informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 67, No. 12, December 2021, pp. 7687–7707
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Employee Responses to Compensation Changes: Evidence from a Sales Firm

Jason Sandvik,[a] Richard Saouma,[b] Nathan Seegert,[c] Christopher Stanton[d,e]

[a] A.B. Freeman School of Business, Tulane University, New Orleans, Louisiana 70118; [b] Eli Broad College of Business, Michigan State University, East Lansing, Michigan 48824; [c] David Eccles School of Business, University of Utah, Salt Lake City, Utah 84112; [d] Harvard Business School, Harvard University, Boston, Massachusetts 02163; [e] National Bureau of Economic Research (NBER), Cambridge, Massachusetts 02138

Contact: jsandvik@tulane.edu, https://orcid.org/0000-0001-5407-1136 (JS); rs2@msu.edu, https://orcid.org/0000-0002-6355-2190 (RS); nathan.seegert@eccles.utah.edu, https://orcid.org/0000-0002-0862-1181 (NS); cstanton@hbs.edu, https://orcid.org/0000-0001-6951-7882 (CS)

Received: April 19, 2019
Revised: April 4, 2020; June 29, 2020; September 8, 2020
Accepted: September 22, 2020
Published Online in Articles in Advance: March 16, 2021

https://doi.org/10.1287/mnsc.2020.3895

Abstract. What are the long-term consequences of compensation changes? Using data from an inbound sales call center, we study employee responses to a compensation change that ultimately reduced take-home pay by 7% for the average affected worker. The change caused a significant increase in the turnover rate of the firm's most productive employees, but the response was relatively muted for less productive workers. On-the-job performance changes were minimal among workers who remained at the firm. We quantify the cost of losing highly productive employees and find that their heightened sensitivity to changes in compensation limits managers' ability to adjust incentives. Our results speak to a driver of compensation rigidity and the difficulty managers face when setting compensation.

History: Accepted by Lamar Pierce, organizations.
Supplemental Material: The data files and online appendices are available at https://doi.org/10.1287/mnsc.2020.3895.

Keywords: compensation • motivation-incentives • personnel • productivity • worker retention

1. Introduction

How will full-time employees respond to unanticipated, adverse compensation changes? Will highly productive workers respond differently than their less productive peers? Can employee responses impact firm performance? When tasked with adjusting compensation, managers balance incentive and retention effects with expense reductions, all while limiting the damage they cause to implicit relational contracts with their employees. Given this difficulty and the uncertain responses of workers, managers largely avoid imposing adverse compensation changes.

The literature has surfaced myriad negative worker responses to explain managers' reluctance to reduce compensation. Behavioral reasons focus on fairness concerns (Fehr et al. 2009, Cohn et al. 2014a) or social preferences such as *warm glow* and social norms (DellaVigna et al. 2016). Other work documents increases in theft and antisocial behavior after unanticipated pay cuts (Greenberg 1990, Giacalone and Greenberg 1997). The research most related to ours emphasizes the importance of compensation practices that retain top talent. For example, Zenger (1992) finds that disproportionately rewarding high performers contributes to their retention, suggesting that managers' ability to revise compensation depends critically on their most productive employees' responsiveness.

The need to retain high performers clearly influences compensation structure, but experiments with compensation in real employment relationships are rare, and an open question is how compensation-retention sensitivity varies over the distribution of worker performance. The underlying issue is that within-firm performance differences across workers can be significant (Lazear 2000, Mas and Moretti 2009, Lazear et al. 2015, Sandvik et al. 2020). If the most productive workers are also the most responsive to adverse compensation changes, then average turnover rates do not adequately capture the full impact on firm performance because of the loss of exceptional talent. This is consistent with Bewley (1998, p. 476), who conducted more than 300 interviews to understand why firms were reluctant to cut pay, even in the face of falling customer demand. Bewley states: "[turnover] among the better workers is especially feared, because they are more valuable and can find new jobs more easily."

We show that the cited concerns of managers are consistent with the responses of an organization's highly productive workers; they quit in response to a reduction in take-home pay. Our empirical setting is a U.S.-based inbound sales call center. The president of one of the six divisions (henceforth Division 1) independently decided to rebalance the division's commission schedule, which led to an 18% decrease in expected commission pay and a 7% reduction in total take-home pay. The realized pay reductions closely matched these expectations. Three months after these changes, the president of a second division enacted similar changes, which led to a decrease in average take-home pay of more than 14%.

To study heterogeneous responses across the employee performance distribution, we use worker-level output data, given to us by the firm, to estimate individual workers' sales productivity before the compensation changes. Individual productivity is widely dispersed, for example, workers at the 75th percentile of the distribution sell about 50% more on a given call than those at the 25th percentile. This large dispersion motivates our investigation of the turnover and effort responses across the worker productivity distribution.

We use three empirical approaches to estimate worker responses. First, we begin with a traditional difference-in-differences estimation, where we compare workers in Division 1, before and after the compensation changes, to workers in untreated control divisions. Importantly, about two months before the compensation changes, Division 1 and the control divisions satisfied the difference-in-differences common-trends assumption. Our second empirical strategy mitigates concerns about long-term trend differences across divisions by focusing on heterogeneous responses for agents of different productivity levels *within* the same division. Our third approach uses survey responses to complement our main results and surface potential mechanisms.

Our main findings point to the importance of the most productive workers' heightened responsiveness to pay changes, as the turnover rate of highly productive agents in Division 1 increased significantly after the compensation changes. Specifically, workers with pre-treatment productivity that was one standard deviation above the mean had between a 40% and 56% increase in attrition, relative to the baseline turnover rate. The average attrition rate of workers in Division 1 did not change, however, as less productive workers decreased their propensity to leave the firm. The loss of human capital from highly productive workers—who contribute significantly more to revenue than their colleagues—had significant consequences for overall profitability. Despite initial savings on compensation expenses, the loss of highly productive agents reduced the firm's operating performance and led to a negative estimated net present value of the changes. The second compensation change, which occurred in Division 2, validates these results. Division 2 contained only veteran, highly productive workers, and after their compensation was adjusted, the turnover rates of these extremely productive workers increased substantially relative to other divisions.

Turnover is rarely instantaneous. When examining Division 1, it took a little over four months for the cumulative loss in sales to outweigh the savings from the reduction in commission payments that resulted from the changes. Of equal importance, we observed virtually no abnormal attrition in the six weeks immediately after the compensation changes—highlighting the fact that workers did not respond to the announcement of the compensation changes by quitting *immediately*. This delay in the onset of turnover allows us to understand how job performance was impacted by the compensation changes, including for those workers who ultimately left the firm.

We find minimal evidence that agents responded to the compensation changes by adjusting their effort. If anything, Division 1 agents may have tried to increase their effort to offset some of the income lost because of the changes in their commission schedule. At first glance, this finding appears inconsistent with basic agency theory in static settings (Jensen and Meckling 1976, Hölmstrom 1979) and with more recent behavioral theories. However, in long-term employment relationships, workers' responses are impacted by income effects, where the desire to offset a portion of lost earnings may offset the desire to reduce effort in response to lower-powered incentives (Ashenfelter and Heckman 1974, Stafford 2015).

Our results underscore the importance of how compensation policy and performance heterogeneity interact in long-term employment relationships. We find that high performers have the greatest turnover sensitivity to compensation reductions, likely because of their superior outside options. High performers' responsiveness to compensation changes is consistent with managers' stated reasons for the rigidity of compensation contracts observed in aggregate data. Our results predict that managers will have more flexibility under labor market institutions that allow compensation policy to be tailored to individual workers.

2. Related Literature and Potential Mechanisms

The two most relevant strands of literature for understanding the relationship between compensation and employee effort and retention are behavioral theories and neoclassical economic foundations. Behavioral theories tend to be tested in short-term settings and focus on effort or output changes in response to compensation changes. Different theories emphasize (wage) fairness (Fehr et al. 2009), social comparisons (Larkin et al. 2012, Cohn et al. 2014b, Obloj and Zenger 2017), and negative reciprocity (Fehr and Falk 1999, Dickson and Fongoni 2019). There is some empirical support for these mechanisms in field experiments, showing that wage reductions reduce output (Kube et al. 2013) or cause attrition from short-term contractual work (Chen and Horton 2016).

By contrast, we find minimal changes in output or effort. The responses we do document, especially on the turnover margin, appear consistent with workers' optimizing their decisions according to their economic interests. We believe there are two main differences between our results and past studies. First, we focus on a long-term employment relationship where the overall change in incentives did not just alter the marginal return to effort but also affected workers' overall income levels. Despite this, past work on long-term employment relationships does show evidence of shirking or reduced production quality in response to perceived insufficient pay raises or unfair compensation (Krueger and Mas 2004, Mas 2006). A key difference in our setting is the presence of performance pay, where workers have an incentive to try and make up lost income with higher effort. Indeed, prior work has shown that piece rate contracts, or contracts with commissions, may have different incentive effects than adjustments to fixed wages (Esteves-Sorenson 2018).

