities. In addition, one might expect that an individual's personality traits—in particular, their degree of extroversion—might influence their decision to receive formal mentorship. We find that none of these additional factors have significant predictive power for the decision to opt out, and *Hiring Score* maintains its explanatory power even when we control for these factors.

In Column (4), we incorporate two final variables into the model: *Call Center Exp.*, which equals one if the agent indicated in the hiring survey that they had previous experience working in a call center, and zero otherwise, and *Sales Experience*, which equals one if the agent indicated in the hiring survey that they had previous experience working in sales, and zero otherwise. The results in Column (4) show that neither *Call Center Exp.* nor *Sales Experience* have significant explanatory power for opting out. While *Hiring Score* continues to be a significant predictor of the opt-out decision, all of these observable characteristics, together, explain less than 7% of the variation in the opt-out decision. This suggests that unobservable characteristics largely drive the decision to forgo mentoring. We also know that the opt-out agents perform worse on average than agents that opt in, and this too could be driven by unobservable characteristics.

4.5 Would Opt-Out Agents Have Benefited from Mentorship?

We can use the estimated Broad-Mentoring and Selective-Mentoring treatment effects, along with the data on the fraction of Selective-Mentoring agents who opt out of receiving a mentor, to estimate the treatment effect of mentorship among agents who opted out. We pre-registered the following procedure for this purpose. Using output measure Y, we define the average treatment effect of mentoring given selection as:

$$\underbrace{\beta_{OptInMentored}}_{ATE|Selection} = E(Y_{OptInMentored}) - E(Y_{OptIn\sim Mentored}).$$

We can then write the unconditional average treatment effect of mentorship as the weighted average of heterogeneous effects with shares π:

$$\underbrace{E(Y_{BroadMentored}) - E(Y_{Broad\sim Mentored})}_{ATE} = \beta_{OptInMentored} \times \pi_{OptIn} \ + \beta_{OptOutMentored} \times \pi_{OptOut}.$$

Rearranging terms, we get,

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

We use the estimated treatment effect of Broad-Mentoring as the estimated ATE, and we use the estimated treatment effect of Selective-Mentoring as the estimated $ATE|Selection$.[22] The values of π come from the proportion of agents who opted out versus opted into mentoring among the Selective-Mentoring cohorts. We show the estimated treatment effect for opt-out agents in Table 6, where standard errors come from 500 block-bootstrap iterations by cohort. The point estimate

[22]We include cohort and day-of-week fixed effects in estimating these treatment effects.

of 1.313 in Column (1) of Panel A implies that opt-out agents would have been able to more than double their overall revenue generation, on average, had they received mentorship. It appears likely, therefore, that opt-out agents in this setting were actually those who would have benefited the most from receiving additional help from a mentor. The table also displays Δ*Effect: Opt-Out – Opt-In*, which is the difference in our inferred treatment effects for those who opt out less the estimated treatment effects for those who we explicitly observe opting in. This allows us to test for treatment effect heterogeneity. We reject equality of treatment effects between those who opt in and those who opt out at the 10% level, when considering ln(Revenue), ln(RPC), and ln(RPH). This indicates a negative correlation between treatment gains and the propensity to select into treatment. Estimates involving Adherence are also positive and statistically significant. As expected, we do not estimate precise heterogeneous treatment effects when considering productivity at the 3–6 month horizon in Panel B.

We also implement a GMM estimator as an alternative to making cross-treatment comparisons to look for heterogeneous effects. This analysis was not pre-registered and as such we report it in Appendix C. The GMM estimator simultaneously searches for common agent-characteristics in the Broad-Mentoring group that match those of the opt-out agents in the Selective-Mentoring group, allowing us to compare productivity effects for similar agents. We let the opt-out probability vary with agent and cohort characteristics, and find similar, substantial heterogeneous treatment effect estimates for non-mentored agents in the Broad-Mentoring branch who were deemed likely to opt out if provided the choice. Specifically, the GMM approach implemented is a systems estimator that recovers an interaction between a latent opt-out probability (estimated in one equation) and treatment effects in a productivity equation. The estimated treatment effect on log revenue for those who would opt out using this approach is 0.84 with a standard error (clustered by cohort) of 0.36.

As an additional test of how the gains to mentorship differ between agents who do and do not seek it out, we use the observable characteristics in Column (2) of Table 5 to estimate the opt-out propensity scores of agents in Board-Mentoring. We then classify Broad-Mentoring agents as either $High_{Opt}$, if their opt-out propensity score is in the top 30% of the distribution, or Low_{Opt}, if their opt-out propensity score is in the bottom 70% of the distribution. We use these thresholds to create sample sizes that (1) are similar to the actual opt-in versus opt-out groups observed in the Selective-Mentoring branch and (2) are still large enough for reliable inference. We then estimate Equation (1) within these subsets of the data with either ln(Revenue), ln(RPC), or ln(RPH) as the independent variable. To determine if the effect of mentorship is significantly different between the $High_{Opt}$ and Low_{Opt} agents, we include a one-zero indicator for $High_{Opt}$ into the model, along with the interaction between this and *Mentored*. The results are displayed in Table A.5 and show that, in the first two months on the sales floor, Broad-Mentoring agents with a high estimated likelihood of opting out had a significantly greater treatment effect of mentorship than did their peers who were less likely to opt out. This finding corroborates the results in Table 6, which imply that the greatest gains to mentorship are realized by agents who opt out of receiving a mentor when given the option to do so. Later, Section 6.1 discusses whether agents' observable characteristics lead to treatment effect heterogeneity, while Section 6.2 discusses whether worker characteristics can be used to improve program targeting.

5 Robustness

In this section we rule out several alternative explanations for our results, and we discuss a host of robustness tests. First, we show that our results are not driven by discouragement or leakage. We then show that the mentoring program does not appear to have crowded out organic mentorship that would have occurred in the program's absence. Finally, we test several additional specifications to the main analysis to assuage robustness concerns, and we show that the effects on productivity are not simply driven by non-random attrition.

5.1 Effects Are Not Driven By Discouragement or Leakage

We first consider the possibility that discouragement or leakage may have affected our estimates of the treatment effects of mentorship. After a number of exercises meant to detect these issues, we find no evidence that these potential concerns alter our findings. However, it is useful to detail how these different channels may have affected our estimates. First, agents who did not receive a mentor may have become discouraged, reducing their performance as a result. Discouragement would cause our estimates to overstate the actual benefits of receiving a mentor because it would negatively affect non-mentored agents, who form the control group for our estimates. Second, agents that did not receive a mentor via our random allocation process may have sought out their own mentor, leading to treatment leakage. A different source of leakage may be where non-mentored agents query 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 both the Broad- and Selective-Mentoring branches, it is possible that agents who opted into the program in the Selective-Mentoring group would be most likely to seek out help, as the framing of this condition gave agents a choice to participate, which would attenuate the estimated treatment effect of mentorship in the Selective-Mentoring group.

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 in the Broad-Mentoring cohorts received information about the mentorship program. Specifically, we asked the internal mentoring staff to privately notify new hires who would receive a mentor about their involvement in the program—dampening the potential for discouragement among non-mentored agents. Second, the firm's internal mentoring staff told agents in the Selective-Mentoring cohorts, all of whom were made aware of the program, that ample opportunities for receiving help were available to those who did not end up receiving a mentor—again, 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, albeit no such feedback was reported back to us.

We test 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 (whose tenure exceeds 18 months).[23] Unlike the experimentally identified results within cohorts and across the Broad-

[23]The non-mentor-eligible hiring cohorts were cohorts of new hires that joined the firm at a time when no mentors were available. As a result, no one in these cohorts was mentored, making them a hold-out group that was not aware of the mentorship program to the same extent as agents in the mentor-eligible hiring

and Selective-Mentoring treatment conditions, this exercise requires a comparison to agents who were not part of the experiment.

Our approach is to compare the productivity of new hires relative to veteran agents in these different conditions. Under the null of no discouragement or leakage, we would expect that new hires who are not mentored but who are in eligible cohorts should have indistinguishable productivity differences (relative to veterans) compared to new hires entering the firm in non-mentor eligible cohorts or prior to the existence of the formal mentoring program. Because these results are not experimentally identified, we use regression adjustments to make conditions comparable between new hires and veterans across the different treatment or hold-out groups.[24] We estimate the following model using ordinary least squares:

$$\begin{aligned}\ln(\text{Revenue})_{i,t} = \alpha &+ \beta_1 \text{New Hire}_i + \beta_2(\text{New Hire} \times \text{Broad})_i + \beta_3(\text{New Hire} \times \text{Selective})_i \\ &+ \beta_4(\text{Mentored} \times \text{Broad})_i + \beta_5(\text{Mentored} \times \text{Selective})_i + \zeta_{j,l,t} + \gamma_{n,l} + \varepsilon_{i,t},\end{aligned} \tag{5}$$

where *New Hire* equals one if the agent has tenure of two months or less and zero for seasoned veterans, *Broad* equals one for agents in Broad-Mentoring and zero otherwise, *Selective* equals one for agents in Selective-Mentoring and zero otherwise, *Mentored* equals one for agents who receive mentorship and zero otherwise. Divisions may experience their own idiosyncratic shocks that can be partially smoothed with different staffing choices that may vary across office, so we include division-by-location-by-date fixed effects, through $\zeta_{j,l,t}$, to provide flexibility. Recruiting tends to occur at the office level, and then agents are allocated to different divisions. To capture differences in new hires across offices, $\gamma_{n,l}$ removes a location-by-new-hire fixed effect. Finally, $\varepsilon_{i,t}$ is an idiosyncratic error term.

The results of this estimation are displayed in Column (1) of Table 7. This table only contains data from the period during the experiment, and is designed to test for discouragement and leakage.

cohorts. These hold out groups were not randomly allocated into the non-eligible condition, but instead were not eligible for mentorship due to operational constraints. Hold outs tended to occur for cohorts hired right after a prior cohort in the same brand and office received treatment because these were times when the supply of mentors was scarce.

[24] 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 5 or more mentor-eligible agents and 5 or more new hires that were not mentor-eligible.

The negative and statistically significant coefficient on *New Hire* suggests that newly hired agents generate approximately 42% less in daily revenue, relative to seasoned veterans. The small and insignificant coefficients on *(New Hire* $\times$ *Broad)* and *(New Hire* $\times$ *Selective)* suggest that newly hired non-mentored agents in mentor-eligible cohorts perform similarly to newly hired agents in non-mentor-eligible cohorts. This result suggests neither discouragement nor leakage are likely an issue in our setting. In addition, we fail to reject that these estimates are jointly equal to zero and that they are equal to each other, suggesting that newly hired, non-mentored agents in Broad-Mentoring performed similarly to newly hired, non-mentored agents in Selective-Mentoring. The positive and statistically significant effect on *(Mentored* $\times$ *Broad)* and the insignificant effect on *(Mentored* $\times$ *Selective)* align with our main mentoring treatment effects discussed in Sections 4.1 and 4.2. Column (2) shows similar results when we use an alternative combination of fixed effects that add additional 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 characteristics, which is important given that randomization of agents into treatments did not occur for seasoned veterans and non-mentor-eligible agents.

Columns (4)–(6) repeat this exercise by defining *New Hire* to equal one if the agent has tenure greater than two months and less than or equal to six months. This aligns with our previous analysis in which we separately assess the short-term, months 1–2, and long-term, months 3–6, effects of mentorship on productivity. The negative point estimate on *New Hire* in Column (4) shows that seasoned veterans still significantly outperform less tenured agents. The small, insignificant coefficients on *(New Hire* $\times$ *Broad)* and *(New Hire* $\times$ *Selective)* further support the notion that discouragement and leakage are not likely driving our estimated mentorship treatment effects.

5.2 Did the Mentoring Program Crowd Out Organic Mentoring?

It is possible that our formal mentoring program may have crowded out organic mentoring that would have occurred in the program's absence. Specifically, if our formal program absorbed all of the potential mentors, then non-mentored agents would perform less well because they did not have the opportunity to naturally find a mentor. As a result, we would expect that non-mentored agents in treatment-eligible cohorts would perform less well than new agents at other times.

To test the possibility that the mentorship program crowded out organic mentoring that would have occurred in the program's absence, we compare the productivity of mentor-eligible new hires to the productivity of new hires who joined the firm prior to the onset of the mentorship program. We continue to use the performance of veterans as a basis for comparison. If the formal mentorship program in our study crowded out organic mentorship opportunities for non-mentored agents—for example, by occupying all potential mentors with formal mentorship duties—then we would expect non-mentored new hires who were mentor-eligible to perform worse than new hires from prior years. We find no such evidence. Table 8 shows results that are similar to those in Table 7, albeit the comparison group is new hires from before the mentorship program, rather than those from contemporaneous hiring cohorts (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* $\times$ *Broad)* and *(New Hire* $\times$ *Selective)* in Columns (1)–(3) suggest that non-mentored, mentor-eligible new hires are no less productive than new hires from previous hiring dispensations, suggesting that crowd-out is not an issue in our setting. In the first two months, the positive coefficients become smaller and statistically insignificant in Column (3) when we adjust for agent characteristics, suggesting there may be composition differences in new hires over time. It is not surprising to find no evidence of crowd-out, as we worked with a company where there was relatively little organic peer-to-peer mentoring at baseline (Sandvik et al., 2020). This further reduces the concern that our formal mentorship program crowded out mentoring that would have occurred naturally.

