

Management Science Internet Appendix to “Should
Human Capital Development Programs be Mandatory
or Voluntary? Evidence from a Field Experiment on
Mentorship”

Jason Sandvik*

Richard Saouma †

Nathan Seegert‡

Christopher Stanton§

*Eller College of Management, University of Arizona

†Eli Broad College of Business, Michigan State University

‡David Eccles School of Business, University of Utah

§Harvard Business School, NBER, and CEPR

Contents

- I.A Alternative Explanations for Treatment Effects 2
 - I.A.1 Experimenter Demand Effects, Hawthorne Effects, Discouragement,
or Information Leakage 2
 - I.A.2 Possibility of Program Crowd-Out 4
 - I.A.3 Retention Effects and Inference About Sales Treatment Effects 4
 - I.A.4 Salience of Differential Treatment 5
 - I.A.5 Belief- and Preference-Based Explanations for Opt-Out Behavior 6
- I.B AEA Pre-Registration Text 22
 - I.B.1 Abstract 22
 - I.B.2 Intervention(s) 22
 - I.B.3 Intervention Start Date 22
 - I.B.4 Intervention End Date 22
 - I.B.5 Primary Outcomes (end points) 22
 - I.B.6 Primary Outcomes (explanation) 22
 - I.B.7 Experimental Design Details 23
 - I.B.8 Cohort Level Randomization 23
 - I.B.9 Within Cohort Randomization 23
 - I.B.10 Compliance Tracking 24
 - I.B.11 Edit June 4, 2019 24
 - I.B.12 Randomization Method 25
 - I.B.13 Randomization Unit 25
 - I.B.14 Was the treatment clustered? 25
 - I.B.15 Sample size: planned number of clusters 25
 - I.B.16 Sample size: planned number of observations 25
 - I.B.17 Sample size (or number of clusters) by treatment arms 25
 - I.B.18 Minimum detectable effect size for main outcomes 25
 - I.B.19 Analysis Plan 25
- I.C Documentation of Instructions to Mentors and Example Worksheet for Structuring Conversations 27
- I.D Pre-Registered Estimation of Heterogeneous Treatment Effects 29
- I.E Worksheet Response Examples 31
- I.F Calculation Details for the Value to the Firm 32

I.A Alternative Explanations for Treatment Effects

In this section, we evaluate threats to the validity of our treatment effect estimates, and we explore a range of possible alternative explanations for differences in treatment effects between the Mandatory- and Voluntary-Conditions.

I.A.1 Experimenter Demand Effects, Hawthorne Effects, Discouragement, or Information Leakage

Our design is a natural field experiment (Harrison and List, 2004), where the participants never met the researchers, limiting experimenter demand effects (i.e., observer bias). To participants, the mentoring program appeared like a normal work activity. While participants were told that outside researchers were analyzing their survey and productivity data, the mentorship program was framed and experienced as a regular part of the firm’s onboarding process. Furthermore, participants were not aware that differences between the Mandatory- and Voluntary-Conditions were the objects of researcher interest.¹ Similarly, Hawthorne effects were not likely in this setting, as sales managers monitor the same performance metrics that we study and provide workers with performance-related feedback. It is unlikely that subject behavior was impacted by the knowledge that outside researchers—with whom the agents never interacted—were tracking their performance. In our prior work with this firm, benign treatments with no productivity impact allowed us to test for Hawthorne and demand effects; we found no evidence of their importance in this setting (Sandvik et al., 2020).

Discouragement and the possible leakage of information due to the design of the program are always important mechanisms to consider in natural field experiments. In our setting, agents who did not receive a mentor may have become discouraged, reducing their performance. Discouragement effects could potentially result in a difference in productivity between mentored and non-mentored agents; we would expect agents who received a mentor to outperform those that did not. In the Voluntary-Condition, where we would expect discouragement effects to be most salient, we find no evidence that agents who opted in and received a mentor outperformed those who did not. In our implementation, we preemptively worked with the company to limit discouragement and information leakage. For example, the protocol called for the staff to privately notify treated workers of their mentor assignment—reducing the salience of unequal treatment and the potential for discouragement among non-mentored workers. We also asked the staff to monitor any complaints or concerns over not being matched to a mentor; no instances of discouragement were ever communicated to the research team. In addition, they never indicated that the content of the mentor-protégé meetings was shared with non-mentored agents.

We formally test the net effect of discouragement and information leakage by comparing the performance of new workers who joined the firm in experimental cohorts with new hires who were not part of the experiment. Under the null of no net discouragement (or encouragement), information leakage, or other SUTVA violations, we would expect the productivity of non-treated workers in experimental cohorts to be equal to the productivity of new hires

¹Subjects were asked to provide informed consent when responding to the new hire survey. The survey was framed around understanding employees’ preferences, work styles, and personality characteristics so that university researchers could help the firm better serve its workforce. The consent protocol did not specify that selection into or out of the mentoring program was the key metric being studied, as this decision was elicited by the firm’s staff.

outside of the experiment. To conduct this test, we compare new hires in experimental cohorts with 1) new hires who were in hold-out cohorts during the time of the experiment and 2) new hires who joined the firm prior to the experiment. Our tests examine the relative performance of new hires versus experienced veterans working in the same divisions at the same time, which enables us to make comparisons over time when sales conditions may differ. This approach removes common time series differences across cohorts, including those due to demand fluctuations, and improves the power of our tests.

We estimate the following model using ordinary least squares:

$$\begin{aligned}
y_{i,t} = & \alpha + \beta_1 \text{New Hire}_i + \beta_2 (\text{New Hire} \times \text{Mandatory})_i \\
& + \beta_3 (\text{New Hire} \times \text{Voluntary})_i \\
& + \beta_4 (\text{New Hire} \times \text{Mandatory} \times \text{Mentored})_i \\
& + \beta_5 (\text{New Hire} \times \text{Voluntary} \times \text{Mentored})_i + \zeta_{j,l,t} + \varepsilon_{i,t},
\end{aligned} \tag{I.A.1}$$

where *New Hire* is an indicator if the agent has tenure of two months or less, *Mandatory* and *Voluntary* are indicators for the Mandatory- and Voluntary-Conditions among program-eligible cohorts, respectively, and *Mentored* is an indicator for those assigned a mentor. $\zeta_{j,l,t}$ captures division-by-location-by-date fixed effects, absorbing fluctuations in call volumes across divisions on particular dates and location-specific shocks that may affect productivity; $\varepsilon_{i,t}$ is an idiosyncratic error term.

Our test of net discouragement/encouragement and information leakage is the joint test that $\beta_2 = \beta_3 = 0$, indicating that the productivity of new hires relative to veterans in the Mandatory- and Voluntary-Conditions is no different than the new hire-to-veteran productivity differences in hold-out cohorts. In the Voluntary-Condition, those who opt out of the program and those who opt in but are not mentored are a non-random group of agents who are pooled together in the $(\text{New Hire} \times \text{Voluntary})_i$ group indicator. We therefore conduct another test of the joint null of zero returns to mentorship and zero net discouragement and information leakage that is robust to non-random selection by testing whether $\beta_3 = \beta_5 = 0$. This test asks whether overall cohort-level productivity in the Voluntary-Condition differs from the productivity of hold-out cohorts.