Our work also contributes to a growing literature on how compensation influences employee retention. The increased turnover rate of highly productive agents in our setting aligns with findings in Krueger and Friebel (2018), who study a reduction in incentive pay at a personnel search firm. They also document changes in output, but the firm they studied increased fixed wages to offset part of the reduction in incentive pay, potentially alleviating income effects and giving rise to effort reductions. Several other studies have considered the ability to attract or retain workers through compensation policy, many of which focus on the effects of stock options (Oyer 2004, Oyer and Schaefer 2005, Aldatmaz et al. 2018). Larkin and Leider (2012) examine how different menus of incentive schemes lead to the selection of sales workers based on their confidence, suggesting convexity in pay helps select highly motivated employees. Campbell et al. (2012) show that high performing lawyers (as measured by their earnings) are less likely to turnover, and Carnahan et al. (2012) show firms that tilt compensation toward high performers show lower turnover at the top of the distribution. We complement these papers by studying employee behavior under different compensation regimes.

Our results on turnover have implications for understanding the link between changes in compensation and monopsony power. The fact that the average worker often does not respond to compensation adjustments is thought to indicate monopsony in the labor market (Manning 2003, Dube et al. 2018). Although the average turnover response to a significant compensation change is negligible in our setting, the increased turnover of highly productive employees caused the change to be net present value (NPV) negative. Accordingly, average turnover rates alone are insufficient to infer whether a firm benefits from exercising labor market power. The remainder of the paper details the setting and our empirical strategy and provides further discussion of results and limitations.

3. Firm Setting and the Compensation Changes

The compensation changes that we study occurred in a U.S.-domiciled, inbound sales call center. It employed more than 2,000 sales agents over the course of our sample period in two main offices and a third smaller office. The agents are organized into six divisions, based on the goods and services (henceforth, products) they sell. The presidents of two different divisions, Division 1 and Division 2, drastically changed the commission schedules of the agents in their divisions, which ultimately led to significant decreases in the average commission and take-home pay of their workers. We briefly provide context here and relegate further details to Online Appendix A.

3.1. Firm Setting

The firm contracts directly with national television, phone, and internet providers to market and sell their products. The different sales divisions are uniquely characterized by the products their agents sell. These divisions are overall similar, with the exception of Division 2, as these agents respond to inquiries from small businesses rather than residential customers. The firm reserves space in Division 2 for its most productive and experienced agents because of the higher profitability associated with small business customers. Division 1 and Division 2 employed 20% and 7% of the firm's sales force, respectively.

An agent's task is to respond to customer needs and to upsell high profit margin products, when appropriate. Sales opportunities are randomly assigned to agents within a division through a queue that assigns agents to calls, making it possible to estimate individual agent productivity after observing a large number of calls for each agent.

3.2. Agent Compensation

Agent pay is made up of a fixed hourly wage, commissions, and occasional small bonuses. New agents have a base hourly wage of approximately 150% of minimum wage, which increases by about $1 per hour annually. Commissions are a significant part of an agent's total compensation. During the eight weeks before the compensation changes occurred, the average Division 1 agent earned $318 per week in commissions, and the average control division agent earned $201. These amounts constituted approximately

30%–40% of agents' overall take-home pay. The president of each division has sole discretion to adjust their sales agents' commission schedules.

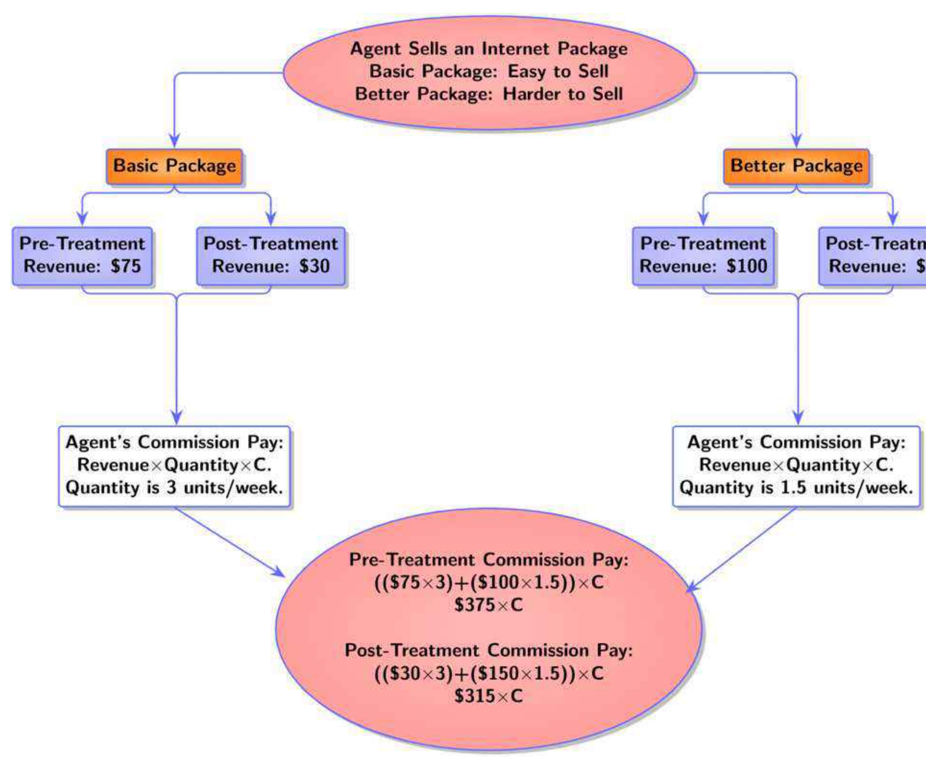
The mapping of products sold by an agent to the commission pay received by the agent—that is, the commission schedule—is determined as follows. Each product has a transfer price assigned to it by the division president, which the firm refers to as *revenues*. These revenue amounts approximate the actual top-line revenue generated for the firm through the sale of the product. For instance, a low-end cable television package may be assigned a revenue of $50, whereas a high-end package may be assigned a revenue of $200. These amounts form the basis for which agents receive commissions, and division presidents set them in a way that (1) rewards agents for each sale they make and (2) provides greater rewards for selling high profit margin products.[1] Throughout the week, agents generate revenue through each sale they make. At the end of the week, these revenue amounts are summed and multiplied by the an agent's commission rate, which is a function of the agent's audited call quality and selling efficiency relative to other agents.[2] The product of an agent's weekly revenue and commission rate determines the weekly commission payment.

3.3. Changes to the Commission Schedule

In November 2016, the president of Division 1 radically recalibrated their agents' commission schedule by changing the transfer prices for the products sold by agents. The commission rate function was not changed, meaning that agents still made the same percentage of revenue on any sale as determined by their relative selling efficiency and call quality, but the revenue amount itself was altered. Prices for customers remained unchanged, as did the per-product top-line accounting revenues realized by the firm. Figure 1 gives an example of changes in the commission schedule for two different types of internet packages.

The Division 1 president tilted the revenue schedule toward high profit margin products—suggesting that agents could earn more by selling these products—but acknowledged that the changes would lead to an overall decrease in commissions. Although management framed the changes as an opportunity for the workers to earn more, survey evidence in Section 3.6 indicates that the agents were aware that they were likely to take home significantly less pay because of the changes. We estimate that the commission schedule changes would reduce the commission

Figure 1. (Color online) Revenue Transfer Price Changes Within the Commission Schedule



Notes. Example of changes in the commission schedule for two different types of internet packages. The pre- and post-treatment revenue transfer prices for the basic package are displayed in the left branch. The pre- and post-treatment revenue transfer prices for the better package are displayed in the right branch. A basic package is easier to sell than a better package, captured by the higher quantity of sales per agent-week: 3 versus 1.5. The agent's commission rate, C, is multiplied by the product of the revenue transfer price and quantity sold to determine the amount of commission pay the agent receives for selling a particular package.

pay of the average Division 1 agent by 18%, holding fixed the pre-treatment period mix of products sold.

According to the firm's management, the changes to the commission schedule were intended to decrease the relatively high commission pay levels that Division 1 agents were earning in the months before the changes. These relatively high commissions were caused by the addition of new territories from which Division 1 agents fielded calls. The inclusion of these new territories (henceforth, the territory shock) significantly increased the average commissions of Division 1 agents. Figure 2(a) shows the evolution of average commissions by division before and after the commission schedule changes. The pre-treatment period, Week −26 to Week 0 (with Week 0 denoting the week before the commission schedule changes), is separated into three periods around the territory shock. The weeks before Week −16 constitute the pre-territory-shock period. The territory shock period runs from Week −16 to Week −8, representing the period of increasing commission levels for Division 1 agents. The period from Week −8 to Week 0 makes up the post-territory-shock period. Division 1 agents' average commission levels increased from $157 in the pre-territory-shock period to $318 at the beginning of the post-territory-shock period. The effects of the territory shock stabilized in the eight weeks before the commission schedule changes.