5.3 Robustness of the Main Mentoring Effects

In this section, we consider several additional specifications that highlight the robustness of our main findings. We report these specifications in the coefficient plots in Figures 4a and 4b for Broad-Mentoring and Selective-Mentoring, respectively. The baseline coefficient and 95% confidence intervals from Panel A of Table 2 and Table 3 are displayed at the top of Figure 4a and Figure 4b, respectively.

First, we consider specifications that do not include cohort fixed effects (second line). Next, we include additional control variables. In the third and forth lines, we report estimates where we include date fixed effects and day-of-week fixed effects, respectively, to capture variation in productivity that is idiosyncratic to a particular day on the sale floor (e.g., accommodating positive Monday demand shocks). Next, we include controls for an agent's age, gender, and marital status. The effect of mentorship on productivity is little changed under these alternative specifications. If we control for sale agents' hiring scores, the sample size decreases, somewhat reducing observational power, but the effect continues to be statistically significant in Figure 4a and the effect size remains similar to previous models.

The next robustness check we consider is whether these results differ between time periods. We administered the mentoring protocol throughout the entirety of 2019. At the beginning of the year, new cohorts were placed into Broad-Mentoring versus Selective-Mentoring based on the hiring cohort's location.[25] In the second half of the year, starting with cohorts hired on May 27, 2019, marking the beginning of our pre-registered intervention period, we began randomly allocating cohorts into either Broad-Mentoring or Selective-Mentoring. If we include a control for the time period in which the cohort was hired, the results remain relatively unchanged. The bottom two lines show that our results hold when imposing two other restrictions on the data: one that removes agents who were initially assigned a mentor but who were unable to participate in the program due to scheduling conflicts,[26] and one that removes observations for which the agent was no longer working for the same division that they were mentored into. The robustness of our main results across all of these different specifications bolsters our conclusion that Broad-Mentoring, but not Selective-Mentoring, significantly improved agent productivity.[27]

[25] This was done to allow the mentoring staff at each location time to get trained on one protocol before administering both.

[26] The firm's mentorship staff worked with the human resource management department to find overlap in the schedules of mentors and protégés that allowed them to be off the phones at the same time, thus facilitating a meeting with minimal coordination difficulties. In some instances, the staff was unable to find such schedule alignment, and these would-be protégés ended up being non-mentored.

[27] We present similar coefficient plots for the other outcome variables in Figures A.2 and A.3.

5.4 Lee Bounds Estimates

The results in Section 4 provide evidence that mentorship significantly increases productivity and early stage retention. The higher productivity due to mentoring captures benefits both from the intensive margin (within-agent increases from agents learning more) and the extensive margin (due to lower/differential attrition). To separate these margins, we estimate treatment effect bounds that account for non-random attrition as proposed by Lee (2009). The key assumption when implementing this approach in our setting is that some mentored agents would have left the firm in the absence of mentorship, but no mentored agents left the firm *because* they were mentored (a traditional monotonicity assumption). Table A.6 reports upper and lower bounds of the estimated treatment effect of Broad-Mentoring 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 0.101 is about half the size of the main effect in Table 2, 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 learning more.

6 Additional Analysis

Next we consider several extensions of our main analysis to better understand the possible mechanisms at play. First, we discuss heterogeneity in the observed treatment effect of mentorship across observable characteristics. We then discuss the implications of our results for the design of other managerial interventions. We leverage worksheet and survey data to investigate additional potential mechanisms. To end, we consider the net present value to the firm of the mentorship program, and we address the external validity of our results.

6.1 Heterogeneous Effects

An important question is whether the factors that determine the opt-out decision also drive gains to mentorship, which may explain why opt-out agents stand to benefit from being mentored. In this section, we consider whether there are heterogeneous mentorship treatment effects based on observable sales agent characteristics (age, gender, martial status, and hiring score). To test for heterogeneous effects, we repeat the analysis used to populate Table 2 by considering the effect of mentorship on the productivity of agents in Broad-Mentoring. In Column (1) of Table A.7, we replicate the ln(Revenue) result from Panel A of Table 2. We then interact *Mentored* with the different characteristics of interest separately in Columns (2)–(5). Column (6) includes all characteristics and interaction terms together in a single model.

The positive and significant point estimates on *Mentored* $\times$ *Age* in Columns (2) and (6) suggest that older sales agents benefit more from mentorship than do their younger colleagues. In contrast, the negative and significant point estimates on *Mentored* $\times$ *Female* in Columns (3) and (6) suggest that female sales agents benefit less from mentorship than do their male colleagues.[28] When we re-estimate this model using an adaptive lasso regression that penalizes over-fitting, we find that the only significant heterogeneous effects of mentorship are based on age and marital status. The penalized coefficient on mentored from the lasso is 0.16, and the coefficients on the interaction of mentored with above-median age and married are 0.16 and -0.12, respectively.[29] This heterogeneity would not have have been known ex-ante without the experiment, and ex-post it appears treatment effect heterogeneity by observable characteristics is less important than whether the program allowed self-selection in participation.

In addition, we consider one other difference across mentored agents, whether their mentor was encouraged to focus on teaching them skills or providing emotional support. Two different versions of the mentor instructions were randomly distributed to the mentors. One placed slightly more emphasis on the importance of befriending protégés and providing them with emotional support. The other emphasized the importance of teaching skills and working through the sales process. This additional variation was introduced in an attempt to understand whether providing social

[28] In line with the insider econometrics approach (Bartel et al., 2004), we shared these results with the firm to get their feelings as to what might be the cause of these heterogeneous effects. It was suggested that the age effect might be attributable to older employees taking the job more seriously, while younger employees—who potentially see the job as temporary—might be less inclined to invest meaningful time into being mentored. On the other hand, in Column (6) we estimate a negative, weakly significant coefficient on *Mentored* $\times$ *Married*, which stands at odds with this interpretation if one assumes that married employees are more likely to seek out permanent employment. The gender effect is potentially due to the fact that nearly all of the mentors were men, and prior research has suggested that gender differences potentially matter in relationships (Athey et al., 2000; Porter and Serra, 2020), though we are cautious in this interpretation.

[29] Note that there are only 5 married agents who are below median age, so the negative coefficient largely offsets the positive effect of age for about 23 percent of the older sample.

support versus the dissemination of skills was most important. However, we do not find evidence that one type of mentorship was more effective than the other. It is possible that the intervention that we implemented—slight changes to the instructions the mentors were given—had too light a touch to have altered mentor behavior toward social support or knowledge transmission.

6.2 Implications for Program Design

The relatively limited treatment effect heterogeneity due to observable worker characteristics has implications for program design. To make a program designer's objective function concrete, consider the task of allocating M mentors to N workers, with $M < N$. The designer's task is to maximize the sum of gains from treatment, or $\sum \tau_i(M)$, where τ_i is the individual heterogeneous treatment effect. One possible design allows workers to self-select into mentorship, with any excess demand rationed in some way. Our finding that broad-based programs are much more effective than selective programs indicates that this strategy would not maximize treatment gains.

Alternatively, a planner might try to improve on random allocation. Could a planner or manager utilize heterogeneity based on observable characteristics to do so? The answer to this question is subtle. First, it was not obvious ex-ante what dimensions of treatment effect heterogeneity would be present, so implementing a broad-based program that is relatively untargeted to learn about heterogeneous effects is likely a critical first step. Key to inference is having at least some data that come from individuals that were not allowed to self-select into the program, as treatment gains and selection propensities are negatively correlated. More generally, our results suggest that having at least one experimental branch with broad-based randomization is necessary to detect treatment effect heterogeneity based on observables. From this data, an adaptive planner may improve on targeting, but our heterogeneous treatment effects analysis suggests that allocation improvements based on demographic targeting would be limited in our firm.

Second, what about targeting based on job performance rather than demographic characteristics? In our context, this type of targeting would come at a cost because workers would need to struggle on their own to make initial performance observable. Our design does not allow us to test the possibility that this targeting rule would raise productivity. This is because mentorship in our experiment is assigned at the beginning of tenure and mentorship effects may differ after some elapsed time on the job.[30]

6.3 Evidence on Potential Mechanisms

Here we explore additional possible differences between mentored agents in Broad- versus Selective-Mentoring that may have contributed to the observed differences in mentorship treatment effects. We specifically consider differences in mentor meeting completion rates and worksheet content. We also compare mentored agents in Broad-Mentoring with high versus low opt-out propensity scores. We then briefly discuss anecdotal evidence from the wrap-up surveys.

6.3.1 Meeting Completion Rates

The first potential mechanism that we explore is the amount of time protégés spent interacting with their mentors. To capture this, we assess the mentorship meeting completion rates of mentor-protégé pairs. Meeting completion rates were generally high, with the average mentor-protégé pair completing 82% of their scheduled meetings, and with over 62% of the pairs meeting every week. Meeting completion rates were significantly higher among Broad-Mentoring agents, 86%, than among Selective-Mentoring agents, 78%. This difference is significant at the 5% level and potentially suggests that the opt-in/opt-out protocol of the Selective-Mentoring branch caused agents to see meeting completion as more optional than mandatory. It may be the case that program framing effects were partially at play in our setting (Hossain and List, 2012; Hong et al., 2015; Englmaier et al., 2017). These framing effects are unlikely to explain the entire difference in treatment effects between Broad-Mentoring and Selective-Mentoring, however, as our main results in Tables 2–4 are robust when controlling for agents' meeting completion rates. Furthermore, the 78% completion rate in Selective-Mentoring means that the vast majority of scheduled meetings still took place.

Next, we consider whether Broad-Mentoring agents with high versus low opt-out propensity scores have different meeting completion rates. In Section 4.5 we showed that agents with a high opt-out likelihood benefited more from mentorship than did agents with a low opt-out likelihood.

[30] We suspect, based on the results in Sandvik et al. (2020), that delay may be beneficial, as we found large treatment effects from a similar intervention among veteran agents who were below-median performers. Again, however, this prior intervention was not targeted toward low performers.

If we compare the mentor meeting completion rates of these agents, we find that agents in the $High_{Opt}$ group completed 91% of there scheduled meetings, whereas agents in the Low_{Opt} group only completed 85% of there meetings. This difference is not statistically significant, though, suggesting that these two groups of protégés did not spend a significantly different amount of time interacting with their mentors. Taken together, our results do not appear to be driven by differences in meeting completion rates.

6.3.2 Worksheet Content

The second test we undertake to understand potential mechanisms is to examine the worksheet content of mentored agents. We do this two different 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 not a perfect measure of the quality of the mentor-protégé meetings, it allows us to proxy for the level of engagement the agents had towards the mentorship protocol. 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.[31] Specifically, we count up 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 do not get classified as either skill or support words are categorized as "other," including stop words.

We begin by comparing the worksheet content of Broad-Mentoring agents and Selective-Mentoring 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 per worksheet variables on the indicator *Broad-Mentoring*, which equals one for agents in Broad-Mentoring and zero for agents in Selective-Mentoring. Panel A of Table A.8 reports the results. We do not find meaningful differences in the number of total

[31] We list the words in each category in Appendix E, along with multiple examples of responses.

words, skill words, or other words recorded by agents in the two mentorship groups. We find some evidence that Broad-Mentoring agents use more support words than do Selective-Mentoring agents. The point estimate represents a 30% greater use of support words, but the effect is only significant at the 10% level. In Panel B, we consider only the mentored agents in Broad-Mentoring to estimate differences in worksheet content between those with a high opt-out propensity score and those with a low opt-out propensity score. We do not find statistically significant differences across any of the word type categories. Taken together, differences in worksheet content do not appear to explain the heterogeneous treatment effects of mentorship. 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 features or characteristics differed by agent. That is, agents who were most likely to opt out of mentorship appeared to benefit more from the same types of mentorship that agents with a low opt-out propensity received.

6.3.3 Wrap-Up Survey Responses

As a final way to understand how mentorship drove the observed productivity benefits, we consider evidence from survey responses. Two weeks after mentors and protégés completed their final mentorship meeting, the protégés were asked by the internal mentoring staff to complete a post-mentorship wrap-up survey. This survey asked them questions about how beneficial they felt their mentoring relationship was and whether they had continued to have contact with their mentor after the formal meetings ended. The completion rates for this survey were quite low (less than 10%), as the firm did not provide any explicit incentives for completing it. As such, we treat these responses as anecdotal evidence of the feelings that some had towards the mentorship program.

We display the protégés' responses to ten different questions in Figure 5, all of which were presented as statements to which respondents were asked to indicate their level of agreement/disagreement. The responses indicate that protégés, on average, felt like they and their mentors both benefited from the mentoring relationship. We do not find evidence that the relationships between mentors and protégés extended beyond the workplace, nor do we find responses that suggest the mentoring relationship distracted agents from reaching their potential. Most importantly, the responses indicated that mentors and protégés continued to interact after the formal relationship ended.

Specifically, 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 four-week protocol ended. This suggests that one benefit of the mentoring program was that it provided new hires with an additional resource—their mentor—to receive help in the future. The average respondent also said that mentorship helped them incorporate important selling tactics into their sales process and that mentorship increased their day-to-day satisfaction at work. This suggests that protégés likely benefited from both an enhanced knowledge of the sales process and a greater level of social support in the office.

6.4 Net Present Value of Mentorship

Next we estimate the returns to mentorship for the firm by considering productivity out to the six month horizon, net of administrative costs associated with the program. The average agent, among the 110 mentored agents in the Broad-Mentoring intervention, experienced over an 18% increase in daily revenue, which results in \$2,400–\$3,100 more in revenue per agent-month, depending on the tenure of the agent. The firm earns this additional revenue net of an 8% commission rate that is paid to sales agents. We then multiply these monthly net-revenue amounts by the number of mentored sales agents still present in the firm each month, over these six months.[32] We discount future cash flows using a 12.5% discount rate, which gives us a present value of the additional revenue earned by mentored agents equal to \$939,094.