The results in Panel A of Table I.A.5 report estimations of productivity relative to these hold-out cohorts. The negative and statistically significant coefficient on *New Hire* in Column (1) suggests that newly hired agents generate approximately 30% ($= e^{-0.354} - 1$) less daily revenue relative to veterans. The small and insignificant coefficients on *New Hire* \times *Mandatory* and *New Hire* \times *Voluntary* suggest that newly hired non-mentored agents in program-eligible cohorts perform like newly hired agents in hold-out cohorts. We thus fail to detect evidence of discouragement or leakage. In addition, we are unable to reject the null that $\beta_2 = \beta_3 = 0$, indicating that newly hired, non-mentored agents in the Mandatory-Condition performed like newly hired, non-mentored agents in the Voluntary-Condition. The significant coefficient on *New Hire* \times *Mandatory* \times *Mentored* and the insignificant coefficient on *New Hire* \times *Voluntary* \times *Mentored* align with the main treatment effect estimates discussed in Section 4. Column (2) shows that our results are robust when controlling for agent demographic characteristics—age, gender, and marital status—which is important, given that randomization of agents into treatments did not occur for veterans and hold-out cohort

agents (we do not have data on referral status or hiring scores for many veteran agents). Columns (3)–(4) repeat this exercise while using IHS(RPC) as the dependent variable. The small, insignificant coefficients on *New Hire* \times *Mandatory* and *New Hire* \times *Voluntary* in both of the columns further support the notion that discouragement and leakage are unlikely drivers of our estimated mentorship treatment effects. In the last row of Panel A, we cannot reject the null that productivity in the voluntary cohorts equals the productivity of hold-out cohorts.

I.A.2 Possibility of Program Crowd-Out

Formal mentoring programs have the potential to crowd out informal arrangements that fill similar functions. In our context, the program may have crowded out informal mentorship if the mentors were already providing informal guidance but stopped once the program began. To test whether our program crowded out mentoring that would have occurred in its absence, we assess whether non-mentored agents in program-eligible cohorts perform less well than new agents who joined the firm prior to the program. We again use the relative performance difference between new hires and veterans as the basis for comparison. If the mentorship program crowded out organic mentoring, then we would expect non-mentored new hires during the program to perform worse than new hires from prior years. We find no such evidence in Panel B of Table I.A.5. These estimations resemble those in Panel A, but the comparison group identifying the baseline *New Hire* indicator is now new hires who enter the firm prior to the experiment, rather than those from contemporaneous hold-out cohorts.² Non-mentored new hires during the time of the program had similar levels of productivity relative to new hires from before the program, suggesting that the program was unlikely to have crowded out organic mentoring.

I.A.3 Retention Effects and Inference About Sales Treatment Effects

In Section 4.2, we showed that, across both conditions, the mentorship program had no impact on retention, which means differential attrition is unlikely to explain the sales revenue treatment effects. We provide additional evidence of this negligible effect in Table I.A.7, which considers retention at longer horizons (i.e., 90–180 days). There are two reasons to probe further on retention. First, retention is consequential for the firm and small differences in retention may change the unit economics of the program. Second, we want to be confident that retention differences across conditions do not explain the treatment effects on sales.

An exercise that helps with both goals is to create what looks like a balanced panel, where we treat a mentorship slot as the unit of analysis and then track what happens to treatment and control agents over a fixed time horizon. This allows us i) to account for the long-term impact of each investment in mentorship and ii) to test whether the results differ from the main analysis, which would indicate that retention differences might drive some of the treatment gains because of imbalance in the panel that happens over time. Specifically, we fill in the productivity of agents who leave the firm with the expected productivity of a replacement for both mentored and non-mentored agents. The total productivity gain to the firm from a mentorship slot is the relative productivity gain of the treated agent while the agent is retained, followed by the productivity of a randomly drawn replacement post-

²Contemporaneous cohorts are not a good comparison group because they would be subject to the same limited supply of informal mentors.

separation. A similar approach is used for the productivity of non-mentored controls.³ When using the mentorship slot as the unit of analysis and filling in replacement productivity, we find results that largely mirror our main results on the unbalanced panel (see Panel A of Table I.A.10). Using the mentorship slot as the unit of analysis, the per-worker benefits to the firm in the Mandatory-Condition remain positive and significant, whereas the analogous firm-benefits are negligible in the Voluntary-Condition. The fact that the estimates are similar suggests that retention differences do not affect our inference about sales productivity treatment effects.

A second approach is to estimate bounds that are robust to non-random attrition, as proposed by Lee (2009). The intuition behind these tests is that one trims the highest and lowest values of the distribution based on the attrition rate from the sample, which gives bounds on the treatment effect for a non-retention related outcome. To the extent that those who are most likely to exit have the lowest values of the productivity variables, this exercise speaks directly to the retention concern. Table I.A.11 reports the upper and lower bounds of the estimated treatment effects on productivity in months 1–2. The results suggest that our estimated treatment effects in the Mandatory-Condition are largely attributable to the intensive margin of agents becoming more productive, and not due to differential retention effects.⁴

I.A.4 Salience of Differential Treatment

Another possibility is that differences in encouragement may have resulted from randomization into treatment, where receiving a mentor may have caused agents to feel special or exceptional because of perceived inequality in access to benefits provided by the firm. These perceptions may be exacerbated if the randomized nature of treatment was less salient to agents in the Mandatory-Condition compared to the Voluntary-Condition.

To test for such encouragement effects, we compare the treatment effects among mentored agents in teams where relatively more or fewer agents were also mentored. Specifically, we compute the fraction of agent-days where an agent may have been aware of differential treatment by working alongside anyone in their hiring cohort who had a different treatment status. If mentored agents felt they were chosen for mentorship for non-random reasons, then they might have changed their effort or buy-in in response to this perception of the firm’s commitment to them. If such perceptions existed, we would expect to see larger treatment effects for agents with fewer mentored teammates, as exposure to more mentored teammates would likely moderate these feelings of being exceptional. We show in Table I.A.12 that the treatment effects of mentorship in the Mandatory-Condition do not vary based on whether or not mentored agents had more exposure to teammates from their same hiring cohort who were also mentored. While we cannot completely rule out the possibility that encouragement effects were at play in our setting, our tests do not detect any evidence of their influence.

³Specifically, use the productivity of a randomly chosen newly hired non-mentor-eligible agent in the same location-division-year-quarter as the departed agent. We then re-estimate our main intention-to-treat regression models. We estimate ordinary least squares regressions with cohort fixed effects in all columns. We conduct 200 iterations of this estimation procedure to bootstrap the standard errors.

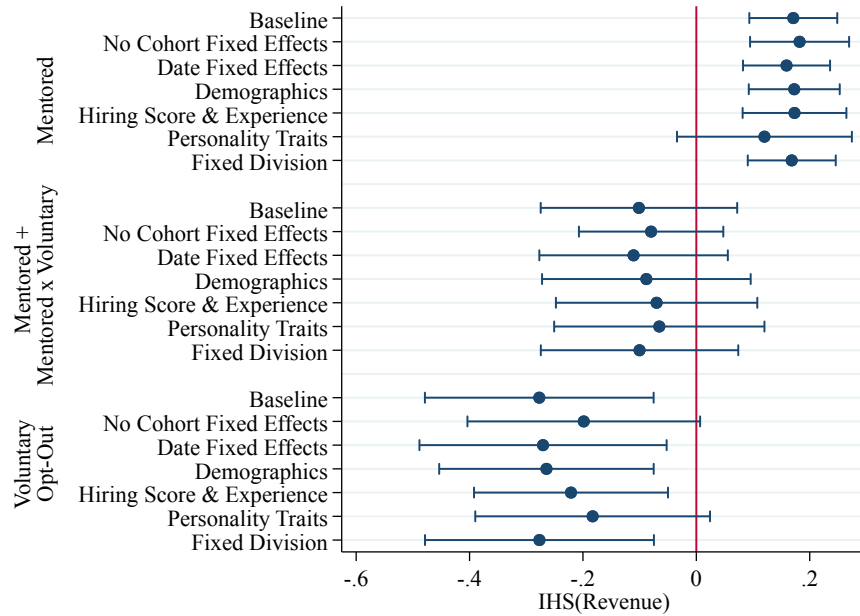
⁴These estimations do not include cohort fixed effects due to limitations in the implementation of the *leebounds* command, reducing their comparability to our main results.