Because we learned of the impending commission schedule changes before they were announced, we followed the insider econometrics approach advocated by Bartel et al. (2004) and interviewed presidents and managers at the firm to assess their predictions for agent reactions. The president and managers in Division 1 believed an agent's responses to the changes would be muted. Other leaders within the firm, however, expressed concern about increased turnover among affected agents. Few sales managers mentioned changes in effort, because, although the strength of incentives would fall, high-powered incentive pay would remain a significant component of agents' total compensation.

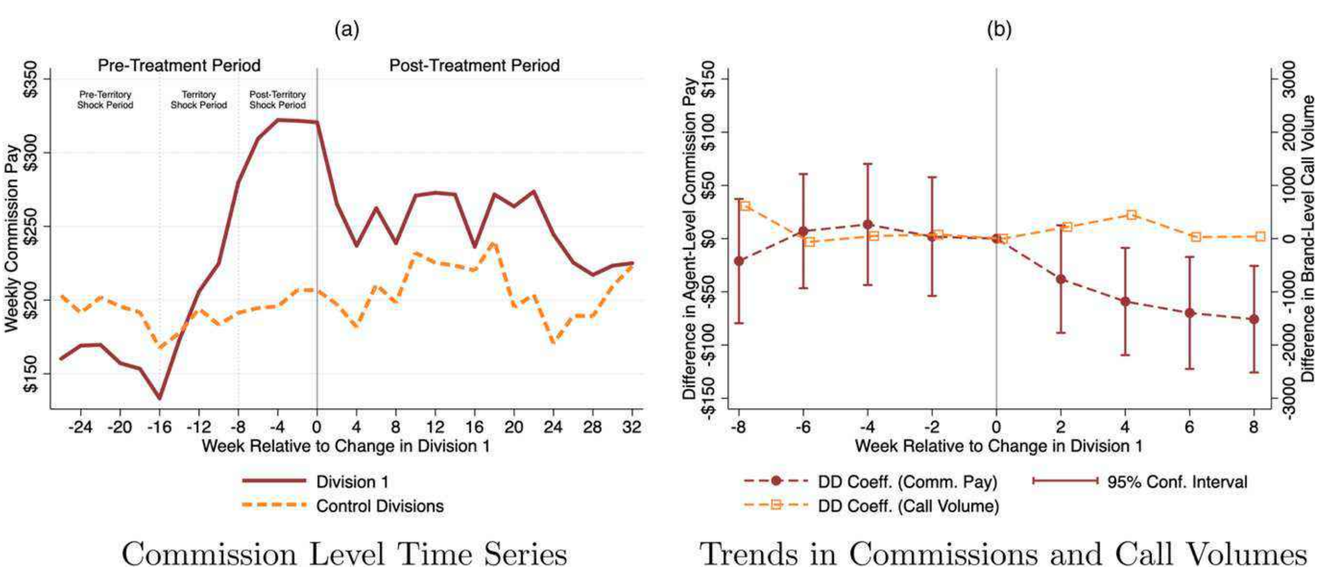
Despite managers' lack of focus on effort, agents did have discretion to meaningfully influence their sales. For example, managers reported that agents would often need to try several different approaches before successfully upselling a customer. A plausible way for agents to reduce their effort is by trying fewer approaches for selling high profit margin products and instead simply fulfilling orders for easier to sell products. Earlier attempts to increase sales with short-term incentives suggest—and subsequent academic experiments confirm—the ability for agents in this firm to adjust their effort (Sandvik et al. 2020). Another example of possible effort adjustment is reducing adherence to one's work schedule by taking more and longer breaks.[3]

Three months after the commission schedule changes occurred in Division 1, the president of Division 2 implemented similar changes.

3.4. Personnel and Productivity Data

We identify the consequences of the commission schedule changes using highly detailed commission,

Figure 2. (Color online) Commissions in Division 1 and the Control Divisions



Notes. (a) Average weekly commission pay levels for agents in Division 1 and the control divisions. The solid vertical line corresponds to the week immediately before the week of the commission schedule changes in Division 1. (b) Difference-in-differences coefficients that capture differential trends in commission pay levels and total call volume between Division 1 and the control divisions in the post-territory-shock period. We fail to reject the null hypothesis that the coefficients in Weeks −8 to 0 jointly equal zero (p = 0.82).

personnel, and productivity data provided by the firm. Division 1 and the control divisions all have consistent data beginning in April 2016. The sample is organized by agent-week and runs through June of 2017. This data set covers 2,033 unique sales agents across 61 weeks, for a total of 39,944 agent-week observations. This data set includes proxies of worker effort—for example, adherence, conversion rate, phone hours, average revenue-per-call (RPC), total revenue generated per week, demographic details—for example, age, race, tenure, gender, marital status, and commission pay data. We refer to this as the immediate sample because it has detailed productivity and commission data in the immediate period surrounding the changes in Division 1. A larger sample, used to study turnover, contains data beginning in July 2015, but it lacks information on sales productivity. We refer to this larger sample as the extended sample. Online Appendix B contains additional details about variable definitions and these samples.

Table 1 displays pre- and post-treatment period summary statistics for the control divisions (Divisions 3–6) in columns 1 and 4, Division 1 in columns 2 and 5, and Division 2 in columns 3 and 6. The pretreatment period is restricted to the post-territory-shock period (Weeks −8 to 0) to highlight the division-level characteristics immediately before the commission schedule changes occurred. We fail to reject the null at the 1% level that Division 1 and control division agents are similar in tenure ($p = 0.758$), age ($p = 0.336$), race ($p = 0.887$), and gender ($p = 0.496$). During the pretreatment period, agents in the control divisions and Division 1 were predominately male (70%–73%). The average agent was 25–26 years old and had been working at the firm for about a year. We do find that agents in the control divisions are significantly more

Table 1. Summary Statistics Pre- and Post-Treatment

Variables	Eight weeks pre-treatment: Control divisions (1)	Treated Division 1 (2)	Treated Division 2 (3)	All weeks post-treatment: Control divisions (4)	Treated Division 1 (5)	Treated Division 2 (6)
Commission	200.91	318.39	502.68	206.92	249.71	308.56
	(184.79)	(283.59)	(333.35)	(189.57)	(230.30)	(226.25)
Adherence	0.80	0.83	0.79	0.81	0.82	0.80
	(0.11)	(0.11)	(0.15)	(0.12)	(0.12)	(0.14)
Conversion	0.26	0.33	0.29	0.25	0.32	0.29
	(0.10)	(0.09)	(0.12)	(0.09)	(0.10)	(0.13)
Log RPC_{Old}	4.11	4.19		4.14	4.09	
	(0.47)	(0.51)		(0.49)	(0.52)	
Log RPC_{New}	4.11	3.96		4.14	3.90	
	(0.47)	(0.50)		(0.49)	(0.50)	
Phone hours	19.89	20.71	17.41	20.41	20.18	15.92
	(7.59)	(7.32)	(6.25)	(8.38)	(7.88)	(7.26)
Total calls	62.35	71.13	49.19	64.58	69.88	45.30
	(26.79)	(27.76)	(19.52)	(28.99)	(29.50)	(21.31)
Tenure (days)	356.56	369.02	672.98	450.36	399.81	608.06
	(419.52)	(389.51)	(558.98)	(505.66)	(411.39)	(594.50)
Age	25.84	25.18	29.71	26.20	25.99	28.33
	(7.15)	(6.50)	(8.64)	(7.32)	(7.75)	(7.98)
Single	0.52	0.68	0.44	0.38	0.44	0.33
	(0.50)	(0.47)	(0.50)	(0.49)	(0.50)	(0.47)
White	0.71	0.72	0.60	0.68	0.66	0.62
	(0.45)	(0.45)	(0.49)	(0.47)	(0.48)	(0.49)
Male	0.70	0.73	0.70	0.73	0.73	0.64
	(0.46)	(0.44)	(0.46)	(0.44)	(0.44)	(0.48)
Agent-weeks	4,024	867	357	13,817	3,474	950
Agents	632	138	51	874	234	89

Notes. This table presents summary statistics for the control divisions, Division 1, and Division 2. The *Commission* measure is average weekly commissions; *Adherence* is a measure of schedule adherence, which captures the amount of time an agent is available to take calls; *Conversion* is the probability of having positive sales revenue on a given call; *Log* RPC_{Old} measures an agent's revenue-per-call (RPC) if the commission schedule had *not* changed; *Log* RPC_{New} measures an agent's revenue-per-call (RPC) if the commission schedule had always been at the new levels. *Phone Hours* capture the amount a time an agent spends talking with customers; and *Total Calls* is the number of calls fielded by an agent each week. Data limitations prevent us from measuring *Log* RPC_{Old} and *Log* RPC_{New} for Division 2. Standard deviations are reported in parentheses.

likely to be married ($p = 0.001$). Division 1 agents have higher adherence ($p < 0.001$), but both groups are at or above the firm's mandatory level of 80%. Both groups spend a similar number of hours talking to customers each week ($p = 0.132$), although Division 1 agents realize higher commissions and greater revenue-per-call ($p < 0.001$), largely because of the territory shock experienced two months before the compensation changes occurred.