We subtract the estimated time costs of taking the mentors and protégés off the phone 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. The average revenue-per-hour for mentors and protégés are \$138.28 and \$92.77, respectively. Mentors were also paid an additional \$10 of "kudos" points for completing each meeting. Together this implies a cost of \$125.53 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

[32] We refer to the survival rate trends in Figure 3a to determine the numbers of agents who remained in the firm at different points in time, and we take the average between the month-begin and month-end employment numbers to capture the number of agents present throughout a given month.

estimated administrative cost of about $88,982, and a net present value of the mentorship program equal to approximately $850,000. It is possible that this is a lower bound on long-term value from the program, as increasing frontline sales worker productivity may allow sales managers to have larger spans of control (Espinosa and Stanton, 2021).

6.5 External validity

Our study contributes to a first wave of empirical evidence on mentorship, and our setting has several strengths in terms of external validity. As part of the first wave of empirical evidence, we made decisions to give us the best chance to empirically test the theory, ensuring high internal validity (List, 2020). Despite these choices, performing our experiment in the field provides several strengths in terms of applying our findings to other settings.

First, the participants in our study appear approximately representative of the population they are drawn from on multiple dimensions, and they are representative of the broader population of workers in the United States. The participants in our study are new workers at a representative call center in Utah. In terms of hourly earnings and gender composition, our study participants are similar to the national- and state-level workers in similar occupations (customer service representatives in SOC code 43405, telemarketers in SOC code 41904, and miscellaneous sales representatives in SOC code 41309), based on data from the 2015–2019 5-year American Community Survey. A comparison of hourly earnings shows that mentor-eligible agents' earnings are quite similar to those in similar roles nationally and at the state level. The average hourly earnings among agents in our sample was about $21. The average worker in similar roles at the national-level earns approximately $23 per hour, and the average state-level worker in these occupations earns about $20 per hour.[33]

As shown in Table 1, approximately 42% of the mentor-eligible agents were female, which is below the national- and state-level averages of 61% and 59%, respectively. However, if we exclude customer service call center jobs and only focus on telemarketers and sales representatives, our sample is more gender diverse than the national- and state-level averages (36% and 32% female, respectively). The average age of the mentor-eligible agents was about 23 years old. The average

[33]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.

age of seasoned veterans was approximately 26 years old. Thus, the agents in our sample are younger than the average age of workers in similar occupations nationally (37 years old) and in the state in which the call centers are located (33 years old). The study firm's workforce management group makes a concerted effort to recruit and hire workers from a broad range of demographic and socioeconomic groups. In general, our study sample appears to be fairly representative of the workforce in general, although because this is an entry-level job, ages are lower than average. To understand how our setting compares to broader occupations, note that among all individuals working for wages in any occupation in the state, average and median hourly earnings are about \$26 and \$19, respectively, with an average age of 37. Forty-six percent of those working for wages in the ACS are female.

Second, our study experienced relatively little selection concerns due to non-compliance. Specifically, protocol non-compliance was uncommon, as mentors and protégés completed about 82% of their scheduled meetings. In Section 5.3, we show that our results are robust to removing non-compliers (i.e., those who never met with their mentor) from the sample. In general, however, meeting participation rates were high, as detailed in Section 6.3.1, which was expected given that the firm scheduled meetings to reduce coordination difficulties and compensated the mentors for each completed meeting.

Third, the task that we asked agents to perform—reflecting on their work, sharing these thoughts with their mentors, and acting on their mentors' advice—was a natural extension of their day-to-day activities. Whereas other experiments might ask participants to perform tasks that are wholly unrelated to their job and current knowledge set, our mentorship protocol provided mentors and protégés with a structured approach to discuss and learn from job-related successes and struggles.

Fourth, our intervention was done at scale, as all new hires within our study firm participated in this large-scale mentorship program, subject to the supply of mentors. Pilot programs often have different features than large-scale programs. For example, a pilot program may benefit from using only the best mentors. In this case, the large-scale program would likely have different results because of the supply of high-quality mentors. These scaling complications are less of a concern in our setting because the mentorship program was implemented as part of the firm's regular hiring process. Hence, our two main takeaways—(1) that formal mentorship is beneficial and (2) that those who benefit the most from mentorship might opt out if given the choice—provide clear implications for organizations considering large-scale mentorship programs.

As a final point regarding external validity, we consider our findings in relation to those in Sandvik et al. (2020). We estimate that mentorship increases newly hired agents' revenue-per-call by 10.5% (see Column (2) of Table 2). Sandvik et al. (2020) implemented a "structured-meetings" protocol that called for a structured conversation early in the week, followed by a second, unstructured conversation over lunch at the end of the week. They find revenue gains in excess of 20% across the entire population of seasoned agents. While the protocols in the two experiments are similar—e.g., each randomly paired sales agents together and asked them to discuss job-specific struggles and successes over four weeks—they differ in several aspects that likely contribute to the difference in effect sizes. Whereas Sandvik et al. (2020) paired together two experienced sales agents, the mentorship program studied here pairs together a newly hired sales agent with a seasoned sales agent, and new employees potentially benefit from peer-effects differently than do existing employees. For example, mentorship may help newly hired sales agents learn how to close entry-level deals (i.e., selling a baseline product), whereas knowledge exchange among veterans may help agents learn how to close high-level deals (i.e., up-selling to premium services packages). In addition, the intervention in Sandvik et al. (2020) was approximately twice as intensive as that described here due to the second, unstructured conversation at the end of each week of treatment. Finally, the mentoring protocol clearly specifies who will be providing information (the mentor), and who will be receiving said information (the protégé), whereas Sandvik et al. (2020) randomly paired employees and treated them as equals. In other words, in the prior study both individuals were potential givers and receivers of knowledge, which may have also contributed to the differing treatment effect sizes.

7 Conclusion

Employees gain human capital through learning on the job. Formal education and even past experience cannot provide individuals with all of the knowledge they need to succeed at their workplace because of the need for firm-specific information or hands-on training. To fill this need, companies employ a host of strategies including on-board training, mentorship programs, and encouraging peer-to-peer learning (Sandvik et al., 2020). The evaluation of these programs, however, is often difficult or impossible using observational data due to inherent selection problems. We provide novel evaluation of a formal mentorship program for new hires by running an experiment to (1) separate treatment and selection issues and (2) identify heterogeneous treatment effects based on selection.

We find that formal mentorship programs can have a large positive effect on productivity. Specifically, we find that individuals in a sales firm who were randomly assigned a mentor had revenues that were 19% higher than agents randomly not assigned a mentor in the first two months on the job. In contrast, we find that agents that opted into the mentorship program and received a mentor did not have higher revenues than agents that also opted in but did not receive a mentor. This finding underscores the practical importance of the potential sources of selection bias from heterogeneous treatment effects. We find that the selection bias in our context is large. Agents who selected into the program, but (randomly) did not receive a mentor, had revenues that were over 30% higher than agents who opted out. Together, these findings suggest that observational investigations may have suggested that the program increased revenues by 30% for the group that opted in even though in truth the effect for that group was zero.

We find that the program would have had the largest effect for agents that opted out, who had lower average productivity metrics than agents who opted in but did not receive a mentor. On-the-job training, therefore, may have the largest impact when it is implemented broadly rather than to a small group that opt in. In this setting, these findings also suggest that on-the-job training is a substitute for ability. Said differently, on-the-job training can help the agents that struggle the most, but these agents might be the *least* likely to seek out the resources for improvement. This selection effect is related to the wellness program participation patterns found in Jones et al. (2019) and the site selection bias identified in Allcott (2015), though the latter estimated a positive relationship between selection into participation and treatment effects, whereas we estimate a negative relationship. There are many reasons why workers may choose to opt out of programs that could improve their productivity. First, those that are struggling may not want to admit their difficulties and may be less likely to ask for help. Second, those that are struggling may not know what they do not know—i.e., they may not realize they are struggling or understand how training/mentorship programs could help them. Finally, those that are struggling may be the least engaged, which could jointly explain the low performance and low uptake of the program in our setting. Taken together, our results suggest that productivity-improving information may be hidden from some workers, but managerial interventions can help trigger the dissemination of it.

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 and warrant future investigations. For example, we find that on-the-job training is a substitute for ability. This may not be a general phenomenon, and in other contexts training may lift productivity most for the best workers. Similarly, we find that lower ability agents tended to opt out of receiving a mentor. In other settings, however, there could be no correlation or the best agents may opt out of mentorship. While the present study uses workers' choice to opt in or opt out of formal mentorship as a fulcrum to evaluate the efficacy of formal workplace mentorship programs, the same RCT approach could potentially shed light on a host of alternative intra-firm management practice issues facing practitioners and scholars. As the number of organizations awash with HRM data grows, so do the opportunities for scholars to assist in the design, interpretation, and value-creation around disentangling causal relationships. We believe our design will be useful for guiding future researchers in embarking on these experiments.

References

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.

Athey, Susan, Christopher Avery, Peter Zemsky. 2000. Mentoring and diversity. *American Economic Review* **90**(4) 765–786.

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. 2009. Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica* **77**(4) 1047–1094.

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.

Bartel, Ann, Casey Ichniowski, Kathryn Shaw. 2004. Using "insider econometrics" to study productivity. *American Economic Review* **94**(2) 217–223.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, John Van Reenen. 2014. Jeea-fbbva lecture 2013: The new empirical economics of management. *Journal of the European Economic Association* **12**(4) 835–876.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, John Van Reenen. 2015. Does management matter in schools? *The Economic Journal* **125**(584) 647–674.

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

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.

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

Chandrasekhar, Arun G, Benjamin Golub, He Yang. 2018. Signaling, shame, and silence in social learning. Tech. rep., National Bureau of Economic Research.

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

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, 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.

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* .

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.

Imbens, Guido W, Jeffrey M Wooldridge. 2009. Recent developments in the econometrics of program evaluation. *Journal of economic literature* **47**(1) 5–86.

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.

Kahn, Lisa B, Fabian Lange. 2014. Employer learning, productivity, and the earnings distribution: Evidence from performance measures. *The Review of Economic Studies* **81**(4) 1575–1613.

Kram, Kathy E. 1988. *Mentoring at work: Developmental relationships in organizational life.*. University Press of America.

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.

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.

Mills, Joyce F, Anna C Mullins. 2008. The california nurse mentor project: Every nurse deserves a mentor. *Nursing Economics* **26**(5) 310.

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

Payne, Stephanie C, Ann H Huffman. 2005. A longitudinal examination of the influence of mentoring on organizational commitment and turnover. *Academy of Management Journal* **48**(1) 158–168.

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.

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* .

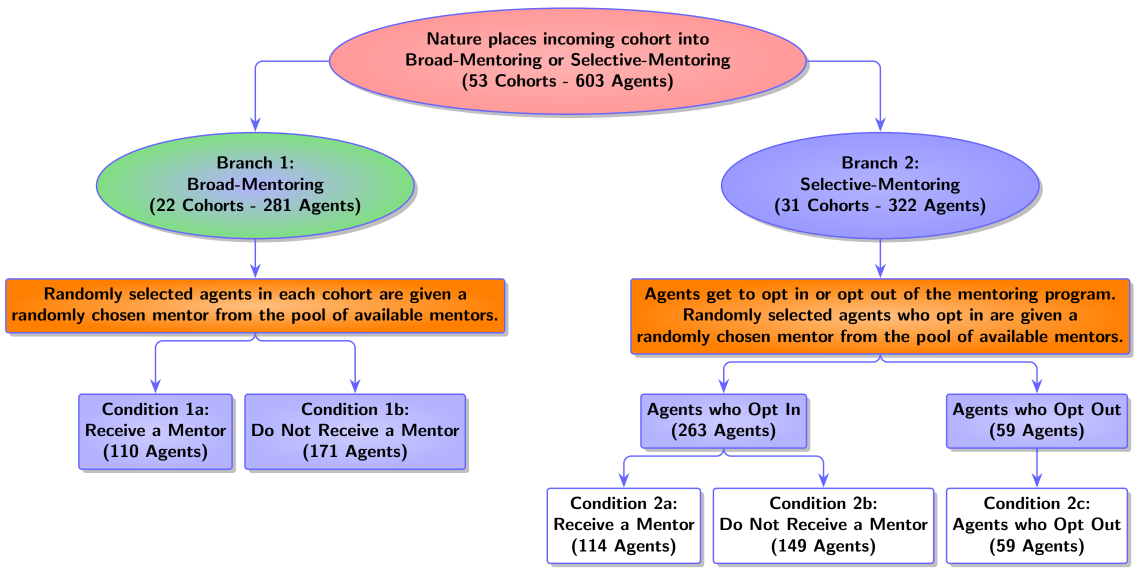
Sandvik, Jason J, Richard E Saouma, Nathan T Seegert, Christopher T Stanton. 2020. Workplace knowledge flows. *The Quarterly Journal of Economics* **135**(3) 1635–1680.

Shaw, Kathryn, Edward P Lazear. 2008. Tenure and output. *Labour Economics* **15**(4) 704–723.