I.A.5 Belief- and Preference-Based Explanations for Opt-Out Behavior

It is possible that beliefs, preferences, or both explain why some agents opt out of the voluntary program. We cannot determine whether miscalibrated beliefs explain the opt-out decision, as we made a design decision to make the program appear as natural as possible. In particular, we did not ask agents in the Voluntary-Condition to report their prior beliefs about program efficacy before soliciting their decision to opt in or opt out due to the concern that reflecting on program benefits may have altered sorting patterns relative to most sign-up procedures that do not ask for such reports. However, our representative survey (see Section 5.2 in the main text) suggests that many workers, approximately 26%, believe similar programs would not benefit them. Accordingly, beliefs around program efficacy may play a role in limiting participation in voluntary training and mentoring programs.

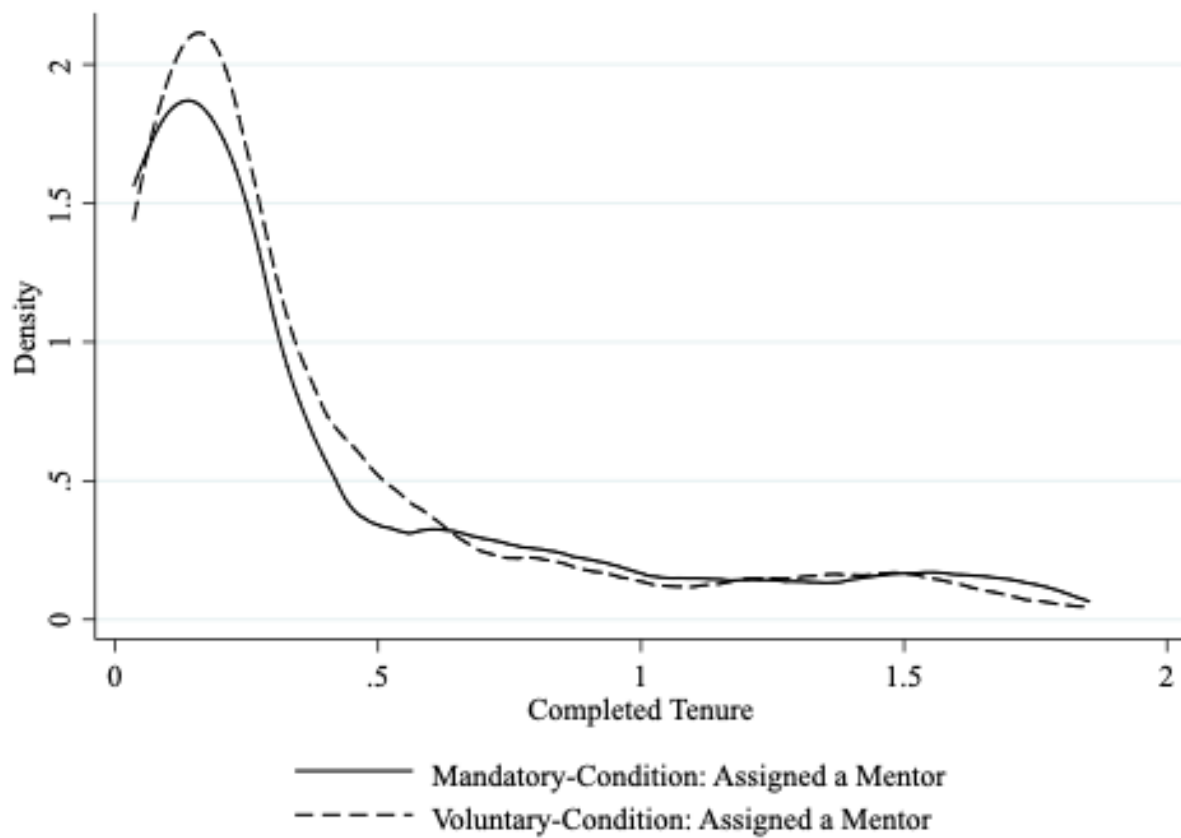
We find limited evidence that workers' preferences explain the decision to opt out, albeit our tests are only indirect. First, to the extent that personality characteristics proxy for preferences across worker types, we show in Table 4 that variation in personality characteristics does little to explain opt-out decisions. Second, we show in Table 6 that most agents in the Mandatory-Condition met with their mentors multiple times, suggesting that an aversion to meeting with more seasoned coworkers is also an unlikely explanation for their opt-out decisions.

Figure I.A.1: Robustness of the Treatment and Selection Effects of Mentoring on Productivity



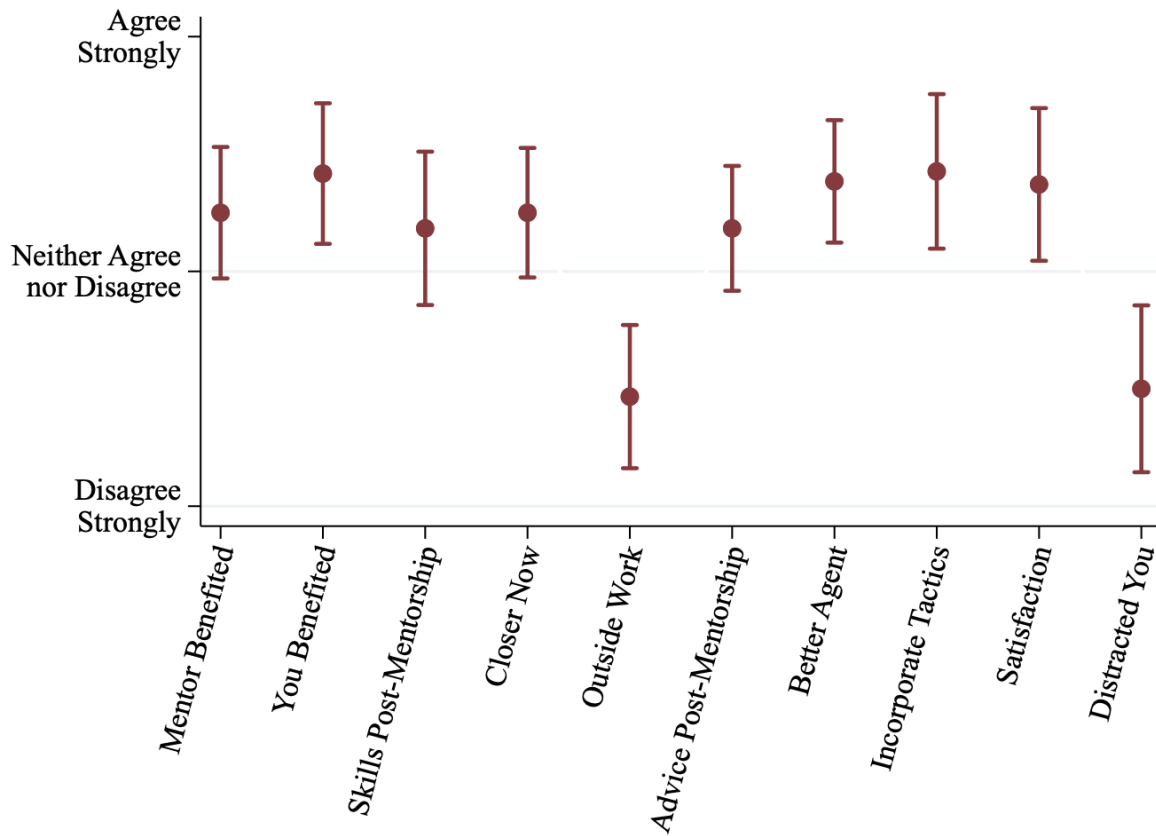
Notes. This figure plots the regression coefficients (and 95% confidence intervals) on *Mentored*, the sum of *Mentored* and *Mentored* \times *Voluntary*, and *Voluntary Opt-Out* from Equation (3). We use $\text{IHS}(\text{Revenue})$ as the dependent variable. The “Baseline” estimation replicates the result from Column (7) of Table 2. The second estimation excludes cohort fixed effects. The third includes date fixed effects. The fourth estimation includes controls for the agent’s demographic characteristics: age, gender, and marital status. The fifth estimation includes additional controls for the agent’s referral status, hiring score, previous call center experience, and previous sales experience. The sixth estimation layers on five more controls for the agent’s personality traits: extroversion, agreeableness, conscientiousness, emotional stability, and openness. Note that these controls are frequently missing for non-mentored agents in the Mandatory-Condition, suggesting these are not good controls for estimating the average treatment effect because the data come from a highly selected sample of agents in the Mandatory-Condition who took the intake survey. The seventh estimation removes observations in which agents are no longer working in the division in which they were initially hired.