Agents in Division 2 earn much more in commissions ($p < 0.001$) because they sell to small businesses rather than residential customers. For expositional ease and because we do not have the full range of performance variables for Division 2, we focus our analysis on estimating the changes in Division 1 relative to control divisions.[4] We discuss the effects of commission changes for Division 2 in Section 5.5.

3.5. Estimating Baseline Agent Productivity

A central test of theories around heterogeneous turnover, emphasized in the adverse selection discussions in Campbell and Kamlani (1997) and Bewley (1998), concerns how agents of different productivity levels respond to a change in their compensation or employment contract. To identify these heterogeneous responses, we begin by estimating agents' productivity (fixed effects) before the commission schedule changes. To do this, we use a fixed effects regression for individual agents, controlling for tenure and division-by-time fixed effects. We extract the individual agent fixed effects and adjust them using a shrinkage procedure designed to limit the influence of measurement error, as in Lazear et al. (2015). These adjusted fixed effects are what we use for the pre-treatment measures of agent productivity. Additional details are provided in Online Appendix B.

Worker fixed effects from the per-period display wide baseline performance variability. For example, average revenue-per-call for Division 1 agents in the top tercile of agent fixed effects was more than 50% higher than the revenue-per-call produced by agents in the bottom tercile. Table A.1 provides additional summary statistics for Division 1 in the pre-treatment period by splits of the sample into terciles of pre-treatment productivity. Agents in the top tercile have higher tenure, in line with the firm retaining highly productive workers. Demographic characteristics also vary across the adjusted worker fixed effects terciles; namely, workers in the highest tercile are older and less likely to be single.

The interpretation of our upcoming analysis would be muddled if the commission schedule changes affected high and low performers differently, due to their selling of different product mixes. We test for this possible confounding factor in Online Appendix B.2 and find that the expected percentage change in commissions was equal across agents in the three terciles of fixed effects in Division 1.

3.6. Surveys of Sentiment and Reactions to the Changes

We conducted a firm-wide survey before the announcement of the changes to gather information regarding agents' sentiment toward the firm. We asked sales agents from all divisions the following three questions. (1) "How likely are you to agree with the following statement: [the firm's] policies, for example on adherence, compensation, and promotion, are justified and fair?" (2) "Suppose your friend is looking for a job, how likely are you to recommend them to apply at [the firm]?" (3) "Do you think you will be promoted in the future?" In addition, we conducted a follow-up survey among agents in Division 1 after the announcement of the changes and before these agents received their first paycheck under the new commission schedule. We asked the same three initial questions and several additional questions related to their perceptions of the commission changes. Additional details of these surveys are provided in Online Appendix B.3.

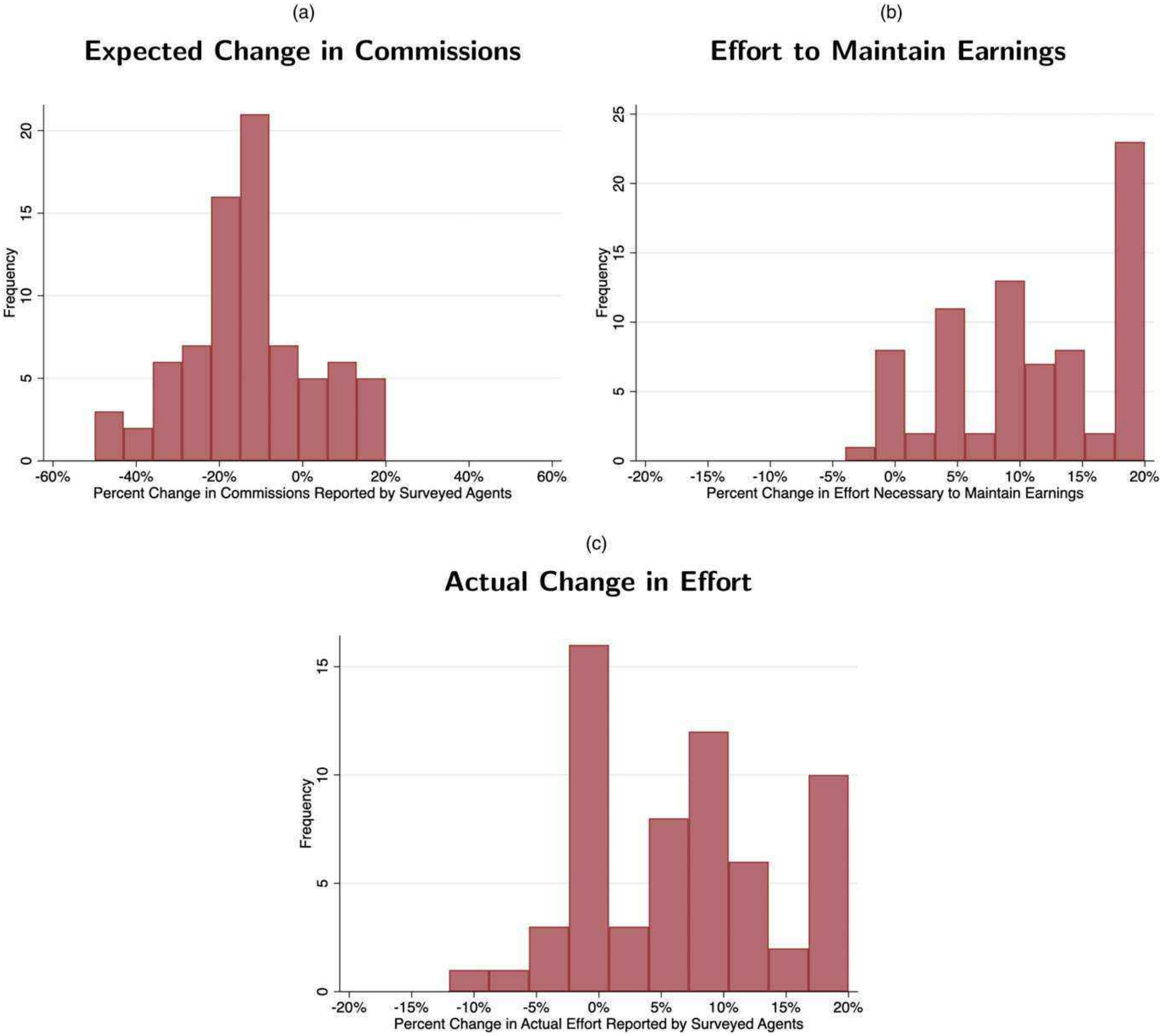
Responses to the follow-up survey conducted after the announcement reveal that the average agent in Division 1 expected their commission pay to decrease by 13% (Figure 3(a)), which approximates our estimate of an 18% decline in commissions. The average agent reported that they *would need to* work 11% harder in response to the changes to maintain their usual commission pay (Figure 3(b)). Agents then reported that *they would*, on average, increase their effort by 7% in response to the changes (Figure 3(c)). Agents' own responses to these last two questions suggest that effort may have actually increased, possibly because of income effects or the desire to maintain their prior earnings.

Several other questions on the follow-up survey asked agents about the motivation for the commission schedule changes. More than 75% of the agents felt the motivation for the changes was clearly communicated by management at the time of the announcement. When asked why the changes occurred, 42% responded with "[management thought] sales reps were overpaid," and 40% responded with "[The firm] needs to make cutbacks to stay in business." The follow-up survey also provides evidence that the changes were unanticipated by the sales agents, because only 2% of the agents say they knew the details of the changes before they were announced.

4. Identifying Assumptions and Common Trends

We are interested in how the commission schedule changes impacted the turnover and effort of the

Figure 3. (Color online) Reported Changes in Commissions and Effort



Notes. (a) Survey responses to a question asking agents what their expected change in commissions would be because of the commission schedule changes. (b) Responses to a question asking how much agents' effort would need to change to maintain their normal level of earnings. (c) Responses to a question of what changes in effort workers actually planned to make. See Section 3.6 for more details.

affected agents. The presence of unaffected divisions motivates the use of a difference-in-differences estimation. Difference-in-differences relies on the assumption that treated and untreated groups follow a common trend in outcomes in the absence of the commission change. To assess the suitability of using other divisions as a control group, we consider trends across several variables. First, we provide evidence of common trends in the attrition rates of agents in Division 1 and the control divisions over many months leading up to the commission schedule changes. We then show common trends in several output measures, which proxy for effort. Finally, we show that agents in Division 1 and the control divisions follow common trends in commission pay after the territory shock and before the commission schedule changes.

To bolster confidence that the common trends assumption is satisfied, we show that there is no divergence in proxies for effort supply (adherence and conversion rates) or effort demand (call volume and phone hours) either before or after the commission change in Division 1. Given that these auxiliary productivity measures do not deviate across divisions suggests that the commission changes in Division 1 were not motivated by future knowledge of call volume changes or other coincident issues that would confound our analysis. That is, the smooth evolution of these measures across treated and control divisions suggests that potentially problematic trend divergence is unlikely in our setting.