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

Figures and Tables

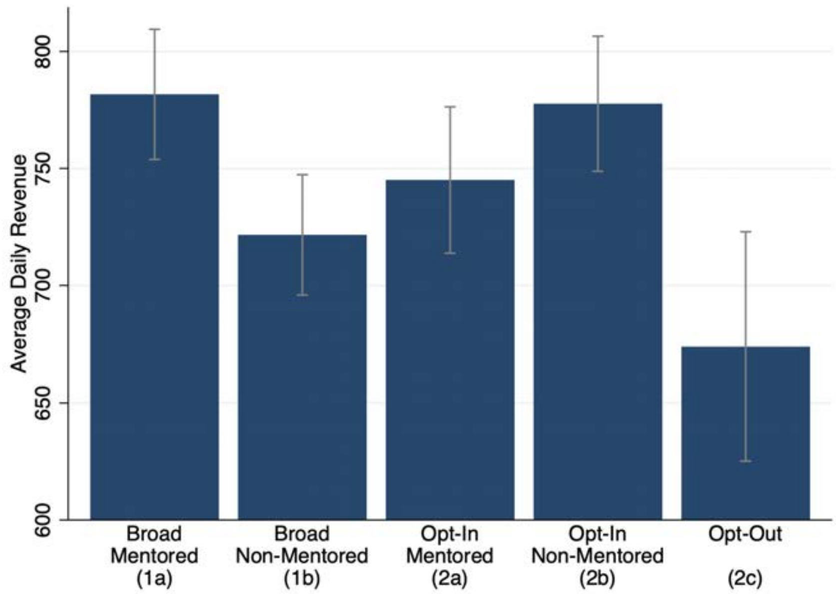
Figure 1: Allocation of Cohorts and Agents to Treatment Conditions



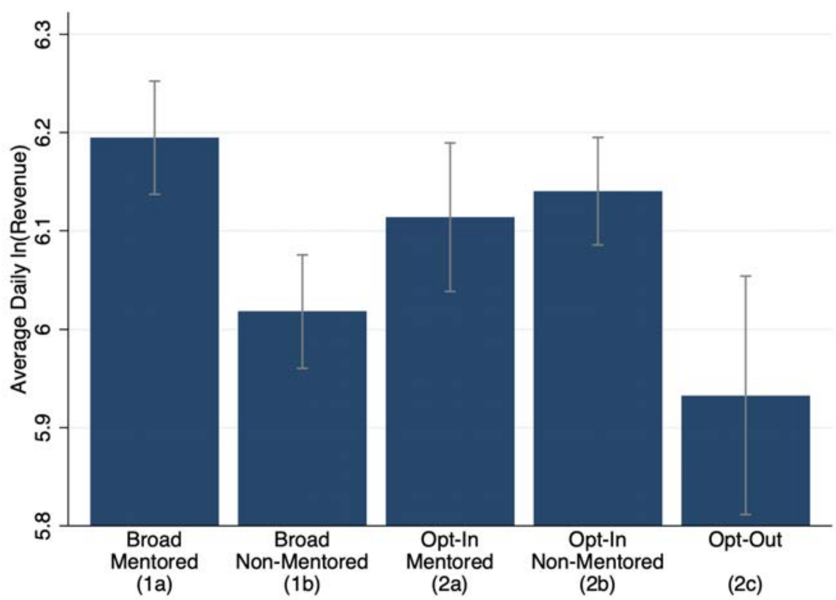
Note: This figure displays the allocation of cohorts to either Broad-Mentoring or Selective-Mentoring, our first level of variation. It then shows the allocation of agents into individual different treatment conditions, our second level of variation.

Figure 2: Effect of Mentoring on Productivity

(a) Average Daily Revenue



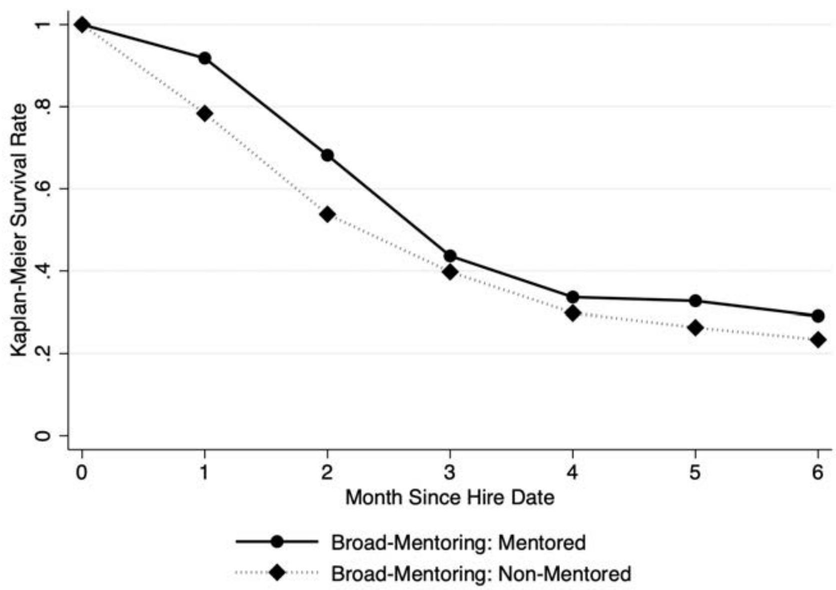
(b) Average Daily ln(Revenue)



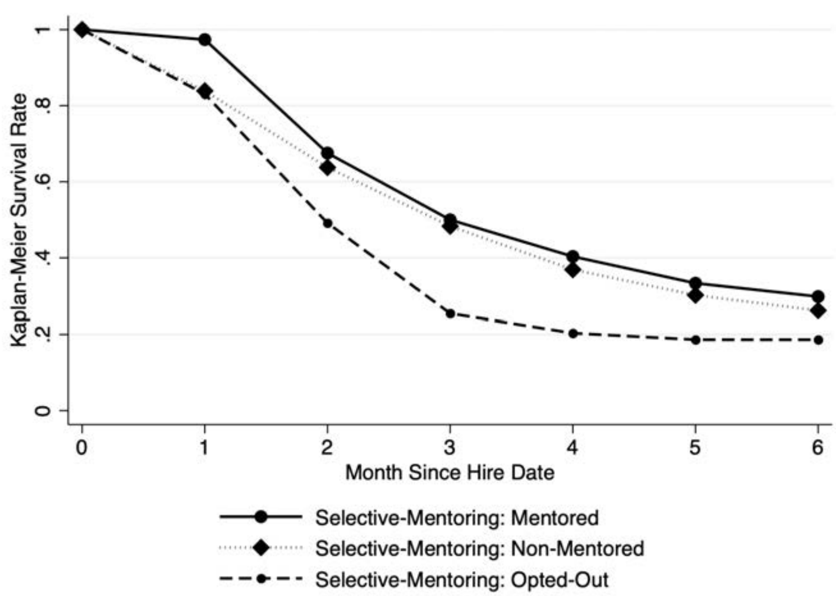
Note: These figures plot the average daily revenue and log revenue, respectively, and 95% confidence intervals for agents' first two months on the sales floor, split by treatment condition. Before we aggregate revenue amounts, we net out cohort fixed effects and then we add in the average productivity 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. Figure (a): (1a)=(1b) p-value<0.001; (2a)=(2b) p-value=0.010; (2b)=(2c) p-value<0.001; (1a)-(1b)=(2a)-(2b) p-value<0.001. Figure (b): (1a)=(1b) p-value<0.001; (2a)=(2b) p-value=0.452; (2b)=(2c) p-value<0.001; (1a)-(1b)=(2a)-(2b) p-value<0.001. Similar bar charts are displayed in Figure A.1 to capture agents' productivity during months 3–6 on the sales floor.

Figure 3: Effect of Mentoring on Retention

(a) Agents in Broad-Mentoring



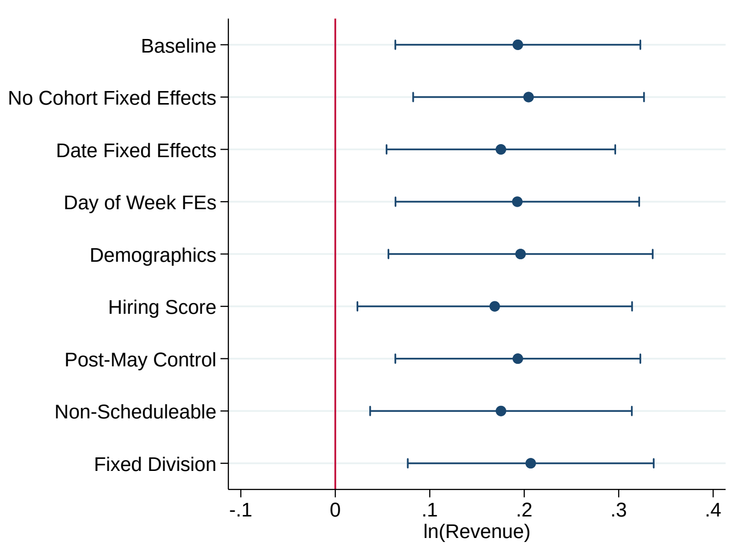
(b) Agents in Selective-Mentoring



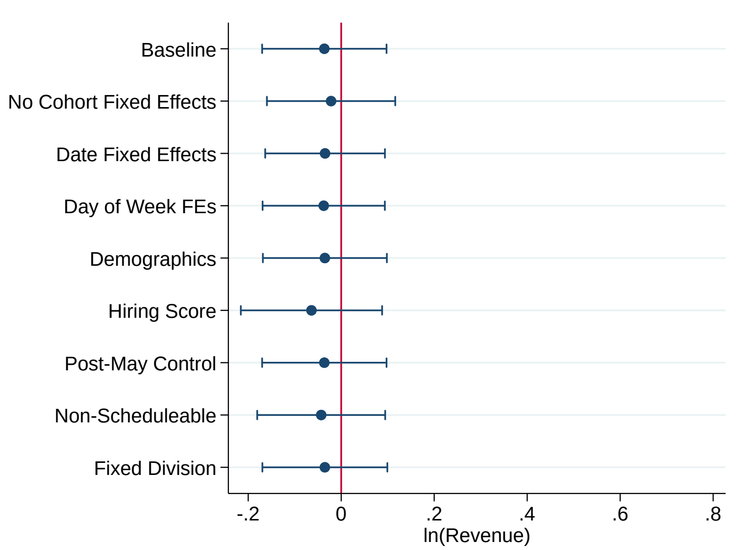
Note: Figure (a) plots Kaplan-Meier survival rates over time for agents in Broad-Mentoring, and Figure (b) considers those in Selective-Mentoring. The survival rate estimator considers a starting point, in our case an agent's hire date, and then, from that time, displays the fraction of agents that remain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

Figure 4: Robustness of the Effect of Mentoring on Revenue Generation

(a) *Mentored* Coefficient: Broad Mentoring

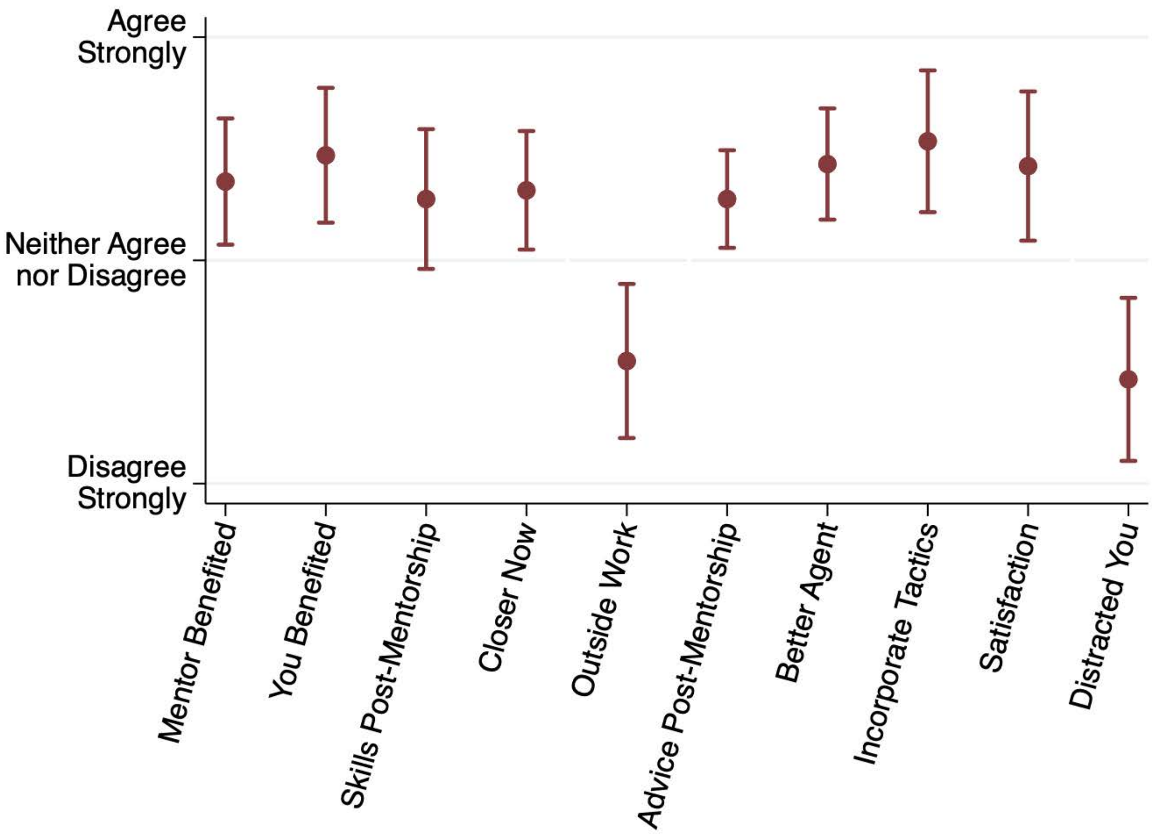


(b) *Mentored* Coefficient: Selective-Mentoring



Note: These figures plot the regression coefficients (and 95% confidence intervals) on *Mentored* from Equation (1) among agents in Broad-Mentoring in Figure (a) and Selective-Mentoring in Figure (b). The "Baseline" estimation replicates the result from Column (1) in Panel A of Table 2 (Table 3) among Broad-Mentoring (Selective-Mentroring) agents. The second estimation excludes cohort fixed effects. The third includes date fixed effects, and the fourth includes day of the week fixed effects. The fifth estimation includes controls for the agent's demographic characteristics, and the sixth controls for the agent's hiring score. The seventh estimation includes a dummy variable that indicates whether the agent's cohort was hired after the May 27th pre-registered start date. The eighth estimation removes agents who were initially assigned a mentor but who were not able to meet with their mentor due to scheduling conflicts. The ninth estimation removes observations in which agents are no longer working in the division in which they were initially hired. We present similar coefficient plots for our other outcome variables in Figures A.2 and A.3.