Figure I.A.2: Tenure of Mentored Agents



Notes. This figure plots the distributions of completed tenure, in years, for each mentored agent in the Mandatory-Condition (solid line) and each mentored agent in the Voluntary-Condition (dashed line). For all agents, the completed tenure is calculated as the difference between their hire date and the date of the last day they are observed in the data, divided by 365.

Figure I.A.3: Responses to Post-Mentorship Survey



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

Table I.A.1: Balance in Mentor Demographics

	Mandatory-Condition	Voluntary-Condition	<i>p</i> -value
	(1)	(2)	(2)–(1)
Mentor Age (yrs.)			
Mean	22.92	23.28	0.732
Std Dev.	(2.57)	(3.20)	
Mentor Female			
Mean	0.07	0.17	0.391
Mentor Married			
Mean	0.14	0.22	0.587
Mentor Tenure			
Mean	1.40	1.12	0.398
Std Dev.	(0.52)	(1.14)	
Number of Protégés	127	155	

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

Table I.A.2: Treatment and Selection Effects of Mentoring (Pilot vs. Post-Pilot)

	Mandatory-Condition (All Agents)		Voluntary-Condition (Opt-In Agents)		Voluntary-Condition (Non-Mentored Agents)		Both Conditions (All Agents)		
	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.180*** (0.010)	0.110*** (0.008)	-0.154 (0.166)	-0.152 (0.113)			0.180*** (0.010)	0.110*** (0.008)	0.174*** (0.032)
Mentored \times Post	-0.033 (0.133)	0.005 (0.096)	0.099 (0.184)	0.096 (0.124)			-0.033 (0.128)	0.005 (0.092)	-0.105 (0.117)
Voluntary Opt-Out					-0.331*** (0.065)	-0.325*** (0.009)	-0.223** (0.092)	-0.177** (0.085)	-0.188*** (0.023)
Voluntary Opt-Out \times Post					-0.063 (0.186)	0.102 (0.103)	-0.225 (0.228)	-0.043 (0.153)	0.028 (0.069)
Mentored \times Voluntary							-0.374** (0.143)	-0.293** (0.106)	-0.325*** (0.026)
Mentored \times Voluntary \times Post							0.186 (0.209)	0.130 (0.150)	0.245** (0.119)
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.028	0.032	0.020	0.046	0.047	0.066	0.026	0.040	0.059
Observations	6,725	6,725	7,569	7,569	4,734	4,734	15,670	15,670	15,670

Notes. This table is structured similarly to Table 2, while allowing us to test whether estimates differ between the pilot and post-pilot data as described in Section 3.4. The sample is composed of agent-day productivity data for all mentor-eligible agents with post-training productivity data. The data covers agents' productivity across their first two months on the job after they complete training. IHS(.) indicates a variable that is transformed by the inverse hyperbolic sine. Revenue ("Rev") is daily total revenue and RPC is revenue per call. *Mentored* equals one for agents who were randomized to receive an available mentor, and zero otherwise, *Voluntary* equals one for agents in the Voluntary-Condition, and zero otherwise, and *Voluntary Opt-Out* equals one for agents who chose to opt out of possibly receiving a mentor, and zero otherwise. *Post* equals one for cohorts that entered the firm on or after May 27th (the post-pilot cohorts), and zero otherwise. The specifications in Columns (1) and (2) include all agents in the Mandatory-Condition. The specifications in Columns (3) and (4) include agents in the Voluntary-Condition who signaled their interest in receiving a mentor (i.e., those who opted in). The specifications in Columns (5) and (6) include agents in the Voluntary-Condition who were not assigned a mentor, including those who opted out of the program. The specifications in Columns (7)–(9) include agents from both conditions. The dependent variable in Column (9), *Index*, is the standardized weighted index of IHS(Revenue), IHS(RPC), IHS(RPH), and Adherence normalized using data from the non-mentored agents in the Mandatory-Condition. We estimate ordinary least squares regressions with cohort fixed effects in all columns. Standard errors are clustered by cohort for those workers entering the experiment after the pilot-period and by pilot-period-by-office for those workers entering during the pilot (this is because the pilot entailed assignment of the Mandatory- and Voluntary-Conditions at the office level), and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.3: Balance in Division Performance

	Mandatory-Condition	Voluntary-Condition	p -value
	(1)	(2)	(2)–(1)
Revenue			
Mean	803.35	846.61	0.267
Std Dev.	(108.52)	(151.9)	
RPC			
Mean	48.50	51.61	0.305
Std Dev.	(9.31)	(11.39)	
RPH			
Mean	128.45	132.9	0.432
Std Dev.	(16.46)	(21.87)	
Calls			
Mean	17.44	17.41	0.928
Std Dev.	(1.06)	(1.42)	
Hours			
Mean	6.25	6.34	0.638
Std Dev.	(0.68)	(0.66)	
Adherence			
Mean	0.87	0.88	0.272
Std Dev.	(0.02)	(0.03)	
Conversion			
Mean	0.24	0.23	0.285
Std Dev.	(0.03)	(0.02)	
Number of Cohorts	21	31	

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

Table I.A.4: Treatment and Selection Effects of Mentoring on Productivity (Log-Transformed Dependent Variables)

	Mandatory-Condition (All Agents)		Voluntary-Condition (Opt-In Agents)		Voluntary-Condition (Non-Mentored Agents)		Both Conditions (All Agents)		
	Log(Rev)	Log(RPC)	Log(Rev)	Log(RPC)	Log(Rev)	Log(RPC)	Log(Rev)	Log(RPC)	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.161*** (0.038)	0.101*** (0.025)	-0.080 (0.076)	-0.079 (0.049)			0.161*** (0.037)	0.101*** (0.024)	0.146*** (0.053)
Voluntary Opt-Out					-0.346*** (0.099)	-0.241*** (0.063)	-0.261*** (0.089)	-0.145*** (0.046)	-0.143*** (0.028)
Mentored \times Voluntary							-0.257*** (0.082)	-0.189*** (0.056)	-0.205*** (0.054)
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.029	0.036	0.021	0.054	0.049	0.072	0.027	0.045	0.059
Observations	6,725	6,725	7,569	7,569	4,734	4,734	15,670	15,670	15,670
<i>p</i> -value: Mentored + Mentored \times Voluntary							0.246	0.119	0.151

Notes. The sample used is composed of agent-day productivity data for all mentor-eligible agents with post-training productivity data. The data covers agents' productivity on their first two months on the job after they complete training. Log(.) indicates a variable that is transformed by the logarithm of one plus the value. Revenue ("Rev") is daily total revenue and RPC is revenue per call. *Mentored* equals one for agents who were randomized to received an available mentor, and zero otherwise, *Voluntary* equals one for agents in the Voluntary-Condition, and zero otherwise, and *Voluntary Opt-Out* equals one for agents who chose to opt out of possibly receiving a mentor, and zero otherwise. The specifications in Columns (1) and (2) include all agents in the Mandatory-Condition. The specifications in Columns (3) and (4) include agents in the Voluntary-Condition who signaled their interest in receiving a mentor (i.e., those who opted in). The specifications in Columns (5) and (6) include agents in the Voluntary-Condition who were not assigned a mentor, including those who opted out of the program. The specifications in Columns (7)–(9) include agents from both conditions. The dependent variable in Column (9), *Index*, is the standardized weighted index of Log(Revenue), Log(RPC), Log(RPH), and Adherence normalized using data from the non-mentored agents in the Mandatory-Condition. We estimate ordinary least squares regressions with cohort fixed effects in all columns. Standard errors are clustered by cohort for those workers entering the experiment after the pilot-period and by pilot-period-by-office for those workers entering during the pilot (this is because the pilot entailed assignment of the Mandatory- and Voluntary-Conditions at the office level), and are reported in parentheses. The bottom row reports the *p*-values from post-estimation tests that the sum of the coefficients on *Mentored* and *Mentored* \times *Voluntary* equals zero. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.5: Tests for SUTVA Violations Comparing New Hire Productivity Relative to Veteran Employees