4.1. Common Trends in Turnover

Our first outcome of interest is the turnover response of Division 1 agents. We graphically assess the pretreatment trends in agent turnover between Division 1 and the control divisions. We use the Kaplan-Meier survival rate estimator, which plots retention rates over time, because it allows visualization of the cumulative nature of turnover. The survival rate estimator considers a starting point and then, from that time, displays the fraction of agents who remain at the firm. This allows us to detect when retention rates diverge and what fraction of the total beginning workforce is affected.

Figure 4 plots the survival rates for agents in Division 1 and the control divisions. To focus on heterogeneous turnover responses, we separately plot the survival rates of high and low performers. Figure 4(a) shows that highly productive workers in both Division 1 and the control divisions follow a similar trend in retention from Month −5 to Month 0. Similarly, low performers in Division 1 have survival rates that closely track those of low performers in the control divisions. The similarity of these survival rate trends suggests that agents in the control divisions provide a valid comparison group to estimate the turnover responses of agents in Division 1.

4.2. Common Trends in Effort

To evaluate the credibility of the assumption of common trends in effort, we estimate time-period differences between Division 1 and control divisions in an event study design and then plot the coefficients and confidence intervals (Fowlie et al. 2018, Cengiz et al. 2019). The functional form is

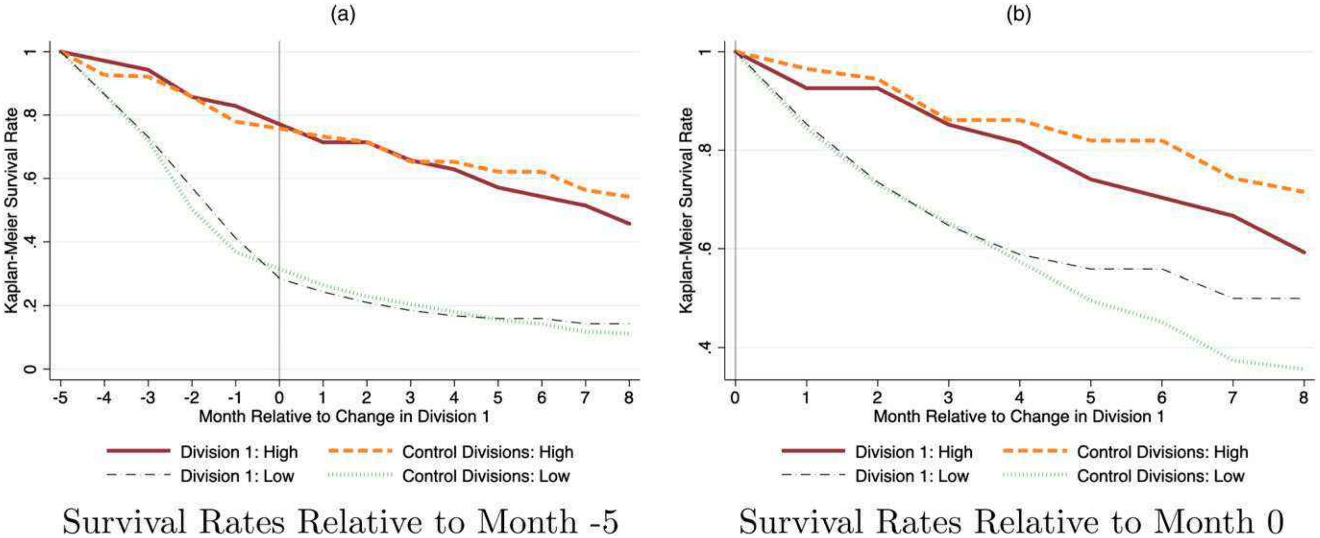
$$y_{i,t} = \sum_t \delta_{i,t}\mathbb{I}(time = t) \times Div1_i + \beta_i Div1_i + \sum_t \lambda_t \mathbb{I}(time = t) + X_{i,t}\Gamma + \sum_j \gamma_j Div_j + \varepsilon_{i,t}. \quad (1)$$

The main coefficients of interest are $\delta_{i,t}$, which capture differences in baseline time effects for agents in Division 1 relative to the common time effects for the control divisions, λ_t. The model also includes controls for location and agent characteristics in $X_{i,t}$ and division fixed effects, Div_j.

We find support for common trends in output-based proxies of effort. We discuss the main results here, but for formal detail on statistical tests and the graphical representations of these estimations, we refer interested readers to Figure 5. When examining adherence and conversion, which are two measures of effort supply, we fail to reject the null hypothesis of divergent trends in the pre-treatment period. We also cannot detect pre-treatment trend differences in call volume and phone hours, which are two measures of demand for worker effort.

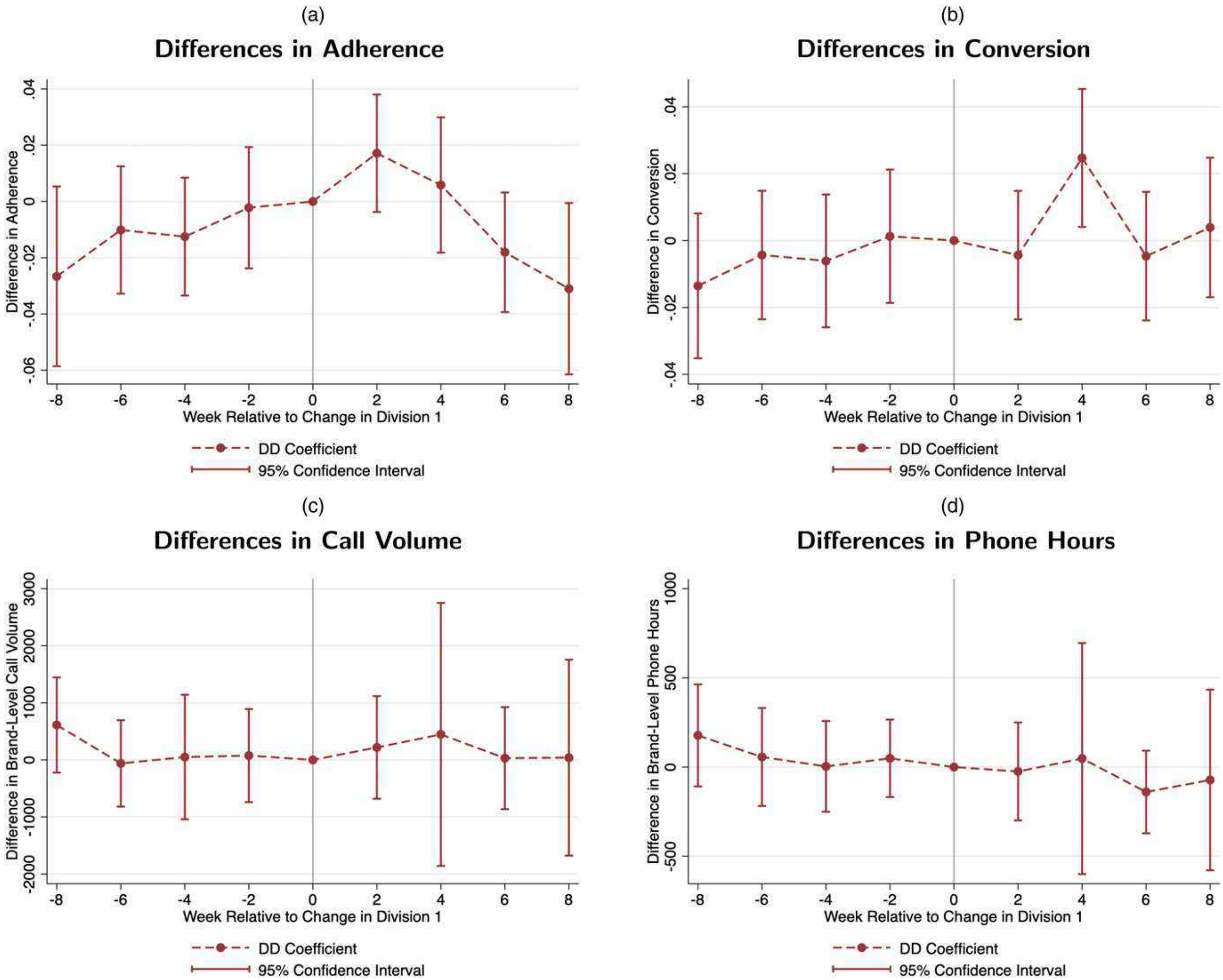
One might worry that trends in effort demand might deviate after the commission changes if the

Figure 4. (Color online) Survival Rates by Productivity



Notes. These figures plot Kaplan-Meier survival rates over time. The survival rate estimator considers a starting point and then, from that time, displays the fraction of agents that remain at the firm. Because turnover can be lumpy, with multiple exits in some weeks and no exits in others, we aggregate survival rates to the monthly level. The sample is split by high and low performers based on whether agents' adjusted worker fixed effects are above or below the median within their division.