Figure 5: Responses to Wrap-Up Survey



Note: This figures plots the average values and 95% confidence intervals for responses to the wrap-up 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 wrap-up survey.

Table 1: Balance in Agent Demographics

Panel A: Cohort-Level Balance in Agent Characteristics

	Broad-Mentoring	Selective-Mentoring	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	
N Cohorts	22	31	

Panel B: Cohort-Level Balance in Ex Ante Productivity

	Broad-Mentoring	Selective-Mentoring	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	

Panel C: Agent-Level Balance in Agent Characteristics

	Broad-Mentoring			Selective-Mentoring			
	Mentored	Non-Mentored	p-value	Mentored	Non-Mentored	p-value	Opted-Out
	(1)	(2)	(2)−(1)	(3)	(4)	(4)−(3)	(5)
Age (yrs.)							
Mean	22.51	23.72	0.195	22.28	22.99	0.359	23.30
Std Dev.	4.43	9.09		5.48	6.83		9.08
Woman							
Mean	0.45	0.42	0.688	0.46	0.42	0.518	0.34
Std Dev.	0.50	0.50		0.50	0.49		0.48
Married							
Mean	0.10	0.15	0.209	0.14	0.17	0.456	0.15
Std Dev.	0.30	0.36		0.35	0.38		0.36
Hiring Score							
Mean	0.83	0.84	0.183	0.85	0.86	0.64	0.83
Std Dev.	0.09	0.08		0.09	0.07		0.09
N Agents	110	171		114	149		59

Notes. In Panel A, we average agent characteristics to the cohort-level, then take averages across cohorts. In Panel B, 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 in Panel B estimate the balance in brand-level productivity measures between cohorts in Broad-Mentoring versus Selective-Mentoring. In Panel C, we take averages across agents within each treatment condition.

Table 2: Effect of Broad-Mentoring on Productivity and Retention

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	0.193***	0.105**	0.113**	0.018*	0.112***	0.110
	(0.062)	(0.039)	(0.048)	(0.009)	(0.036)	(0.076)
Cohort FE	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.030	0.038	0.026	0.070	0.069	0.034
Observations	6,744	6,744	6,744	6,744	281	281

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	0.179	0.076	0.104	0.005	0.025	0.036
	(0.108)	(0.064)	(0.095)	(0.007)	(0.070)	(0.071)
Cohort FE	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.043	0.062	0.040	0.098	-0.040	-0.012
Observations	5,993	5,993	5,993	5,993	281	281

Notes. The sample used in Columns (1)–(4) is composed of agent-day productivity data for agents in Broad-Mentoring. Panel A uses data from 269 unique agents, 12 less than the number who were randomized in or out of mentorship, due to turnover during training—i.e., before joining the sales floor. Panel B uses data from 133 unique agents due to attrition in the first two months on the sales floor. *Mentored* equals one for agents who received mentorship, 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. 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 3: Effect of Selective-Mentoring on Productivity and Retention

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	-0.036	-0.050	-0.032	0.003	0.116***	0.008
	(0.066)	(0.044)	(0.059)	(0.005)	(0.037)	(0.058)
Selective Opt-In	0.318***	0.191***	0.209**	0.009	0.051	0.163*
	(0.101)	(0.056)	(0.078)	(0.013)	(0.065)	(0.084)
Cohort FE	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.028	0.053	0.028	0.144	0.092	0.108
Observations	8,393	8,393	8,393	8,393	322	322

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	0.024	-0.024	0.021	0.008	0.009	0.020
	(0.084)	(0.052)	(0.065)	(0.007)	(0.064)	(0.071)
Selective Opt-In	0.050	-0.011	-0.021	0.012	0.246**	0.209**
	(0.102)	(0.041)	(0.081)	(0.013)	(0.091)	(0.081)
Cohort FE	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.055	0.054	0.072	0.065	0.050	0.015
Observations	7,082	7,082	7,082	7,082	322	322

Notes. The sample used in Columns (1)–(4) is composed of agent-day productivity data for agents in Selective-Mentoring. Panel A uses data from 309 unique agents, 13 less than the number who were randomized in or out of mentorship, due to turnover during training—i.e., before joining the sales floor. Panel B uses data from 171 unique agents due to attrition in the first two months on the sales floor. *Mentored* equals one for agents who received mentorship, and zero otherwise. *Selective Opt-In* equals one for agents who chose to opt into the mentorship program—providing them with the possibility, but no guarantee, of receiving mentorship—and zero otherwise. We estimate ordinary least squares regressions in all columns. Columns (5)–(6) use data with a single observation per unique 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. 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 4: Differential Effects of Broad-Mentoring and Selective-Mentoring on Productivity and Retention

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 × Broad	0.241***	0.163***	0.156**	0.014	-0.014	0.100
	(0.089)	(0.058)	(0.076)	(0.011)	(0.051)	(0.095)
Mentored	-0.047	-0.058	-0.043	0.004	0.126***	0.010
	(0.064)	(0.043)	(0.059)	(0.006)	(0.036)	(0.059)
Cohort FE	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.026	0.045	0.025	0.110	0.063	0.057
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 × Broad	0.156	0.099	0.082	-0.004	0.023	0.022
	(0.135)	(0.082)	(0.114)	(0.010)	(0.095)	(0.100)
Mentored	0.023	-0.023	0.022	0.009	0.001	0.014
	(0.084)	(0.052)	(0.065)	(0.007)	(0.065)	(0.071)
Cohort FE	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.050	0.051	0.057	0.083	-0.015	-0.000
Observations	12,270	12,270	12,270	12,270	544	544

Notes. The sample used in Columns (1)–(4) is composed of agent-day productivity data for all mentor-eligible agents other than those who opt-out in the Selective-Mentoring cohorts. Panel A uses data from 523 unique agents, 21 less than the number who were randomized in or out of mentorship, due to turnover during training—i.e., before joining the sales floor. Panel B uses data from 284 unique agents due to attrition in the first two months on the sales floor. *Mentored* equals one for agents who received mentorship, and zero otherwise. *Broad* equals one for agents in Broad-Mentoring and zero for agents in Selective-Mentoring. We estimate ordinary least squares regressions in all columns. Columns (5)–(6) use data with a single observation per unique 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. 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 5: Determinants of Opting Out

Dep. Variable	= 1 if Opted Out			
	(1)	(2)	(3)	(4)
Age	0.011	0.015	0.021	0.014
	(0.022)	(0.022)	(0.028)	(0.029)
Female	-0.388	-0.413	-0.038	-0.079
	(0.291)	(0.285)	(0.312)	(0.319)
Married	-0.156	-0.082	0.144	0.071
	(0.358)	(0.357)	(0.459)	(0.450)
Hiring Score		-4.060**	-4.270**	-4.740**
		(1.704)	(1.861)	(1.924)
SLC			0.006	0.330
			(0.468)	(0.438)
Referral			0.119	0.213
			(0.423)	(0.461)
Extroversion			-0.145	-0.168
			(0.175)	(0.188)
Agreeableness			0.022	-0.030
			(0.206)	(0.201)
Conscientiousness			0.007	0.102
			(0.290)	(0.298)
Emotional Stability			0.047	0.029
			(0.170)	(0.180)
Openness			-0.011	0.085
			(0.254)	(0.251)
Call Center Exp.				0.727
				(0.536)
Sales Experience				-0.043
				(0.561)
Division Fixed Effects			✓	✓
Pse. R-Square	0.007	0.024	0.032	0.066
Observations	322	322	295	295

Notes. This sample is restricted to the 322 agents in the Selective-Mentoring cohorts. 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. Limited personality data reduces the sample sizes in Columns (3) and (4). Columns (3) and (4) include fixed effects for agents in the two largest divisions. In Column (4) we include indicator variables that capture previous work experience. Only agents hired after May 27th were asked about their previous call center and sales experience, so we fill in missing values as zero. We also include a dummy equal to one for agents with missing work experience data, and zero otherwise, as a control in Column (4). Standard errors are clustered by cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Estimated Treatment Effect of Mentoring Among Opt-Out Agents

Panel A: Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence
	(1)	(2)	(3)	(4)
Opt-Out Effect	1.313**	0.734**	0.833*	0.138**
	(0.530)	(0.342)	(0.422)	(0.063)
ΔEffect: Opt-Out − Opt-In	1.015*	0.642*	0.768*	0.151**
	(0.533)	(0.344)	(0.420)	(0.060)

Panel B: Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence
	(1)	(2)	(3)	(4)
Opt-Out Effect	0.916	0.606	0.504	-0.007
	(0.825)	(0.501)	(0.676)	(0.051)
ΔEffect: Opt-Out − Opt-In	0.981	0.676	0.629	-0.012
	(0.825)	(0.496)	(0.680)	(0.052)

Notes. The results in this table show estimates of the treatment effect among agents who opt out of mentorship and a comparison of this treatment effect to the treatment effect among agents in Selective-Mentoring who opt in. To estimate standard errors, we block-bootstrap by cohort (N = 53) over the whole procedure, with 500 bootstrap replications for each column. Standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 7: Testing for Discouragement and Leakage

	Months 1–2			Months 3–6		
	(1)	(2)	(3)	(4)	(5)	(6)
New Hire	-0.421***			-0.206**		
	(0.103)			(0.091)		
New Hire × Broad	0.022	0.070	0.077	0.009	0.101	0.108
	(0.094)	(0.101)	(0.103)	(0.097)	(0.103)	(0.101)
New Hire × Selective	0.076	0.095	0.094	0.033	0.131	0.131
	(0.082)	(0.088)	(0.087)	(0.090)	(0.099)	(0.093)
Mentored × Broad	0.226***	0.231***	0.218***	0.191**	0.164*	0.144
	(0.071)	(0.071)	(0.075)	(0.093)	(0.093)	(0.096)
Mentored × Selective	-0.035	-0.038	-0.026	0.010	0.018	0.052
	(0.074)	(0.074)	(0.073)	(0.089)	(0.087)	(0.085)
Age			0.005			0.007*
			(0.004)			(0.004)
Female			-0.170***			-0.194***
			(0.052)			(0.055)
Married			0.014			0.055
			(0.060)			(0.055)
Hiring Score			0.931**			0.734*
			(0.400)			(0.402)
Division-Location-Date FE	✓	✓	✓	✓	✓	✓
Location-New Hire FE	✓			✓		
Division-Location-New Hire FE		✓	✓		✓	✓
New Hire-Date FE		✓	✓		✓	✓
Adj. R-Square	0.084	0.089	0.100	0.082	0.087	0.100
Observations	41,867	41,858	41,858	43,009	42,998	42,998
$\text{New}_B = 0$, $\text{New}_S = 0$	0.636	0.551	0.545	0.931	0.393	0.344
New_B - $\text{New}_S = 0$	0.548	0.790	0.854	0.805	0.757	0.805
New_B - New_S + $\text{Men}_B = 0$	0.046	0.021	0.025	0.074	0.150	0.189

Notes. This table 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 specifications are described in Section 5.1. The dependent variable is ln(Revenue) in all specifications. In the bottom three rows, New_B stands for new hire in Broad-Mentoring, New_S stands for new hire in Selective-Mentoring, and Men_B stands for mentored in Broad-Mentoring. Standard errors are clustered by agent and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8: Testing for Crowd-Out