Panel A: Discouragement/Leakage Tests Between Mentor-Eligible and Hold-Out Cohorts				
	IHS(Revenue)		IHS(RPC)	
	(1)	(2)	(3)	(4)
New Hire	-0.354***	-0.275**	-0.431***	-0.393***
	(0.113)	(0.119)	(0.070)	(0.072)
New Hire × Mandatory	-0.025	-0.041	-0.033	-0.044
	(0.117)	(0.120)	(0.069)	(0.070)
New Hire × Voluntary	0.007	0.002	0.047	0.039
	(0.109)	(0.111)	(0.068)	(0.068)
New Hire × Mandatory × Mentored	0.191***	0.199***	0.111***	0.115***
	(0.047)	(0.047)	(0.034)	(0.034)
New Hire × Voluntary × Mentored	-0.067	-0.041	-0.067	-0.050
	(0.067)	(0.067)	(0.042)	(0.042)
Division-Location-Date FE	✓	✓	✓	✓
Demographic Controls		✓		✓
Adj. R-Square	0.073	0.076	0.096	0.099
Observations	47,803	47,766	47,803	47,766
New Hire _{Man} = 0, New Hire _{Vol} = 0	0.954	0.907	0.445	0.425
New Hire _{Vol} = 0, Mentored _{Vol} = 0	0.564	0.800	0.281	0.500

Panel B: Crowdout Tests Between Mentor-Eligible and Pre-Experimental Cohorts				
	IHS(Revenue)		IHS(RPC)	
	(1)	(2)	(3)	(4)
New Hire	-0.400***	-0.370***	-0.399***	-0.394***
	(0.088)	(0.091)	(0.057)	(0.059)
New Hire × Mandatory	0.009	0.008	-0.074	-0.070
	(0.113)	(0.115)	(0.069)	(0.070)
New Hire × Voluntary	0.027	0.026	-0.015	-0.012
	(0.116)	(0.116)	(0.075)	(0.074)
New Hire × Mandatory × Mentored	0.166***	0.170***	0.093**	0.094**
	(0.062)	(0.065)	(0.039)	(0.040)
New Hire × Voluntary × Mentored	-0.070	-0.057	-0.069	-0.060
	(0.067)	(0.068)	(0.044)	(0.045)
Division-Location-Date FE	✓	✓	✓	✓
Demographic Controls		✓		✓
Adj. R-Square	0.102	0.103	0.131	0.132
Observations	75,094	75,094	75,094	75,094
New Hire _{Man} = 0, New Hire _{Vol} = 0	0.971	0.971	0.475	0.513
New Hire _{Vol} = 0, Mentored _{Vol} = 0	0.543	0.695	0.162	0.275

Notes. This table reports tests of the net effect of discouragement, leakage, and crowd-out by comparing the performance of three groups of agents to seasoned veterans who began working at the firm prior to the onset of the experiment: (1) new hires who were in mentor-eligible hiring cohorts; (2) new hires who were in hold-out cohorts that were not eligible for mentorship during the time of the experiment (Panel A); and (3) new hires who entered the firm before the experiment (Panel B). We test whether the performance of treatment-eligible new hires differs from those of non-treatment eligible new hires. Agents who opt out in the Voluntary-Condition are included (see the text for parameter interpretation given this sample). The dependent variable is IHS(Revenue) in Columns (1)–(2) and IHS(RPC) in Column (3)–(4). In the bottom two rows, *NewHire_{Man}* stands for new hire in the Mandatory-Condition, *NewHire_{Vol}* stands for new hire in the Voluntary-Condition, and *Mentored_{Vol}* stands for new hire who was mentored in the Voluntary-Condition. All specifications include division-by-location-by-date fixed effects. Columns (2) and (4) control for agent age, gender, and marital status. Standard errors are clustered by cohort for those workers entering the experiment after the pilot-period and by pilot-period-by-office for those workers entering during the pilot (this is because the pilot entailed assignment of the Mandatory- and Voluntary-Conditions at the office level), and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.6: Long-Term Treatment and Selection Effects of Mentoring on Productivity

	Mandatory-Condition (All Agents)		Voluntary-Condition (Opt-In Agents)		Voluntary-Condition (Non-Mentored Agents)		Both Conditions (All Agents)		
	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.079	0.051	-0.016	-0.053			0.079	0.051	0.055
<i>standard errors</i>	(0.083)	(0.066)	(0.111)	(0.073)			(0.080)	(0.063)	(0.048)
<i>sharpened q-value</i>							[0.329]	[0.372]	[0.315]
Voluntary Opt-Out					0.127	0.075	0.037	0.051	-0.057
<i>standard errors</i>					(0.161)	(0.088)	(0.102)	(0.066)	(0.093)
<i>sharpened q-value</i>							[0.561]	[0.372]	[0.408]
Mentored \times Voluntary							-0.099	-0.108	-0.069
<i>standard errors</i>							(0.149)	(0.104)	(0.074)
<i>sharpened q-value</i>							[0.404]	[0.329]	[0.329]
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.040	0.057	0.046	0.035	0.042	0.053	0.043	0.051	0.049
Observations	5,815	5,815	6,608	6,608	3,874	3,874	13,238	13,238	13,238
<i>p-value: Mentored + Mentored \times Voluntary</i>							0.852	0.432	0.803

Notes. The sample used is composed of agent-day productivity data for all mentor-eligible agents with post-training productivity data. The data covers agents' productivity across their third to sixth months on the job after they complete training. IHS(.) indicates a variable that is transformed by the inverse hyperbolic sine. Revenue ("Rev") is daily total revenue and RPC is revenue per call. *Mentored* equals one for agents who were randomized to receive an available mentor, and zero otherwise, *Voluntary* equals one for agents in the Voluntary-Condition, and zero otherwise, and *Voluntary Opt-Out* equals one for agents who chose to opt out of possibly receiving a mentor, and zero otherwise. The specifications in Columns (1) and (2) include all agents in the Mandatory-Condition. The specifications in Columns (3) and (4) include agents in the Voluntary-Condition who signaled their interest in receiving a mentor (i.e., those who opted in). The specifications in Columns (5) and (6) include agents in the Voluntary-Condition who were not assigned a mentor, including those who opted out of the program. The specifications in Columns (7)–(9) include agents from both conditions. The dependent variable in Column (9), *Index*, is the standardized weighted index of IHS(Revenue), IHS(RPC), IHS(RPH), and Adherence normalized using data from the non-mentored agents in the Mandatory-Condition. We estimate ordinary least squares regressions with cohort fixed effects in all columns. Standard errors are clustered by cohort for those workers entering the experiment after the pilot-period and by pilot-period-by-office for those workers entering during the pilot (this is because the pilot entailed assignment of the Mandatory- and Voluntary-Conditions at the office level), and are reported in parentheses. Sharpened q-values that adjust for the false discovery rate are presented in brackets, following Anderson (2008). The bottom row reports the *p*-values from post-estimation tests that the sum of the coefficients on *Mentored* and *Mentored \times Voluntary* equals zero. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.7: Longer-Term Treatment Effects on Retention

	Both Conditions (All Agents)			
	Tenure ₉₀	Tenure ₁₂₀	Tenure ₁₅₀	Tenure ₁₈₀
	(1)	(2)	(3)	(4)
Mentored	-0.061 (0.076)	-0.028 (0.095)	0.016 (0.069)	0.005 (0.088)
Voluntary Opt-Out	-0.242*** (0.075)	-0.213** (0.080)	-0.157** (0.065)	-0.094 (0.062)
Mentored \times Voluntary	0.003 (0.093)	-0.010 (0.116)	-0.047 (0.087)	-0.026 (0.091)
Cohort Fixed Effects	✓	✓	✓	✓
Adj. R-Square	-0.003	0.005	0.003	0.013
Observations	591	591	591	591
Mean Value of Tenure _{<i>i</i>}	0.45	0.35	0.30	0.26
<i>p</i> -value: Mentored + Mentored \times Voluntary	0.288	0.571	0.551	0.658