Figure 5. (Color online) Trends in Proxies for Effort Supply and Effort Demand



Notes. The coefficients in these figures are estimates of $\delta_{i,t}$ from Equation (1), using different outcome variables of interest. Adherence and conversion are the two proxies for an agent's supply of effort. Call volume and phone hours are the proxies for customers' demand for worker effort. To improve the readability of these figures, we aggregate data into biweekly bins. The *p*-values of tests that the Week −8 to Week 0 point estimates are jointly equal to zero are 0.45, 0.70, 0.36, and 0.76 for adherence, conversion, call volume, and phone hours, respectively. The *p*-values of tests that the Week 0 to Week 8 point estimates are jointly zero are 0.98 and 0.76 for call volume and phone hours, respectively.

treatment is correlated with managements' forecasts of how the environment might evolve. This does not appear to be the case. To show that call volumes and the amount of time spent working with customers do not change coincidentally with the commission schedule changes, we test that the point estimates are jointly equal to zero in the eight weeks after the effective date of the change. We cannot detect any divergence in call-based measures of effort demand that occurred simultaneously with the commission changes in Division 1. Thus, any change in worker output that we observe is not likely because of reduced call volumes or time spent talking to customers. This suggests the effects of the territory shock, which had previously shifted demand, were permanent and had stabilized in the two months preceding the commission schedule changes.

4.3. Common Trends in Commissions

Figure 2(a) shows that commission pay in Division 1 and the control divisions follows a common trend before the territory shock, despite differences in levels. The commission levels of Division 1 agents deviate from this common trend during the territory shock period, but they appear to level off in the post-territory-shock period and again track the commission trends of agents in the control divisions. The implementation of the commission schedule changes again shocks the trend of Division 1 commission levels after Week 0, but the two groups appear to follow similar trends from Week 4 through at least Week 16. There is relatively little movement in the control divisions in the immediate aftermath of the Division 1 compensation changes, suggesting that the changes had limited spillover effects into other divisions. We further discuss tests that show limited spillovers to the control divisions in Online Appendix C.

In our setting, we expect common trends in commission pay levels after the territory shock. We focus on this eight-week period immediately prior to the commission schedule changes because we know that trends differed during the territory shock period. Figure 2(b) plots the coefficients, $\delta_{i,t}$, estimated using Equation (1), just as was done to assess the common trends in effort. We find evidence of common trends in commission levels when we plot the coefficients across time. The point estimates in Weeks −8 to −2 are all close to zero, and zero always exists within the 95% confidence intervals around these points. Furthermore, we fail to reject the null hypothesis that the coefficients jointly equal zero ($p = 0.82$). This result suggests that Division 1 and the control divisions followed common trends in commission pay levels in the two months before the commission schedule changes occurred. We overlay a plot of differences in brand-level call volume to show that the observed differences in commission levels after the commission schedule changes are not driven by brand-specific variations in call volume. Managers of the firm confirm that, in the absence of the commission schedule changes, agents in Division 1 would have continued to realize the high commission levels they enjoyed in the post-territory-shock period.

The analysis suggests that trends in commission levels caused by the territory shock in Division 1 are not a major concern for our empirical approach. Instead, the most likely issue for interpreting estimates in light of the territory shock is the loss of *balance* between Division 1 and other divisions because the territory shock potentially changed agents' reference points or caused the job to become relatively more attractive than it had been beforehand. Because our counterfactual compares agents with better jobs and higher earnings to the control agents, our estimates of turnover and effort responses are likely lower bounds for the consequences that managers would otherwise anticipate when adjusting pay.

5. Results and Exploration of Mechanisms

This section details Division 1 agents' turnover and effort responses to the changes in their commission schedule. We also estimate the firm-level effects of the observed worker responses. We then consider the role of sentiment for our findings and discuss the turnover effects of Division 2 agents. Our empirical analysis is motivated by a theoretical model, which, for brevity, we discuss in Online Appendix D. The model provides context for our estimates by showing that whether a compensation change is profitable depends on (1) the cost changes that affect the firm's wage bill, (2) changes in workers' effort, and (3) changes in the composition of the workforce, due to asymmetric turnover based on agent productivity.

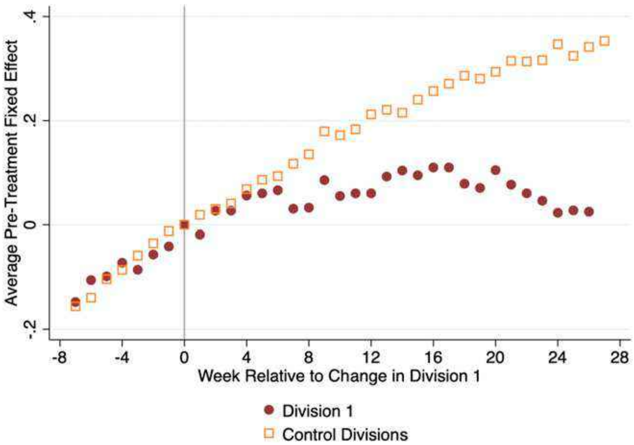
5.1. Turnover Responses

In Section 4.1, we introduced Figure 4(a) to show the common trends in attrition between agents in Division 1 and control divisions in the months before the commission schedule changes. This figure also shows that the survival rates of highly productive agents in Division 1 break from those of highly productive agents in the control divisions in the post-treatment period. This figure conditions on agents who were present at the firm several months before the commission schedule changes, which is useful as a diagnostic tool for pretrends. Figure 4(b), on the other hand, considers survival rates relative to the sample of agents present in each group in what is labeled Month 0 (October 2016), the calendar month immediately before the changes occurred. After the commission schedule changes, the survival rate of high performers in Division 1 decreases, relative to that of high performers in the control divisions, whereas the survival rate of low performers appears to increase. This is preliminary evidence of a heterogeneous turnover effect, wherein highly productive agents in Division 1 are more likely to leave the firm in response to the commission schedule changes.

Figure 6 presents how this differential turnover influences the composition of agents who remain by plotting the average z-score of adjusted worker fixed effects for Division 1 and the control divisions. As in many sales firms, there is positive selection by worker quality over time, captured by the upward trend in average adjusted worker fixed effects in all divisions in the pre-treatment period (represented by points to the left of the vertical line). There is then clear evidence that average worker quality begins to deteriorate in Division 1 several weeks after the commission schedule changes. By 24 weeks after the change, the average fixed effects for Division 1 fall by more than 0.3 standard deviations relative to control divisions. This divergence in adjusted worker fixed effects provides graphical evidence that, in response to the commission schedule changes, agents with high pre-treatment productivity exited the firm at a higher rate than did agents with low pre-treatment productivity.

With this evidence in hand, we formally examine turnover by using a difference-in-differences estimator. These estimations use the extended sample, which includes additional data predating the immediate sample by at least a full calendar year for each

Figure 6. (Color online) Adjusted Worker Fixed Effects Before and After the Compensation Changes



Notes. Average adjusted worker fixed effects for Division 1 and the control divisions after taking a z-score transformation. Because the fixed effects are calculated before the commission changes, the data are limited to agents who were at the firm before the commission changes. Adjusted worker fixed effects are calculated from a regression of log commissions on worker dummy variables, division-by-week dummy variables, and a cubic spline in tenure. We then correct for sampling variation using the method in Lazear et al. (2015). The series are normalized to correspond at the announcement date, which is depicted by the vertical line.

division. Our goal is to identify how turnover was affected when agents' commission schedule changed and how the turnover probability differs based on agent productivity. In an analysis of turnover, it is necessary to account for how the baseline probability of leaving the firm changes with worker tenure (Bartel and Borjas 1981). We use a very flexible specification for how the usual probability of leaving the firm changes with tenure by including a flexible function, $g(Tenure)$, in the model. We specify this function as a fifth-order polynomial, providing enough flexibility to capture the possibility that workers with longer tenures are less likely to leave and that the relationship between attrition and tenure has several inflection points.[5] This function is distinct from time fixed effects, which are meant to capture calendar time shocks, like seasonality, that affect all workers. We then include combinations of division and time fixed effects to capture permanent heterogeneity across divisions and seasonal shocks that may be correlated with treatment. The model we estimate is

$$\begin{aligned} Turnover_{i,t} = Div_j &+ \delta_1(Treated_i \times Post_t) \\ &+ \delta_2(Treated_i \times Post_t \times Prod_i) \\ &+ \beta_1(Prod_i) + \beta_2(Treated_i \times Prod_i) \\ &+ \beta_3(Post_t \times Prod_i) + TimeControls \\ &+ g(Tenure) + \beta_4 X_{i,t} + \varepsilon_{i,t}. \end{aligned} \quad (2)$$

The dependent variable, $Turnover_{i,t}$, is an indicator that the week in question is worker i's last week in the firm. After the worker leaves, he or she is no longer included in the sample. The dependent variable is thus the instantaneous turnover probability, or hazard, given that the worker was at the firm in the week in question. The parameter δ_1 captures the average change in turnover probability of agents in Division 1, conditional on tenure and time controls, after the commission schedule changes occurred. This is indicated by $Post_t$, the post-treatment indicator, being interacted with $Treated_i$. We include division fixed effects, Div_j, to control for division-level differences in attrition. The matrix $X_{i,t}$ has a third-order polynomial in age, along with fixed effects for ethnicity, gender, call center location, and marital status. The separate tenure splines and age polynomials allow the effects of experience within the firm and total labor market experience to differ. We include baseline measures of worker productivity, captured by $Prod_i$, and its interaction with postevent indicators. To identify differences in productivity, we use the standardized z-score of adjusted worker fixed effects in the pre-treatment period. We use z-scores to standardize the adjusted fixed effects across Division 1 and the control divisions. This approach also facilitates the interpretation of the parameters, as a unit change in the z-score, $Prod_i$, corresponds to a standard deviation of the underlying productivity measure.