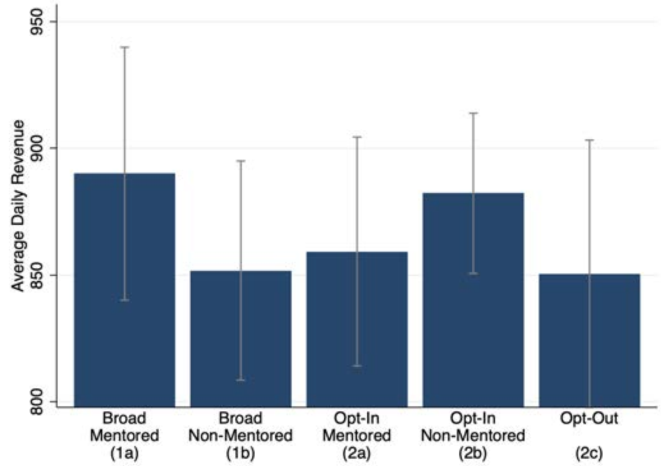
	Months 1–2			Months 3–6		
	(1)	(2)	(3)	(4)	(5)	(6)
New Hire	-0.508***			-0.141**		
	(0.054)			(0.057)		
New Hire × Broad	0.215**	0.217**	0.080	-0.091	-0.100	-0.175
	(0.093)	(0.094)	(0.096)	(0.116)	(0.112)	(0.106)
New Hire × Selective	0.228***	0.252***	0.107	-0.062	-0.054	-0.125
	(0.087)	(0.091)	(0.089)	(0.101)	(0.100)	(0.093)
Mentored × Broad	0.260***	0.251***	0.229***	0.244**	0.221**	0.209**
	(0.075)	(0.074)	(0.076)	(0.104)	(0.104)	(0.102)
Mentored × Selective	-0.031	-0.030	-0.023	0.036	0.046	0.059
	(0.073)	(0.073)	(0.071)	(0.089)	(0.087)	(0.081)
Age			0.002			0.002
			(0.003)			(0.003)
Female			-0.082**			-0.112***
			(0.036)			(0.038)
Married			0.042			0.074
			(0.049)			(0.049)
Hiring Score			0.819***			0.746**
			(0.252)			(0.293)
Division-Location-Date FE	✓	✓	✓	✓	✓	✓
Location-New Hire FE	✓			✓		
Division-Location-New Hire FE		✓	✓		✓	✓
Adj. R-Square	0.160	0.161	0.166	0.125	0.127	0.132
Observations	72,913	72,913	72,913	73,423	73,423	73,423
$\text{New}_B = 0$, $\text{New}_S = 0$	0.016	0.013	0.475	0.716	0.668	0.221
New_B - $\text{New}_S = 0$	0.881	0.696	0.767	0.785	0.654	0.610
New_B - New_S + $\text{Men}_B = 0$	0.004	0.011	0.017	0.021	0.059	0.074

Notes. This table reports tests of the net effect of crowd-out 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 who were hired prior to the onset of the mentoring program; and (3) seasoned veterans who began working at the firm before the onset of the mentorship program. The specifications are described in Section 5.2. The dependent variable is ln(Revenue) in all specifications. In the bottom three rows, New_B stands for new hire in Broad-Mentoring, New_S stands for new hire in Selective-Mentoring, and Men_B stands for mentored in Broad-Mentoring. Standard errors are clustered by agent and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

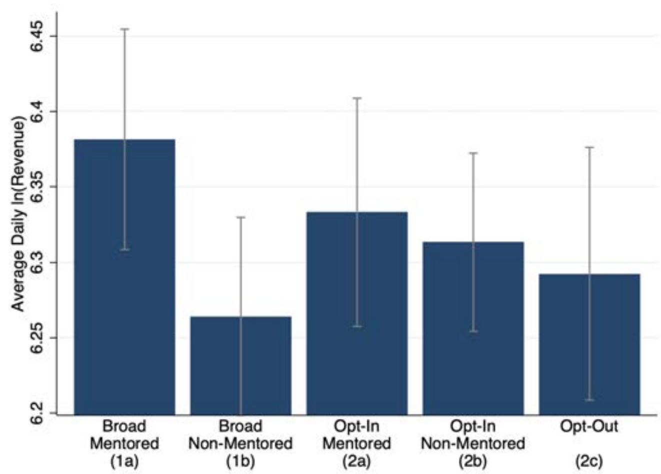
A Appendix Figures and Tables

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

(a) Average Daily Revenue



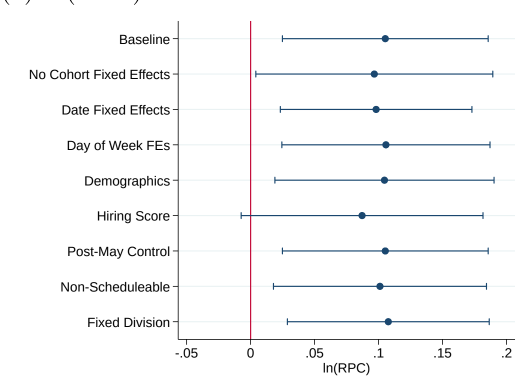
(b) Average Daily ln(Revenue)



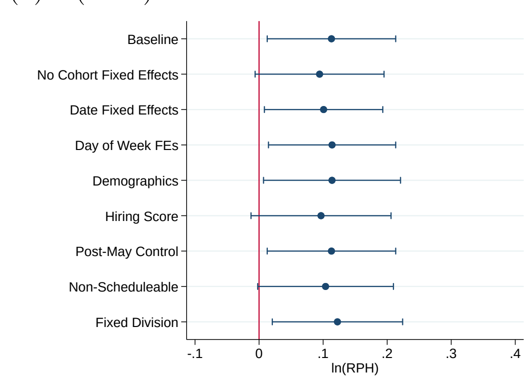
Note: These figures plot the average daily revenue and log revenue, respectively, and 95% confidence intervals for agents' third to sixth months on the sales floor, split by treatment condition. Before we aggregate revenue amounts, we net out cohort fixed effects and then we add in the average across all agents that did not receive a mentor as a baseline.

Figure A.2: Coefficient Plots: Effect of Broad-Mentoring on Productivity in Months 1–2

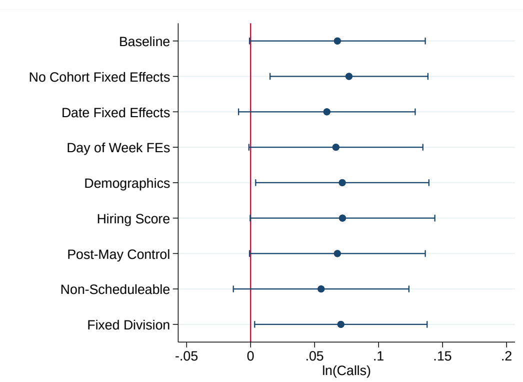
(a) ln(RPC)



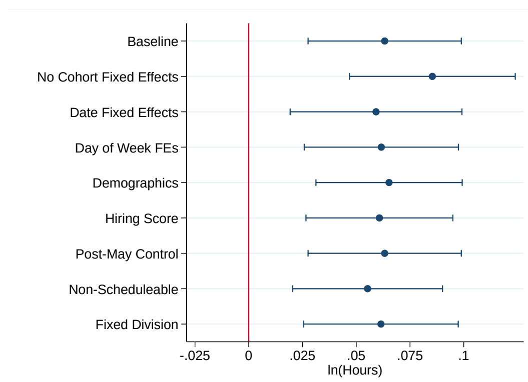
(b) ln(RPH)



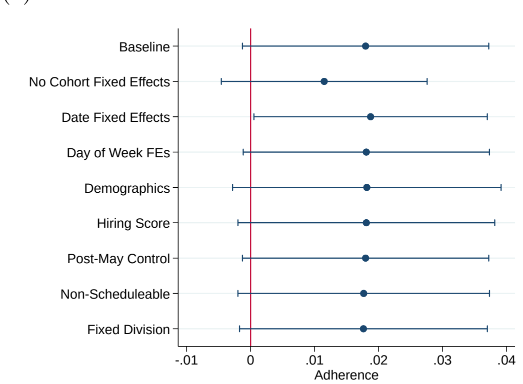
(c) ln(Calls)



(d) ln(Hours)



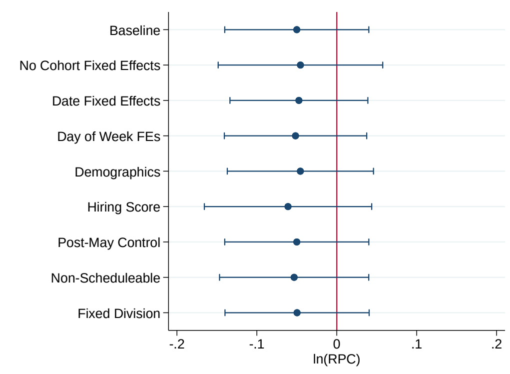
(e) Adherence



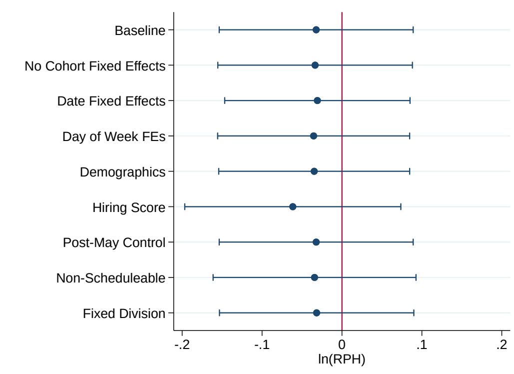
Notes: These figures plot the regression coefficients (and 95% confidence intervals) on *Mentored* from Equation (1) among agents in Broad-Mentoring. The "Baseline" estimations come from an estimation of Equation (1). The second estimation excludes cohort fixed effects. The third includes date fixed effects, and the fourth includes day of the week fixed effects. The fifth estimation includes controls for the agent's demographic characteristics, and the sixth controls for the agent's hiring score. The seventh estimation includes a dummy variable that indicates whether the agent's cohort was hired after the May 27th pre-registered start date. The eighth estimation removes agents who were initially assigned a mentor but who were not able to meet with their mentor due to scheduling conflicts. The ninth estimation removes observations in which agents are no longer working in the division in which they were initially hired.

Figure A.3: Coefficient Plots: Effect of Selective-Mentoring on Productivity in Months 1–2

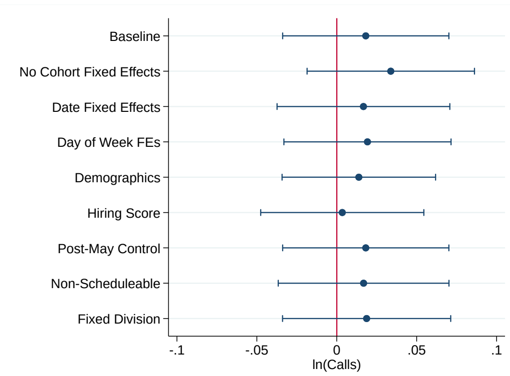
(a) ln(RPC)



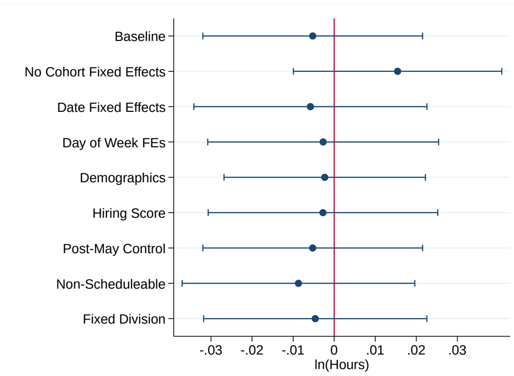
(b) ln(RPH)



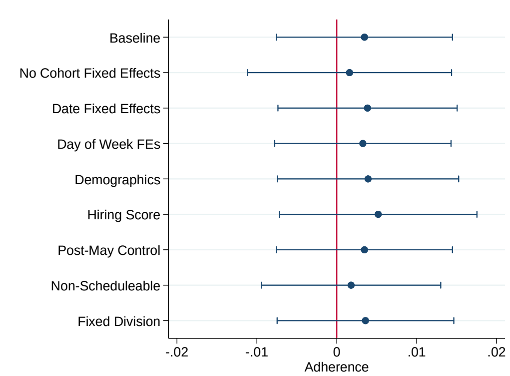
(c) ln(Calls)



(d) ln(Hours)



(e) Adherence



Notes: These figures plot the regression coefficients (and 95% confidence intervals) on *Mentored* from Equation (1) among agents in Selective-Mentoring. The "Baseline" estimations come from an estimation of Equation (2). The second estimation excludes cohort fixed effects. The third includes date fixed effects, and the fourth includes day of the week fixed effects. The fifth estimation includes controls for the agent's demographic characteristics, and the sixth controls for the agent's hiring score. The seventh estimation includes a dummy variable that indicates whether the agent's cohort was hired after the May 27th pre-registered start date. The eighth estimation removes agents who were initially assigned a mentor but who were not able to meet with their mentor due to scheduling conflicts. The ninth estimation removes observations in which agents are no longer working in the division in which they were initially hired.

Table A.1: Balance in Mentor Demographics

	Broad-Mentoring	Selective-Mentoring	p-value
	(1)	(2)	(2)−(1)
Mentor Age (yrs.)			
Mean	24.09	23.70	0.531
Std Dev.	4.72	4.56	
Mentor Woman			
Mean	0.28	0.23	0.358
Std Dev.	0.45	0.42	
Mentor Married			
Mean	0.07	0.17	0.031
Std Dev.	0.26	0.37	
Mentor Tenure			
Mean	1.33	1.22	0.396
Std Dev.	0.84	1.10	
Number of Protégés	110	114	

Notes. In this table we report average characteristics of the agents who mentored protégés in Broad-Mentoring in Column (1) and of the agents who mentored protégés in Selective-Mentoring in Column (2). The p-values from the difference-in-means tests are reported in the rightmost column. Mentors were not designated exclusively to either of the mentoring branches. In other words, a mentor's first protégé could have been assigned to Broad-Mentoring, whereas their second protégé could have been assigned to Selective-Mentoring. Mentors were never informed as to whether their protégés were from Broad-Mentoring or Selective-Mentoring cohorts.