Notes. The sample used is composed of a single observation per agent, among all mentor-eligible agents with post-training productivity data. $Tenure_i$ equals one for agents who remain with the firm for at least i days after their hire date, and zero otherwise. *Mentored* equals one for agents who were randomized to receive an available mentor, and zero otherwise, *Voluntary* equals one for agents in the Voluntary-Condition, and zero otherwise, and *Voluntary Opt-Out* equals one for agents who chose to opt out of possibly receiving a mentor, and zero otherwise. The specifications in all columns include agents from both conditions. We estimate ordinary least squares regressions with cohort fixed effects in all columns. Standard errors are clustered by cohort for those workers entering the experiment after the pilot-period and by pilot-period-by-office for those workers entering during the pilot (this is because the pilot entailed assignment of the Mandatory- and Voluntary-Conditions at the office level), and are reported in parentheses. The penultimate row reports the average value of $Tenure_i$ for the sample of agents used in the specification within that column. The bottom row reports the *p*-values from post-estimation tests that the sum of the coefficients on *Mentored* and *Mentored* \times *Voluntary* equals zero. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.8: Determinants of Program Opt-Out and the Relationship Between Opting Out, Productivity, and Worker Characteristics (using Raw Hiring Scores)

Dep. Variable	= 1 if Opted Out				IHS(Revenue)		IHS(RPC)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	0.002 (0.002)	0.001 (0.002)	0.001 (0.001)	-0.001 (0.003)	-0.003 (0.005)	-0.005 (0.005)	-0.002 (0.003)
Female	-0.061** (0.029)	-0.008 (0.029)	-0.004 (0.032)	-0.016 (0.031)	-0.113 (0.134)	-0.101 (0.123)	-0.092 (0.081)
Married	-0.014 (0.051)	-0.022 (0.059)	-0.020 (0.056)	0.001 (0.050)	0.077 (0.156)	0.125 (0.185)	0.109 (0.124)
Hiring Score	-0.625** (0.268)	-0.534** (0.218)	-0.529** (0.211)	-0.537** (0.219)	2.571** (0.969)	2.766*** (0.806)	2.085*** (0.466)
Location 1	-0.059 (0.081)	0.040 (0.053)	0.033 (0.052)	0.021 (0.052)			
Referral	-0.023 (0.037)	0.019 (0.048)	0.015 (0.050)	0.010 (0.043)	-0.199 (0.121)	-0.189 (0.112)	-0.084 (0.062)
Call Center Exp.		0.093 (0.062)	0.092* (0.054)	0.098* (0.054)		0.400** (0.187)	0.318** (0.113)
Sales Experience		0.000 (0.058)	0.000 (0.055)	0.001 (0.056)		0.005 (0.217)	-0.038 (0.119)
High Extroversion			-0.030 (0.041)	-0.031 (0.042)		0.206 (0.126)	0.161 (0.095)
High Agreeableness			-0.040 (0.026)	-0.041 (0.026)		-0.196* (0.103)	-0.103 (0.085)
High Conscientiousness			-0.056 (0.049)	-0.054 (0.048)		-0.045 (0.112)	-0.057 (0.077)
High Emotional Stability			0.042 (0.050)	0.041 (0.051)		-0.151 (0.101)	-0.063 (0.056)
High Openness			0.023 (0.039)	0.026 (0.038)		0.066 (0.132)	0.082 (0.086)
Missing Survey				0.285*** (0.073)		-0.439** (0.180)	-0.288** (0.121)
Division Fixed Effects		✓	✓	✓			
Cohort Fixed Effects					✓	✓	✓
(Pse.) R-Square	0.035	0.063	0.078	0.204	0.061	0.074	0.096
Observations	365	341	341	365	4,734	4,734	4,734

Notes. The sample in Columns (1)–(4) is restricted to the 365 agents in the Voluntary-Condition, including those who quit before they completed training. The dependent variable is an indicator that equals one if the agent opted out. The coefficients capture the marginal effects from logistic regressions of different potential predictors of the choice to opt out. Experience and personality factors were collected via survey. We split personality scores on the sample median. Column (4) includes agents who did not complete the new hire survey, which we account for with a *Missing Survey* indicator. In Columns (5)–(6), we use the sample of agents in the Voluntary-Condition who were not mentored, and we regress IHS(Revenue), the inverse hyperbolic sine of daily revenue, on agents’ characteristics. In Column (7), we use the sample of agents in the Voluntary-Condition who were not mentored, and we regress IHS(RPC), the inverse hyperbolic sine of daily revenue-per-call, on agents’ characteristics. In Columns (2)–(4), the Division Fixed Effects indicators reflect the inclusion of controls for whether the agent works in the first largest division, the second largest division, or one of the other smaller divisions. Standard errors are clustered by cohort for those workers entering the experiment after the pilot-period and by pilot-period-by-office for those workers entering during the pilot (this is because the pilot entailed assignment of the Mandatory- and Voluntary-Conditions at the office level). We report marginal effects and delta-method standard errors in parentheses in Columns (1)–(4). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.9: Differences in Worksheet Content Across Conditions in the Experiment

	Total Words per Worksheet	Skill Words per Worksheet	Support Words per Worksheet	Other Words per Worksheet
	(1)	(2)	(3)	(4)
Mandatory-Condition	-0.880 (2.506)	0.357 (0.286)	0.126* (0.076)	-1.362 (2.412)
Adj. R-Square	-0.005	0.003	0.010	-0.004
Observations	172	172	172	172
Mean DV	46.94	4.22	0.43	42.29

Notes. This table considers differences in worksheet content between protégés in the Mandatory-Condition and those in the Voluntary-Condition. We have completed worksheet data for 172 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 from that agent. For each worksheet, we identify the number of words in the responses that relate to job-specific skills or knowledge (*Skill*), those that relate to receiving support, encouragement, and friendship (*Support*), and those that are neither related to skill nor support (*Other*), which include stop words. These become the dependent variables in our regression specifications of worksheet content on mentorship type. Robust standard errors are reported in parentheses. The mean of the dependent variable is listed below the observation count line. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.10: Productivity Treatment Effects When a Mentorship Slot is the Unit of Analysis

Panel A: Months 1-2									
	Mandatory-Condition (All Agents)		Voluntary-Condition (Opt-In Agents)		Voluntary-Condition (Non-Mentored Agents)		Both Conditions (All Agents)		
	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.141*** (0.053)	0.108*** (0.033)	-0.069 (0.053)	-0.050 (0.036)			0.141*** (0.053)	0.108*** (0.033)	0.114*** (0.018)
Voluntary Opt-Out					-0.276*** (0.064)	-0.162*** (0.047)	-0.199*** (0.061)	-0.097** (0.043)	-0.077*** (0.027)
Mentored \times Voluntary							-0.232*** (0.084)	-0.171*** (0.053)	-0.149*** (0.029)
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.023	0.034	0.025	0.043	0.036	0.046	0.024	0.036	0.078
Observations	10,216	10,216	9,986	9,986	6,588	6,588	22,451	22,451	22,451
<i>p</i> -value: Mentored + Mentored \times Voluntary							0.346	0.345	0.433

Panel B: Months 1-6									
	Mandatory-Condition (All Agents)		Voluntary-Condition (Opt-In Agents)		Voluntary-Condition (Non-Mentored Agents)		Both Conditions (All Agents)		
	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	IHS(Rev)	IHS(RPC)	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mentored	0.056* (0.031)	0.043** (0.021)	-0.028 (0.035)	-0.023 (0.024)			0.056* (0.031)	0.043** (0.021)	0.054*** (0.019)
Voluntary Opt-Out					-0.116*** (0.038)	-0.070** (0.027)	-0.094** (0.036)	-0.035 (0.026)	-0.036* (0.021)
Mentored \times Voluntary							-0.093* (0.050)	-0.074** (0.034)	-0.071*** (0.026)
Cohort Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R-Square	0.013	0.021	0.019	0.026	0.029	0.036	0.018	0.026	0.078
Observations	27,920	27,920	28,569	28,569	19,070	19,070	63,242	63,242	63,242
<i>p</i> -value: Mentored + Mentored \times Voluntary							0.432	0.351	0.426