Table 2 displays the turnover responses of agents in Division 1, relative to those in the control divisions. The different columns correspond to different combinations of *TimeControls* to account for a variety of possible temporal differences across divisions. Across all specifications, highly productive workers in Division 1 became more likely to leave after the commission schedule changes. The point estimates on $Treated \times Post \times Prod$ across columns 1–3 indicate that Division 1 agents with pre-treatment productivity one standard deviation above the mean had turnover rates that increased by 1.5–2.1 percentage points in a given week, compared with Division 1 agents with average pre-treatment productivity. This turnover increase is relative to an overall sample mean of about 3.7%, indicating that agents one standard deviation above the mean had between a 40% and 56% increase in attrition from the sample average.

These turnover effects are precisely estimated when clustering by the identity of a worker's manager. If we instead cluster standard errors at the division level, which was the level of the treatment, our standard errors are similar. However, we do not report these standard errors, because test statistics based on them are misleading because of having few divisions and only one treated group. Instead we conduct robust statistical tests using a combined randomization inference and wild bootstrap procedure designed to estimate critical regions under clustering with few

Table 2. Linear Probability Model Estimates of Turnover Responses

	Last week in firm				
	(1)	(2)	(3)	(4)	(5)
Treated × *Post* × *Prod*	0.021**	0.015**	0.016*	0.013**	0.012*
	(0.007)	(0.005)	(0.007)	(0.005)	(0.005)
Treated × *Post*	–0.006	–0.006		–0.002	
	(0.004)	(0.007)		(0.010)	
Treated × *Placebo* × *Prod*	–0.006		–0.002		
	(0.004)		(0.004)		
Treated × *Placebo*	0.000				
	(0.004)				
Week fixed effects	✓	✓		✓	
Division × *Week-of-Year* fixed effects		✓			
Week × *Division* fixed effects			✓		✓
Post-territory-shock period				✓	✓
Observations	51,497	51,497	51,497	19,689	19,689
Mean turnover probability in Division 1			0.037		
p-value on *Treated* × *Post* × *Prod*	0.017	0.081	0.096	0.036	0.040
p-value on *Treated* × *Post*	0.482	0.316		0.133	

Notes. The dependent variable is an indicator that equals one if it is the worker's last week at the firm. The sample includes all current employees in Division 1 and the control divisions with nonmissing data. Estimates come from a linear probability model that captures changes in the turnover probability for the existing workforce. Each model includes a fifth-order polynomial for workers' tenure to account for a potentially arbitrary baseline relationship between tenure and turnover. *Prod* refers an agent's sales z-score, which is the standardized measure of an agent's pre-treatment productivity estimated as their adjusted worker fixed effect according to the procedure in Lazear et al. (2015). For additional details, see Section 3.5 and Online Appendix B. The specification in column 2 includes division by week-of-year fixed effects to account for seasonality. The specification in column 3 includes week by division fixed effects. Columns 4 and 5 use a shortened pre-treatment period that only includes the weeks of data after the territory shock period. *Placebo* is an indicator for the date 52 weeks before the treatment date. Standard errors are clustered by manager (in parentheses). The p-values in the bottom two lines are computed after clustering by division and applying the wild cluster bootstrap randomization inference procedure in MacKinnon and Webb (2018). We use the t-statistic version of the procedure that imposes the null hypothesis.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

treated clusters (MacKinnon and Webb 2018).[6] The p-values from these tests are displayed in the bottom rows of Table 2 for δ_1 and δ_2.

Table 2 also includes placebo tests to assess whether the observed attrition patterns would have occurred at a different time. The most natural prior time to test is the exact date in the prior calendar year, and the specification in column 1 includes placebo indicators that are dummies for the period one year before the announcement week. This tests whether the observed turnover effect is because of the commission schedule changes or annual patterns in turnover. The zero coefficients on *Treated* × *Placebo* × *Prod* and *Treated* × *Placebo* indicate that the turnover patterns overall and by agent productivity level did not diverge between Division 1 and the control divisions at the same time in the past.

Columns 2 and 3 show that the estimates are robust to the inclusion of different combinations of time, week-of-year, and division fixed effects. The division by week-of-year fixed effects in column 2 compares division-level turnover rates across calendar years, which captures possible seasonality by division and guards against the possibility that Division 1 had a similar seasonal change in turnover in the prior year. The point estimate of 0.015 suggests that workers ±1 standard deviation around the mean had post-treatment turnover rates of 5.2% and 2.2%, respectively. This difference provides strong evidence of a heterogeneous turnover response across the distribution of worker productivity. Column 3 includes division-by-time fixed effects, which only allows us to identify heterogeneous turnover by productivity. The advantage of these estimates is that we do not need to rely on common trends by division but only common trends in turnover by different productivity groups.[7] The similarity of the overall estimates in columns 1 and 2 and those that do not depend on common trends at the division level (column 3) add credibility to the identifying assumptions. This heterogeneous effect also holds when we restrict the sample to begin eight weeks before the commission schedule changes occurred (columns 4 and 5), which removes the weeks before and during the territory shock period. The results are also robust to whether workers' productivity fixed effects include or omit their tenure.

Taken together, this evidence suggests that highly productive agents in Division 1 were more likely to leave the firm in response to the commission schedule changes, as predicted by managers in the prior literature (Campbell and Kamlani 1997, Bewley 1998).

We note that, although high performers were more likely to leave the firm following the change, the average turnover rate was unchanged. The point estimates on *Treated* × *Post* are not precisely estimated in any of the specifications. The inability to reject that the main effects are zero for Division 1 indicates that overall turnover did not increase among these agents. Instead, only agents who were highly productive in the pre-treatment period increased their likelihood of leaving. This is consistent with the trends in survival rates, depicted in Figure 4(b). This figure shows that high performers have an *increased* likelihood of quitting (i.e., a decreased survival rate), whereas low performers have a *decreased* likelihood of quitting. The departure of high performers potentially increased the placement of low performing agents in the selling efficiency quintiles used to determine commission rates. Consequently, the implicit contract improved for low performers with the attrition of high performers, increasing their incentive to stay in the firm. These offsetting effects provide intuition as to why we do not observe a significant average turnover effect among agents in Division 1.

The results from multiple additional placebo tests highlight the robustness of our turnover response estimations. Following Gubler et al. (2018), we perform 50 placebo simulations for the turnover response estimation using randomized treatment groups and treatment dates over agents. The coefficient for 46 of the 50 placebos is smaller in size and less statistically significant than the estimated coefficient. This is approximately what one would expect from the placebo tests given the statistical significance of the estimate. We also repeat this procedure by randomizing treatment over different control divisions. In this setup, our estimate is larger than all placebo estimates. The results of these placebo estimates are displayed in Figure A.2, (a) and (b), respectively.

5.2. Effort Responses

Having found evidence of heterogeneity in the turnover of agents in Division 1, we next investigate whether these agents altered their effort in response to the commission schedule changes. The first specifications for estimating the effects of the commission schedule changes on worker effort are difference-in-differences regressions with the following form:

$$y_{i,t} = \alpha_i + Div_j + Trend_j + \delta_1(Treated_i \times Post_t) + \lambda_t + \beta_1 X_{i,t} + \varepsilon_{i,t}. \tag{3}$$

The model includes time (week) fixed effects, λ_t, and division fixed effects, Div_j. Some specifications include an individual fixed effect α_i, and some include division-specific time trends, $Trend_j$. To account for the potential that different trends across divisions bias the estimates, we check the robustness of our results by using a propensity score reweighting estimator to match control division agents who were on similar trends as those in Division 1 before the commission schedule changes occurred. This approach aims to better balance treated and control agents, based on levels of and changes in compensation over the entire pre-treatment period.[8] In addition, we verify our results by reducing the sample to a balanced panel of agents who are present in the sample before July 2016 and after April 2017. This ensures that we capture variation in agents' behavior before and after the commission schedule changes and not just changes in the composition of workers.[9]

The results of the difference-in-differences estimations using Equation (3) are contained in columns 1–5 of Table 3. We use data from the eight weeks before and the eight weeks after the commission schedule changes to estimate agents' effort responses. In Table A.3, we show that our results are robust to the inclusion of all the pre- and post-treatment data. Panel A of Table 3 contains results for agents' adherence and shows that, on average, agents in Division 1 do not reduce their adherence in response to the commission schedule changes. We also find negligible differences in agents' conversion rates (Panel B). The null results in Panels A and B are robust to the inclusion of agent fixed effects (column 2), the inclusion of division-specific time trends (column 3), the use of a reweighting estimator (column 4), and the use of a balanced panel (column 5). These findings align with the graphical evidence presented in Figure 5, (a) and (b). They suggest that agents did not avoid calls nor did they reduce their sales conversion efforts; that is, we find little evidence of effort adjustment after the commission schedule changes.