Table A.2: Effect of Broad-Mentoring on Labor Supply

Panel A: Labor Supply in Months 1–2

	ln(Calls)	ln(Hours)	ln(Calls/Hour)
	(1)	(2)	(3)
Mentored	0.068*	0.063***	0.005
	(0.033)	(0.017)	(0.025)
Cohort FE	✓	✓	✓
Adj. R-Square	0.096	0.078	0.099
Observations	6,744	6,744	6,744

Panel B: Labor Supply in Months 3–6

	ln(Calls)	ln(Hours)	ln(Calls/Hour)
	(1)	(2)	(3)
Mentored	0.082*	0.060**	0.022
	(0.046)	(0.023)	(0.037)
Cohort FE	✓	✓	✓
Adj. R-Square	0.089	0.069	0.154
Observations	5,993	5,993	5,993

Notes. The sample is composed of agent-day productivity data for agents in Broad-Mentoring. Panel A uses data from 269 unique agents, 12 less than the number who were randomized in or out of mentorship, due to turnover during training—i.e., before joining the sales floor. Panel B uses data from 133 unique agents due to attrition in the first two months on the sales floor. *Mentored* equals one for agents who received mentorship, 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.3: Effect of Selective-Mentoring on Labor Supply

Panel A: Labor Supply in Months 1–2

	ln(Calls)	ln(Hours)	ln(Calls/Hour)
	(1)	(2)	(3)
Mentored	0.018	-0.005	0.023
	(0.025)	(0.013)	(0.025)
Selective Opt-In	0.072	0.078**	-0.006
	(0.047)	(0.037)	(0.029)
Cohort FE	✓	✓	✓
Adj. R-Square	0.146	0.082	0.235
Observations	8,393	8,393	8,393

Panel B: Labor Supply in Months 3–6

	ln(Calls)	ln(Hours)	ln(Calls/Hour)
	(1)	(2)	(3)
Mentored	0.034	-0.012	0.046
	(0.044)	(0.040)	(0.037)
Selective Opt-In	0.055	0.073	-0.019
	(0.073)	(0.060)	(0.056)
Cohort FE	✓	✓	✓
Adj. R-Square	0.125	0.059	0.239
Observations	7,082	7,082	7,082

Notes. The sample is composed of agent-day productivity data for agents in Selective-Mentoring. Panel A uses data from 309 unique agents, 13 less than the number who were randomized in or out of mentorship, due to turnover during training—i.e., before joining the sales floor. Panel B uses data from 171 unique agents due to attrition in the first two months on the sales floor. *Mentored* equals one for agents who received mentorship, and zero otherwise. *Selective Opt-In* equals one for agents who chose to opt into the mentorship program—providing them with the possibility, but no guarantee, of receiving mentorship—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.4: Differential Effects of Broad-Mentoring and Selective-Mentoring on Labor Supply

Panel A: Labor Supply in Months 1–2

	ln(Calls)	ln(Hours)	ln(Calls/Hour)
	(1)	(2)	(3)
Mentored × Broad	0.050	0.067***	-0.017
	(0.042)	(0.021)	(0.036)
Mentored	0.018	-0.004	0.022
	(0.026)	(0.012)	(0.026)
Cohort FE	✓	✓	✓
Adj. R-Square	0.118	0.073	0.161
Observations	13,779	13,779	13,779

Panel B: Labor Supply in Months 3–6

	ln(Calls)	ln(Hours)	ln(Calls/Hour)
	(1)	(2)	(3)
Mentored × Broad	0.051	0.074	-0.023
	(0.064)	(0.046)	(0.052)
Mentored	0.032	-0.014	0.045
	(0.044)	(0.040)	(0.037)
Cohort FE	✓	✓	✓
Adj. R-Square	0.115	0.069	0.208
Observations	12,270	12,270	12,270

Notes. The sample is composed of agent-day productivity data for all mentor-eligible agents other than those who opt-out in the Selective-Mentoring cohorts. Panel A uses data from 523 unique agents, 21 less than the number who were randomized in or out of mentorship, due to turnover during training—i.e., before joining the sales floor. Panel B uses data from 284 unique agents due to attrition in the first two months on the sales floor. *Mentored* equals one for agents who received mentorship, and zero otherwise. *Broad* equals one for agents in Broad-Mentoring and zero for agents in Selective-Mentoring. 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: Effect of Mentoring on Productivity by Predicted Opt-Out

Panel A: Productivity in Months 1–2

	ln(Revenue)			ln(RPC)			ln(RPH)		
	$High_{Opt}$	Low_{Opt}	All	$High_{Opt}$	Low_{Opt}	All	$High_{Opt}$	Low_{Opt}	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.395***	0.009	0.018	0.232***	-0.022	-0.006	0.243**	-0.012	0.001
	(0.131)	(0.080)	(0.076)	(0.071)	(0.051)	(0.052)	(0.099)	(0.066)	(0.063)
Mentored × $High_{Opt}$			0.417***			0.264***			0.266**
			(0.139)			(0.081)			(0.107)
$High_{Opt}$			-0.125			-0.035			-0.047
			(0.110)			(0.064)			(0.078)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.055	0.035	0.036	0.053	0.044	0.044	0.038	0.035	0.031
Observations	2,193	4,128	6,321	2,193	4,128	6,321	2,193	4,128	6,321

Panel B: Productivity in Months 3–6

	ln(Revenue)			ln(RPC)			ln(RPH)		
	$High_{Opt}$	Low_{Opt}	All	$High_{Opt}$	Low_{Opt}	All	$High_{Opt}$	Low_{Opt}	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.247	0.221	0.154	0.133	0.154	0.081	0.093	0.166	0.113
	(0.193)	(0.166)	(0.147)	(0.146)	(0.104)	(0.090)	(0.154)	(0.145)	(0.122)
Mentored × $High_{Opt}$			0.146			0.033			0.043
			(0.202)			(0.109)			(0.135)
$High_{Opt}$			0.064			0.083			0.086
			(0.147)			(0.095)			(0.111)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.083	0.051	0.050	0.096	0.059	0.068	0.059	0.048	0.045
Observations	2,034	3,526	5,561	2,034	3,526	5,561	2,034	3,526	5,561

Notes. The sample consists of agents in Broad-Mentoring for whom we are able to estimate an opt-out propensity score (i.e., they have non-missing demographic and hiring score data). We estimate agents' opt-out propensity scores via the *psmatch2* command in Stata and by matching the observable characteristics of agents in Broad-Mentoring to those of agents in Selective-Mentoring who opt out. After estimating this propensity score, we place Broad-Mentoring agents into $High_{Opt}$, if their propensity score of opting out is in the top 30% of the propensity score distribution, and we place agents with a propensity score in the bottom 70% into Low_{Opt}, indicating that they had a low likelihood to opt out. We use these thresholds to create sample sizes that are analogous to the actual opt-in versus opt-out groups and are still large enough for reliable inference. We then estimate Equation (1) within these subsets of the data with either *ln(Revenue)*, *ln(RPC)*, or *ln(RPH)* as the independent variable. To determine if the effect of mentorship is significantly different between the $High_{Opt}$ and Low_{Opt} agents, we include a one-zero indicator for $High_{Opt}$ into the model, along with the interaction between this and *Mentored*. 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: Lee Bounds Estimates of the Effect of Mentoring on Productivity

Panel A: Broad-Mentoring, Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	ln(Calls)	ln(Hours)	Adherence	Conversion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mentored_{lower}	0.101**	-0.005	0.001	0.028**	0.064***	0.002	-0.009***
	(0.040)	(0.028)	(0.031)	(0.013)	(0.008)	(0.003)	(0.004)
Mentored_{upper}	0.513***	0.294***	0.325***	0.170***	0.141***	0.036***	0.028***
	(0.035)	(0.026)	(0.028)	(0.013)	(0.008)	(0.003)	(0.004)
Observations	7,816	7,816	7,816	7,816	7,816	7,816	7,816

Panel B: Broad-Mentoring, Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	ln(Calls)	ln(Hours)	Adherence	Conversion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mentored_{lower}	0.166***	0.051**	0.065**	0.073***	0.067***	-0.008***	-0.003
	(0.035)	(0.025)	(0.027)	(0.015)	(0.011)	(0.003)	(0.004)
Mentored_{upper}	0.386***	0.216***	0.237***	0.133***	0.132***	0.008***	0.020***
	(0.033)	(0.025)	(0.027)	(0.016)	(0.012)	(0.002)	(0.004)
Observations	6,592	6,592	6,592	6,592	6,592	6,592	6,592

Panel C: Selective-Mentoring, Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	ln(Calls)	ln(Hours)	Adherence	Conversion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mentored_{lower}	-0.051	-0.062*	-0.054	0.022	0.007	-0.002	-0.011***
	(0.058)	(0.035)	(0.042)	(0.020)	(0.014)	(0.005)	(0.004)
Mentored_{upper}	-0.013	-0.036	-0.026	0.034***	0.015*	0.002	-0.006
	(0.038)	(0.028)	(0.030)	(0.013)	(0.008)	(0.003)	(0.005)
Observations	7,756	7,756	7,756	7,756	7,756	7,756	7,756

Panel D: Selective-Mentoring, Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	ln(Calls)	ln(Hours)	Adherence	Conversion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mentored_{lower}	0.002	-0.045	-0.009	0.032**	-0.001	-0.006**	-0.012***
	(0.061)	(0.037)	(0.045)	(0.016)	(0.013)	(0.003)	(0.004)
Mentored_{upper}	0.203***	0.086***	0.143***	0.093***	0.050***	0.006***	0.006
	(0.037)	(0.026)	(0.029)	(0.015)	(0.011)	(0.002)	(0.004)
Observations	6,872	6,872	6,872	6,872	6,872	6,872	6,872

Notes. Panels A and B use agent-day productivity data for agents in Broad-Mentoring. Panels C and D use agent-day productivity data for agents in Selective-Mentoring, 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.

Table A.7: Heterogeneous Effects of Broad-Mentoring on Productivity in Months 1–2

	(1)	(2)	(3)	(4)	(5)	(6)
Mentored	0.193***	-0.871*	0.320***	0.231***	0.414	-0.230
	(0.062)	(0.429)	(0.089)	(0.071)	(0.774)	(0.872)
Mentored × Age		0.046**				0.046**
		(0.020)				(0.019)
Age		-0.007				-0.010
		(0.012)				(0.010)
Mentored × Female			-0.329**			-0.367**
			(0.144)			(0.148)
Female			0.075			0.096
			(0.131)			(0.125)
Mentored × Married				-0.277		-0.497*
				(0.250)		(0.278)
Married				0.305		0.366*
				(0.178)		(0.188)
Mentored × Hiring Score					-0.297	-0.538
					(0.944)	(0.680)
Hiring Score					-0.070	0.288
					(0.711)	(0.535)
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.030	0.035	0.032	0.032	0.032	0.044
Observations	6,744	6,744	6,744	6,744	6,321	6,321

Notes. This sample is composed of agent-day productivity data for agents in Broad-Mentoring. The dependent variable in all specifications is *ln(Revenue)*. *Mentored* equals one for agents who received mentorship, and zero otherwise. We estimate ordinary least squares regressions in all columns. Column (1) replicates the result from Column (1) of Table 2. Columns (2)–(5) each separately include a different observable characteristic into the model, along with the interaction between this characteristic and *Mentored*. Column (6) includes all characteristics and interaction terms together in a single model. 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.8: Differences in Worksheet Content

Panel A: Broad-Mentoring versus Selective-Mentoring

	Total Words per Worksheet	Skill Words per Worksheet	Support Words per Worksheet	Other Words per Worksheet
	(1)	(2)	(3)	(4)
Broad-Mentoring	-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

Panel B: High versus Low Opt-Out Likelihood

	Total Words per Worksheet	Skill Words per Worksheet	Support Words per Worksheet	Other Words per Worksheet
	(1)	(2)	(3)	(4)
$\text{High}_{Opt-Out}$	-0.173	0.303	-0.072	-0.404
	(3.654)	(0.418)	(0.120)	(3.475)
Adj. R-Square	-0.012	-0.007	-0.008	-0.012
Observations	85	85	85	85
Mean DV	47.14	4.39	0.504	42.252

Notes. Panel A of this table considers differences in worksheet content between protégés in Broad-Mentoring and those in Selective-Mentoring. Panel B reports the relations between Broad-Mentoring agents' opt-out propensity score ($High_{Opt}$) and the content of their completed Protégé Worksheets. We estimate the opt-out propensity score exactly as described in Section 4.5. Of the 110 agents in Broad-Mentoring who received mentorship, we have completed worksheet content for 93 of them, and we have non-missing demographic data (which allows us to estimate an opt-out propensity score) for 85. 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 and opt-out likelihood. Robust standard errors are reported in parentheses. The mean of the dependent variable is listed below the observation count line in each panel. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

B Main Results with Rehires Included

Table B.1: Balance in Agent Demographics

Panel A: Cohort-Level Balance in Agent Characteristics

	Broad-Mentoring	Selective-Mentoring	p-value
	(1)	(2)	(2)−(1)
Age (yrs.)			
Mean	22.86	23.11	0.696
Std Dev.	2.22	2.26	
Woman			
Mean	0.44	0.39	0.276
Std Dev.	0.15	0.18	
Married			
Mean	0.14	0.15	0.916
Std Dev.	0.10	0.15	
Hiring Score			
Mean	0.84	0.85	0.199
Std Dev.	0.04	0.04	
Rehired			
Mean	0.07	0.08	0.709
Std Dev.	0.08	0.08	
N Cohorts	22	31	