Notes. The results in this table show estimates of the treatment effects of mentorship when a slot (i.e., an occupied position in either the treatment or control group during the experiment) is the unit of analysis. This enables us to track long-term performance and account for attrition. To do so, we form a panel that is “balanced” by taking the actual productivity of agents when they remain at the firm, while replacing the productivity measure with the expected productivity of imputed replacement agents if the agents have left. In other words, for mentor-eligible agents who leave the firm before the two-month mark (six-month mark) in Panel A (Panel B), we extend the time series of their productivity provision to two (six) months and replace their post-termination productivity values with the productivity of a randomly newly hired replacement agent. Specifically, use the productivity of a randomly chosen newly hired non-mentor-eligible agent in the same location-division-year-quarter as the departed agent. We then re-estimate our main intention-to-treat regression models. We estimate ordinary least squares regressions with cohort fixed effects in all columns. We conduct 200 iterations of this estimation procedure to bootstrap the standard errors. The bottom row reports the *p*-values from post-estimation tests that the sum of the coefficients on *Mentored* and *Mentored \times Voluntary* equals zero. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.11: Lee Bounds Estimates of the Effect of Mentoring on Productivity

	Mandatory-Condition		Voluntary-Condition	
	IHS(Revenue)	IHS(RPC)	IHS(Revenue)	IHS(RPC)
	(1)	(2)	(3)	(4)
Mentored _{lower}	0.169*** (0.046)	0.094*** (0.035)	-0.222*** (0.076)	-0.131*** (0.048)
Mentored _{upper}	0.208** (0.083)	0.122** (0.051)	-0.058 (0.042)	-0.018 (0.031)
Observations	8,408	8,408	9,231	9,231

Notes. This table uses agent-day productivity data for agents in the Mandatory-Condition in Columns (1) and (2) and for agents in the Voluntary-Condition, excluding those agents who opt out of the program, in Columns (3) and (4). The sample is a balanced panel of agents' productivity through their first two months of tenure, augmented with missing data for those who exit the firm during the panel. We estimate treatment effect bounds that account for non-random attrition as proposed by Lee (2009). These estimations do not include cohort fixed effects, as the *leebounds* command cannot accommodate fixed effects, which causes the estimates to differ from those in our main results. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table I.A.12: Tests for Variation in Mentoring Treatment Effects Based on Exposure to Teammates with the Same or Different Treatment Status

	IHS(Revenue)	IHS(RPC)
	(1)	(2)
Mentored	0.154*** (0.035)	0.101*** (0.019)
Mentored \times Same Treatment	0.030 (0.176)	0.022 (0.076)
Cohort FE	✓	✓
Adj. R-Square	0.033	0.036
Observations	6,725	6,725

Notes. The sample consists of agents in the Mandatory-Condition. After training, agents joined different teams, after which they would begin the mentoring program. For each agent, mentored or not, we generate an intermediate variable (not the regressor), *Same Teammate*, which equals one if, on that day, all the agent's teammates from their hiring cohort had the same mentored/non-mentored treatment designation as their own. Otherwise, *Same Teammate* equals 0 for that agent-day. We then compute the rolling average of the *Same Teammate* variable for each agent, from their first day on the job after training to the present day, and we label this *Same Treatment*. The variable *Same Treatment* captures the fraction of days on the job historically that a new hire worked alongside teammates from their same hiring cohort who had the exact same mentored/non-mentored treatment designation as themselves. So, a mentored (non-mentored) agent with *Same Treatment* = 1 has never worked alongside a non-mentored (mentored) teammate from their same hiring cohort and would have no perceptions of differential treatment. The mean of *Same Treatment* in the sample is 0.22. All specifications include cohort fixed effects. Standard errors are clustered by cohort for those workers entering the experiment after the pilot-period and by pilot-period-by-office for those workers entering during the pilot (this is because the pilot entailed assignment of the Mandatory- and Voluntary-Conditions at the office level), and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

I.B 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.⁵

I.B.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.

I.B.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.

I.B.3 Intervention Start Date

2019-05-27

I.B.4 Intervention End Date

2019-12-20

I.B.5 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.).⁶

I.B.6 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

⁵We received IRB approval at all of our respective institutions prior to the onset of the mentoring program. See IRB19-0769.

⁶Following our prior paper, we use an approximate logarithmic transformation of RPC (and revenue). In this case, we use IHS() rather than logs because we have daily data that includes cases where these newly hired agents make zero sales. Our results, as shown in Table I.A.4, are similar if we use the original pre-registered log() measure.

both agent’s firm-specific knowledge and their individual effort.⁷ Worker turnover measures whether the interventions changed the agents’ propensity to leave the firm. Log of completed tenure is a different measure of retention that has been used in the prior literature and the attendance measure provides an adjacent measure of agent effort.⁸ Finally, engagement measures are hypothesized to be forward looking measures of productivity.⁹

I.B.7 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:

I.B.8 Cohort Level Randomization

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

I.B.9 Within Cohort Randomization

For cohorts in Mandatory-Condition, new hires will receive a mentor with probability 50%. This will be communicated privately between sales floor staff and the individual workers. Agents in the Mandatory-Condition who do not receive a mentor will not receive communication regarding the program. For cohorts in Voluntary-Condition, sales floor staff verbally explain the firm’s mentorship program, answer questions, and provide each agent a confidential ballot where they can decide whether or not to enter a lottery which randomly

⁷RPC was the primary endpoint based on our experience analyzing the productivity of veteran agents within the firm (Sandvik et al., 2020), but seminar participants asked us about other margins of adjustment (like time use) that is captured by total revenue and is more relevant for new agents compared to more experienced workers, which is why we report both metrics.

⁸In practice, we find no differences in completed tenure and simplify the analysis by using indicators for completed tenure rather than log completed tenure.

⁹As noted below, the firm changed the cadence of collecting engagement metrics (which were planned for 5 weeks after training completion), so we do not have these measures for many cohorts.

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.

I.B.10 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.¹⁰ 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].¹¹ 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.”

I.B.11 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 exists, but for whom historical data is available) were “hold out” cohorts who knew nothing about mentoring.¹² A “sentiment survey” will be administered to all agents in their 5th week on the sales floor.¹³ This will be one week after mentored agents finish hiring. We will gather information on their feelings towards the onboarding 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.

¹⁰As 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.

¹¹The post-mentorship survey completion rates of mentors and protégés were very poor, so we do not have meaningful data for this question. Anecdotally, the average responses of both sets of individuals suggests that protégés’ RPC would have been lower in the absence of mentorship, but the inference is not precise.

¹²Further consultation with the firm’s staff meant that the hold-out procedure was not isolated to a single division or office but instead involved rotation of some cohorts out of the program. More detail about hold-out cohorts is provided in the main text.

¹³We 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.

I.B.12 Randomization Method

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

I.B.13 Randomization Unit

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

I.B.14 Was the treatment clustered?

Yes

I.B.15 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.

I.B.16 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.

I.B.17 Sample size (or number of clusters) by treatment arms

Please see design field.

I.B.18 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).

I.B.19 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.¹⁴ 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:

¹⁴Mentors received instructions that either put more emphasis on the supportive nature of the program or the skills-building nature of the program. We detect no differences between instruction type. Because of this and for brevity, we omit this indicator from the models in our treatment effects tests.