Panel B: Cohort-Level Balance in Ex Ante Productivity

	Broad-Mentoring	Selective-Mentoring	p-value
	(1)	(2)	(2)−(1)
Revenue			
Mean	785.75	814.56	0.419
Std Dev.	101.52	142.04	
RPC			
Mean	48.78	51.10	0.376
Std Dev.	8.00	10.14	
RPH			
Mean	119.58	123.70	0.443
Std Dev.	13.95	22.03	
Calls			
Mean	17.10	17.10	0.994
Std Dev.	0.96	1.25	
Hours			
Mean	6.57	6.59	0.848
Std Dev.	0.39	0.39	
Adherence			
Mean	0.82	0.84	0.209
Std Dev.	0.04	0.04	
Conversion			
Mean	0.23	0.22	0.528
Std Dev.	0.03	0.03	
N Cohorts	22	31	

Panel C: Agent-Level Balance in Agent Characteristics

	Broad-Mentoring			Selective-Mentoring			
	Mentored	Non-Mentored	p-value	Mentored	Non-Mentored	p-value	Opted-Out
	(1)	(2)	(2)−(1)	(3)	(4)	(4)−(3)	(5)
Age (yrs.)							
Mean	22.50	23.63	0.196	22.41	23.04	0.405	23.50
Std Dev.	4.36	8.69		5.52	6.74		8.34
Woman							
Mean	0.44	0.42	0.703	0.44	0.42	0.820	0.33
Std Dev.	0.50	0.50		0.50	0.50		0.47
Married							
Mean	0.10	0.15	0.282	0.15	0.18	0.555	0.13
Std Dev.	0.31	0.36		0.36	0.38		0.33
Hiring Score							
Mean	0.83	0.84	0.135	0.85	0.86	0.636	0.84
Std Dev.	0.09	0.08		0.09	0.07		0.09
Rehired							
Mean	0.04	0.10	0.076	0.04	0.09	0.148	0.18
Std Dev.	0.20	0.30		0.20	0.28		0.39
N Agents	115	190		119	163		72

Notes. In Panel A, we average agent characteristics to the cohort-level, then take averages across cohorts. In Panel B, 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 in Panel B estimate the balance in brand-level productivity measures between cohorts in Broad-Mentoring versus Selective-Mentoring. In Panel C, we take averages across agents within each treatment condition.

Table B.2: Effect of Broad-Mentoring on Productivity

Panel A: Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	ln(Calls)	ln(Hours)	ln(Calls/Hour)	$Tenure_1$	$Tenure_2$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.208***	0.099**	0.123**	0.024**	0.088**	0.068***	0.021	0.106***	0.105
	(0.062)	(0.039)	(0.049)	(0.010)	(0.033)	(0.018)	(0.024)	(0.034)	(0.073)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.027	0.032	0.022	0.075	0.100	0.088	0.101	0.062	0.029
Observations	7,353	7,353	7,353	7,353	7,353	7,353	7,353	305	305

Panel B: Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	ln(Calls)	ln(Hours)	ln(Calls/Hour)	$Tenure_3$	$Tenure_4$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.154	0.064	0.072	0.006	0.072	0.069***	0.003	0.041	0.048
	(0.097)	(0.057)	(0.092)	(0.007)	(0.044)	(0.023)	(0.042)	(0.066)	(0.069)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.039	0.059	0.038	0.093	0.083	0.058	0.145	-0.032	-0.011
Observations	6,224	6,224	6,224	6,224	6,224	6,224	6,224	305	305

Notes. This sample is composed of agent-day productivity data for agents in Broad-Mentoring. *Mentored* equals one for agents who received mentorship, and zero otherwise. We estimate ordinary least squares regressions in all columns. Columns (8)–(9) use data with a single observation per unique 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. 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 B.3: Effect of Selective-Mentoring on Productivity

Panel A: Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	ln(Calls)	ln(Hours)	ln(Calls/Hour)	Tenure_1	Tenure_2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	-0.021	-0.035	-0.021	0.005	0.014	-0.004	0.018	0.120***	0.026
	(0.074)	(0.048)	(0.066)	(0.005)	(0.028)	(0.013)	(0.026)	(0.038)	(0.057)
Selective Opt-In	0.239***	0.115*	0.130*	0.005	0.079*	0.084***	-0.005	0.031	0.069
	(0.086)	(0.061)	(0.075)	(0.012)	(0.041)	(0.030)	(0.039)	(0.058)	(0.079)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.032	0.054	0.033	0.150	0.152	0.081	0.231	0.106	0.107
Observations	9,072	9,072	9,072	9,072	9,072	9,072	9,072	354	354

Panel B: Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	ln(Calls)	ln(Hours)	ln(Calls/Hour)	Tenure_3	Tenure_4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.023	-0.017	0.021	0.010	0.025	-0.012	0.037	0.015	0.022
	(0.087)	(0.054)	(0.067)	(0.006)	(0.048)	(0.038)	(0.042)	(0.063)	(0.073)
Selective Opt-In	0.037	-0.001	-0.008	0.016	0.030	0.048	-0.018	0.129	0.106
	(0.092)	(0.061)	(0.094)	(0.010)	(0.057)	(0.055)	(0.055)	(0.097)	(0.095)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.050	0.054	0.069	0.067	0.109	0.048	0.218	0.039	0.009
Observations	8,021	8,021	8,021	8,021	8,021	8,021	8,021	354	354

Notes. This sample is composed of agent-day productivity data for agents in Selective-Mentoring. *Mentored* equals one for agents who received mentorship, and zero otherwise. *Selective Opt-In* equals one for agents who chose to opt into the mentorship program—providing them with the possibility, but no guarantee, of receiving mentorship—and zero otherwise. We estimate ordinary least squares regressions in all columns. Columns (8)–(9) use data with a single observation per unique 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. 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 B.4: Differential Effects of Broad-Mentoring and Selective-Mentoring on Productivity

Panel A: Productivity in Months 1–2

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	ln(Calls)	ln(Hours)	ln(Calls/Hour)	$Tenure_1$	$Tenure_2$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored × Broad	0.236**	0.141**	0.151*	0.018	0.072*	0.070***	0.002	-0.026	0.077
	(0.096)	(0.062)	(0.081)	(0.011)	(0.042)	(0.021)	(0.035)	(0.050)	(0.093)
Mentored	-0.027	-0.042	-0.028	0.005	0.016	-0.002	0.018	0.131***	0.028
	(0.074)	(0.049)	(0.066)	(0.006)	(0.027)	(0.012)	(0.026)	(0.037)	(0.059)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.025	0.040	0.024	0.118	0.119	0.076	0.156	0.061	0.054
Observations	14,754	14,754	14,754	14,754	14,754	14,754	14,754	587	587

Panel B: Productivity in Months 3–6

	ln(Revenue)	ln(RPC)	ln(RPH)	Adherence	ln(Calls)	ln(Hours)	ln(Calls/Hour)	$Tenure_3$	$Tenure_4$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored × Broad	0.138	0.086	0.055	-0.003	0.048	0.085*	-0.037	0.032	0.031
	(0.129)	(0.078)	(0.113)	(0.010)	(0.065)	(0.045)	(0.058)	(0.092)	(0.100)
Mentored	0.016	-0.022	0.017	0.009	0.024	-0.016	0.040	0.009	0.017
	(0.087)	(0.054)	(0.067)	(0.006)	(0.049)	(0.038)	(0.041)	(0.065)	(0.073)
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.046	0.049	0.055	0.079	0.100	0.060	0.192	-0.008	0.007
Observations	12,881	12,881	12,881	12,881	12,881	12,881	12,881	587	587

Notes. This sample is composed of agent-day productivity data for all mentor-eligible agents other than those who opt-out in the Selective-Mentoring cohorts. *Mentored* equals one for agents who received mentorship, and zero otherwise. *Broad* equals one for agents in Broad-Mentoring and zero for agents in Selective-Mentoring. We estimate ordinary least squares regressions in all columns. Columns (8)–(9) use data with a single observation per unique 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. 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 B.5: Determinants of Opting Out

Dep. Variable	= 1 if Opted Out			
	(1)	(2)	(3)	(4)
Age	0.016	0.019	0.022	0.012
	(0.023)	(0.024)	(0.027)	(0.030)
Female	-0.394	-0.387	0.047	0.026
	(0.248)	(0.245)	(0.308)	(0.319)
Married	-0.485	-0.427	-0.179	-0.232
	(0.372)	(0.372)	(0.448)	(0.423)
Rehire	1.081***	1.198***	1.031**	0.990*
	(0.386)	(0.400)	(0.438)	(0.565)
Hiring Score		-3.385**	-3.661**	-4.183**
		(1.593)	(1.726)	(1.821)
SLC			0.291	0.614
			(0.444)	(0.418)
Referral			-0.146	-0.048
			(0.353)	(0.382)
Extroversion			-0.084	-0.089
			(0.149)	(0.159)
Agreeableness			-0.067	-0.127
			(0.151)	(0.146)
Conscientiousness			-0.089	-0.009
			(0.242)	(0.247)
Emotional Stability			0.081	0.073
			(0.152)	(0.160)
Openness			0.091	0.172
			(0.257)	(0.260)
Call Center Exp.				0.721
				(0.540)
Sales Experience				-0.043
				(0.527)
Division Fixed Effects			✓	✓
Pse. R-Square	0.032	0.045	0.045	0.081
Observations	354	354	322	322

Notes. This sample is restricted to the 354 agents in the Selective-Mentoring cohorts. 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. Limited personality data reduces the sample sizes in Columns (3) and (4). Columns (3) and (4) include fixed effects for agents in the two largest divisions. In Column (4) we include indicator variables that capture previous work experience. Only agents hired after May 27th were asked about their previous call center and sales experience, so we fill in missing values as zero, except for rehired agents who receive value of one for both previous call center experience and sales experience. We also include a dummy equal to one for agents with missing work experience data, and zero otherwise, as a control in Column (4). Standard errors are clustered by cohort and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

C Generalized Method of Moments Estimator

The GMM approach allows us to build in additional flexibility for estimating the opt-out probability, which may have differences across cohorts due to sampling variation. In particular, we form moment conditions that correspond to two equations. The first equation accounts for the opt-in probability among Selective-Mentoring cohorts based on the moment conditions,

$$E[x_i \times (OptIn_i - \exp(x_i\delta)/(1 + \exp(x_i\delta))].$$

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

$$\begin{aligned} E[z_i \times (Y - Cohort_i - \beta_{OptIn}(Broad \times \pi_{OptIn}(x_i, \delta) + SelectiveOptIn) \\ -\beta_{OptInMentored}Mentored(Broad \times \pi_{OptIn}(x_i, \delta) + SelectiveOptIn) \\ -\beta_{OptOutMentored}Broad \times (1 - \pi(x_i, \delta)))]. \end{aligned}$$

The instruments z_i for the second equation are indicators for *Broad* and *Mentored*, *Selective Opt-In*, and *Selective Opt-In* and *Mentored*, along with cohort dummies (in practice we use the within transformation). We estimate the system jointly and allow x_i to include Age, Hiring Score, Female, and a time trend that captures when the cohort was hired and onboarded.

D Documentation

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

Protégé Worksheet (Week 2)

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

Weekly Self-Reflection:
What was your **big win** from Week 1? Does your mentor know this yet?

Do you feel **comfortable reaching out** to your mentor? coach? manager? Why or why not?

Do you feel you have been **provided the tools** you need to be successful? Why or why not?

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

Protégé Worksheet (Week 3)

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

Weekly Self-Reflection:
What was your **big win** from Week 2? Does your mentor know this yet?

How are you **adjusting to your work environment**? What does your team **expect of you**?

Do you understand your **department's compliance rules**? Why or why not?

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

Protégé Worksheet (Week 4)

Protégé: ____

Mentor: ____________________________________ Number of times mentor has reached out: _____

Date: ____

Weekly Self-Reflection:

What was your **big win** from Week 3? Does your mentor know this yet?

What is **something new you have learned** since being on the phones?

Do you understand **’s attendance policy**? What **motivates you** to work hard?

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

E 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 A: 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 B: 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

F 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.

F.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.

F.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.

F.2.1 Intervention Start Date

2019-05-27

F.2.2 Intervention End Date

2019-12-20

F.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.).

F.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.[34] Worker turnover measures whether

[34] 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).

the interventions changed the agents' propensity to leave the firm. Log of completed tenure is a different measure of retention that 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.

F.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:

F.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: Broad-Mentoring or Selective-Mentoring. 40% of the cohorts will be in the Broad-Mentoring group and 60% of the cohorts will be in the Selective-Mentoring group.

F.4.2 Within Cohort Randomization

For cohorts in Broad-Mentoring, new hires will receive a mentor with probability 50%. This will be communicated privately between sales floor staff and the individual workers. Agents in Broad-Mentoring cohorts who do not receive a mentor will not receive communication regarding the program. For cohorts in Selective-Mentoring, 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.

F.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.[35] 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].[36] 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."

F.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.[37] 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.

F.7 Randomization Method

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

F.8 Randomization Unit

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

[35] 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.

[36] The wrap-up 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.

[37] 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.

F.9 Was the treatment clustered?

Yes

F.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.

F.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.

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

Please see design field.

F.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).

F.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.[38] 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:

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

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

[38] As described in Section 6.1, differences in mentor instructions did not have a meaningful impact on the treatment effect of mentorship. Because of this and for brevity, we omit this indicator from the models in our heterogeneous treatment effects tests.

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 proteges 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.).[39] 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.

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 protege. This regression will include mentor fixed effects and mentor tenure.[40]

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

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