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

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 regression of Y on a dummy for receiving a mentor and cohort fixed effects in cohorts that have (entirely) randomly assigned mentoring. This yields:

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

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

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

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

We plan to validate these estimates using the electronic survey responses collected after the protégé graduates from the formal mentoring program, approximately 4 weeks following the initial onboarding instruction (e.g. how to use the systems, enroll for benefits, etc.).¹⁵ 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.¹⁶

Finally, to assess whether mentoring detracts from—or improves sales—for the mentor, we will regress mentor log RPC and other sales measures on indicators demarking whether the mentor is eligible to mentor but has not yet done so, whether they have previously mentored in the program, or whether they are actively mentoring a protégé. This regression will include mentor fixed effects and mentor tenure.¹⁷

¹⁵As mentioned earlier, we were not able to administer this survey.

¹⁶This is a very high-dimensional exercise, and the most interesting potential differences (like females mentoring females) ended up having very small cell sizes, reducing statistical power for these tests.

¹⁷Tests that compare the characteristics of mentors and protégés, and those that look at the impact of mentorship on mentor productivity, are likely to be discussed in a separate article.

I.C Documentation of Instructions to Mentors and Example Worksheet for Structuring Conversations

Mentor Instructions

What is a Mentor?

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

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

What will You Do as a Mentor?

As a mentor, you will do the following:

1. You will meet with your protégé at **least once a week**.
 - a. Before meeting, your protégé will complete the Protégé Worksheet.
 - i. If he/she has not completed it, you will kindly help him/her do so.
 - b. During your meeting, you and your protégé will discuss his/her responses. You should also take this time to do the following:
 - i. Impart knowledge and skill by explaining, giving useful examples, and demonstrating processes, and asking thought-provoking questions.
 - ii. Discuss actions you've taken to become a successful sales agent.
 - iii. Provide him/her with any tips and sales tactics that help you overcome customer concerns and that help you up-sell to better services.
 - iv. Practice the designated sales protocol with them and help them gain a strong understand of the products, services, and bundles available.
2. After meeting with your protégé, you will deliver the finished worksheet to _____.
 - a. _____ will initial and timestamp the worksheet and make a record that you completed your weekly meeting responsibility.
3. Every two weeks, you will be asked to complete an on-line survey.
 - a. These questions are meant to gauge the progress of your protégé and the overall benefit of the mentoring relationship.
 - b. Please answer these questions honestly, as they are not meant to punish but, instead, to help _____ assess the effectiveness of the mentorship program.

Protégé Worksheet (Week 1)

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

Weekly Self-Reflection:

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

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

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

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

Weekly Goal:

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

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

For Mentors to Respond:

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

Protégé's Initials

Mentor's Initials

Intern's Initials &
Timestamp

I.D Pre-Registered Estimation of Heterogeneous Treatment Effects

We can use the estimated Mandatory-Condition and Voluntary-Condition productivity treatment effects, along with the data on the fraction of Voluntary-Condition agents who opt out of receiving a mentor, to estimate the treatment effect of mentorship among opt-out agents. We pre-registered the following procedure for this purpose. Using productivity measure Y , we define the conditional average treatment effect of mentoring given selection into participation as the difference in expected production between mentored and non-mentored agents conditional on opting in:

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

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

$$\begin{aligned} ATE &= E(Y_{MandatoryMentored}) - E(Y_{Mandatory\sim Mentored}) = \\ &\beta_{OptInMentored} \times \pi_{OptIn} + \beta_{OptOutMentored} \times \pi_{OptOut}. \end{aligned}$$

Rearranging terms, we get,

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

We use the estimated treatment effect in the Mandatory-Condition as the estimated ATE , and we use the estimated treatment effect in the Voluntary-Condition as the estimated $ATE|OptIn$.¹⁸ The values of π come from the proportion of agents who opted out in the Voluntary-Condition. We show the estimated treatment effect for opt-out agents in Table I.D.1, where standard errors come from 500 block-bootstrap iterations by cohort. The point estimate of 1.207 in Column (1) of Panel A implies that opt-out agents would have more than doubled their overall revenue generation, on average, had they received mentorship. Based on this analysis, opt-out agents were those who would have benefited the most from receiving mentorship. That is, program participation is negatively correlated with treatment gains.

¹⁸We include cohort fixed effects in estimating these treatment effects.

Table I.D.1: Estimated Treatment Effect of Mentoring Among Opt-Out Agents

Pre-Registered Estimates of Opt-Out Treatment Effects		
	IHS(Revenue)	IHS(RPC)
	(1)	(2)
Opt-Out Mentored Effect	1.422** (0.648)	1.159*** (0.417)

Notes. This table reports estimates of the treatment effect of mentorship among agents who would have opted out of the program using the pre-registered estimator described above. This estimator imposes that the full difference between the Mandatory- and Voluntary-Condition comes from heterogeneous treatment effects and selection. To estimate standard errors, we block-bootstrap by cohort ($N = 52$) over the whole procedure, with 500 bootstrap replications for each column. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

I.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 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]
· Check in with my coach and be confident

Panel D: Words Associated with Sales Skills and Knowledge

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

Panel E: Words Associated with Receiving Support

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

I.F Calculation Details for the Value to the Firm

This section details how we calculate the NPV to the firm from the mentorship program. In all calculations, we compare the actual stream of revenues generated by treated employees net of commission pay to the actual stream of revenues generated by control employees, again net of commission pay. To account for potential retention differences and to get at the long-term effect, we assume that each treatment is a discrete investment that happens only once. We then use the treatment or control worker’s slot as the unit of analysis, where we track each worker over a six-month horizon. If a worker leaves, we continue to track that worker’s shifts going forward, but we fill in their productivity with the expected net revenue of a replacement employee.

To estimate additional revenues, using a treatment event as the unit of analysis, we take the estimated treatment effect in the Mandatory-Condition across the first six months of tenure, $5.75\% = e^{0.056} - 1$ (from Column (1) in Table I.A.10, Panel B), and multiply it by the average daily revenue of agents in non-mentored slots, \$735, yielding a daily revenue increase of about \$42 over the first six months of tenure. There were 127 mentored agents in the Mandatory-Condition, resulting in \$882 more in revenue per agent-month based on agents’ total days of work. The firm earns this additional revenue net of an 8% commission rate paid to sales agents. Hourly pay is invariant to productivity and did not change with the program. We multiply these monthly net-revenue amounts by six, the number of months, and by 127, the number of mentorship slots. We conservatively assume this additional revenue is realized at the end of the year and discount the future cash flow using a 12.5% discount rate, which gives us a present value of the additional revenue earned by mentored agents equal to approximately \$550,000.

Second, we consider the costs. These costs include the estimated time costs of taking the mentors off the phone (protégés’ opportunity costs are included in the revenue treatment effects) and administrative costs. Mentors and protégés spent 30 minutes in the mentorship meetings each week. Revenue per hour for mentors averaged \$148 and they were paid an additional \$10 of “kudos” points for completing each meeting. Together this implies a cost of \$84 per meeting or \$30,000 total, assuming that each treated agent met with their mentor 2.5 times. We include the administrative costs of the two internal mentorship staff members who oversaw the program in the two locations, estimated to total approximately \$67,000, (generously) assuming that mentoring administration accounted for 50% of their workload. This leads to total costs of about \$97,000. Together, this leads to a net present value of the program equal to approximately \$453,000, which may be a lower bound if more productive agents allow sales managers to have larger spans of control (Espinosa and Stanton, 2021).

The estimate of potential gains from reallocating agents in the Voluntary-Condition to the Mandatory-Condition is done as follows: There are 155 agents to reallocate. Assuming the full treatment effects translate, we can use the previous estimates, which allocate overhead and opportunity costs, and scale by the relative number of agents to get roughly $\frac{155}{127} \times \$453,000 = \$553,000$. We then account for the fact that overhead was already allocated to the Mandatory-Condition, yielding a gain of \$620,000. This number can be scaled by the treatment effect difference due to self-selection and treatment effect heterogeneity.

Finally, the switch from a voluntary to a mandatory program would have led to an increase in gross commissions to agents of around \$65,620 ($= \$42 \times 21 \times 6 \times 0.08 \times 155$).