We further consider changes in agents' effort by considering two additional proxies of worker sales effort, log revenue-per-call if (1) the commission schedule had *not* changed (Panel C) and (2) if the commission schedule had always been at the new levels (Panel D). In these specifications, we take the revenue transfer prices as given, based on the respective commission schedule regime and apply these pseudo-revenues to the volume of products sold. We find minimal evidence of changes in log revenue-per-call at both the old and new revenue levels.[10] The positive estimates in Panel D suggest that agents might have increased effort after the commission schedule changes, potentially to compensate for income

Table 3. Estimates of Effort Responses

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Adherence to schedule							
Treated × Post	0.003	0.006	0.025	0.022	0.018	0.029*	
	(0.009)	(0.010)	(0.014)	(0.019)	(0.019)	(0.013)	
Treated × Post × Prod						–0.005	–0.004
						(0.005)	(0.006)
Observations	8,647	8,647	8,647	7,570	3,979	8,647	8,647
Panel B: Conversion rate							
Treated × Post	0.004	0.002	0.010	0.006	0.008	0.016*	
	(0.007)	(0.005)	(0.007)	(0.008)	(0.007)	(0.007)	
Treated × Post × Prod						–0.020***	–0.020***
						(0.005)	(0.005)
Observations	8,283	8,283	8,283	6,903	3,743	8,283	8,283
Panel C: Log RPC at old prices							
Treated × Post	–0.025	–0.039	0.005	–0.016	0.005	0.022	
	(0.031)	(0.025)	(0.033)	(0.036)	(0.048)	(0.032)	
Treated × Post × Prod						–0.041	–0.042
						(0.024)	(0.024)
Observations	9,229	9,229	9,229	7,840	4,126	9,229	9,229
Panel D: Log RPC at new prices							
Treated × Post	0.027	0.007	0.027	0.002	0.016	0.046	
	(0.032)	(0.025)	(0.036)	(0.039)	(0.054)	(0.035)	
Treated × Post × Prod						–0.048	–0.049
						(0.024)	(0.025)
Observations	9,229	9,229	9,229	7,840	4,126	9,229	9,229
Week fixed effects	✓	✓	✓	✓	✓	✓	
Agent fixed effects		✓	✓	✓	✓	✓	✓
Division trend controls			✓	✓	✓	✓	
Week × division fixed effects							✓
Reweighted				✓			
Balanced sample					✓		

Notes. The sample includes all current employees in Division 1 and the control divisions with nonmissing data. The models in columns 1–6 include fixed effects for week, division, and office location. All models include cubic splines for tenure and a cubic polynomial in age. The ordinary least squares regression in column 1 includes dummies for ethnicity, gender, and marital status. The specifications in columns 2–7 include individual fixed effects. Columns 3–6 include division-specific trend controls. The specification in column 4 uses a reweighting estimator based on the propensity score for being in Division 1 (see Online Appendix C.1). The balanced panel in column 5 restricts to workers who are present prior to July 2016 and after April 2017. Columns 6 and 7 consider heterogeneous responses based on worker productivity, and column 7 omits week fixed effects and division-specific trend controls and instead includes week by division fixed effects. Differing numbers of observations across panels reflect differences in data availability. The sample used restricts to eight weeks of pre-treatment data and eight weeks of post-treatment data. The results are similar when all available pre- and post-treatment data are used (Table A.3). Reported standard errors are clustered by manager.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

they stood to lose. However, this finding is not precisely estimated in any of the specifications.

We find limited evidence that agents were able to substantially shift from low- to high-margin products. If workers were substituting to higher margin products under the new commission structure, we would have expected to see substantial divergence between the results using the old and new revenue schedules in Panels C and D. Instead, both sets of estimates include 0 in the confidence intervals, suggesting minimal ability to substitute to higher margin products.[11] In some specifications, however, we are able to reject the null that the coefficients using new or old prices are the same. Specifically, in columns 1 and 2, we reject equality at the 1% level, and in column 3, we reject equality at the 10% level. However, in columns 4 and 5, which use a reweighting procedure and a balanced sample, respectively, we cannot reject equality at the 10% level. Comparing these differences, their estimated magnitudes are generally small.[12] Substitution to higher margin products would have implicitly reduced the magnitude of the compensation changes experienced by the agents. As a result, the relationship between the compensation changes and turnover that we estimate is likely a lower bound for the turnover that would have materialized absent product substitution.

A second specification identifies heterogeneous effort responses across agents, based on their pretreatment productivity.[13] This specification includes interactions of productivity pre-treatment $Prod_i$, with variables in the model in Equation (3):

$$\begin{aligned} y_{i,t} = \alpha_i &+ \left(Div_j \times \lambda_t\right) + \delta_1(Treated_i \times Post_t) \\ &+ \delta_2(Treated_i \times Post_t \times Prod_i) \\ &+ \beta_1 Prod_i + \beta_2(Post_t \times Prod_i) \\ &+ \beta_3(Treated_i \times Prod_i) + \beta_4 X_{i,t} + \varepsilon_{i,t}. \end{aligned} \tag{4}$$

Column 6 of Table 3 reports both δ_1 and δ_2 from Equation (4), capturing the fact that highly productive agents may have different effort responses on some dimensions. Column 7 identifies only the parameter δ_2 by including division-by-time fixed effects in the model. We do not find a heterogeneous reduction in adherence across agents of varying productivity levels (Panel A), which suggests that neither high performers nor low performers responded to the commission schedule changes by avoiding calls or disregarding their schedules. In Panel B, we find that the conversion rates of highly productive agents decreased, relative to those of less productive agents. Similarly, the negative coefficients on *Treated* × *Post* × *Prod* in Panels C and D suggest that high performers may have reduced their revenue generation per call, relative to the average agent, but the effects are not precisely estimated.

A potential concern with the estimations of productivity changes is the possibility of mean reversion, which may be amplified in short time periods. Several empirical facts suggest a limited role for mean reversion in the productivity data we present, but we acknowledge the possibility. First, it is unlikely that mean reversion drives these results because Equation (4) accounts for this through $\beta_2(Post_t \times Prod_i)$. The parameter δ_2 on $(Treated_i \times Post_t \times Prod_i)$ thus captures any deviation from natural agent productivity mean reversion in the postperiod.[14] In addition, mean reversion would likely bias us toward finding substantial changes in effort because sales in Division 1 began at a higher level than in the control divisions, and we would thus expect Division 1 sales to fall under mean reversion. Given the modest size of the estimated sales reductions, we expect that mean reversion is unlikely to be driving these results. Taken together, the results in Table 3 imply that agents of all productivity levels had rather muted effort responses to the commission changes.

Finally, none of these results on effort report the wild cluster bootstrap randomization p-values and because of the general insignificance of the findings. Figure A.4 shows placebo tests where we repeat our effort estimation procedure using different control divisions as the chosen *treated* division. As the figure shows, we cannot reject the null of zero changes in effort when control divisions proxy for the treated division.

5.3. Implications for Profitability

Having estimated both the turnover and effort responses of agents in Division 1, we next estimate the overall return on investment stemming from the compensation changes. Although the firm initially saved money as the result of paying fewer commissions in Division 1, over time, the lost revenue from high performing agents who left the firm outweighed the initial compensation savings.

At the outset, the cost savings from the commission schedule changes looked attractive, saving the firm about \$0.68 in compensation expense per call. Because turnover was minimal in the first few weeks after the changes, there was no offsetting reduction in revenue. However, about two months after the changes (eight weeks), the workforce composition effect reduced average revenue-per-call by \$0.58 compared with Week 8 labor cost savings of \$0.71. Over time, the decrease in the average revenue-per-call grew more quickly than the cost savings. About four months post-treatment (18 weeks) is the inflection point where the change became unprofitable. Six months after the changes, the firm's gross margin per-call fell by more than 1.7 percentage points.

To put these numbers into context, we estimate the total net present value of the commission schedule change by multiplying the per-call numbers by the actual number of calls per week. Using just a six-month horizon, the present value of the commission schedule changes totaled negative \$75,500. We emphasize that this estimate likely understates the impact for the firm because we do not include the costs of training new hires. Additionally, our analysis does not consider the spillover effects associated with losing high performers. Previous work has shown that high-performing employees are an important resource for raising the productivity of others (Sandvik et al. 2020), so the loss of highly productive workers likely had a deleterious effect on long-term productivity, beyond the six-month horizon. We provide details behind these NPV calculations in Online Appendix B.4.

5.4. Commission Schedule Changes, Worker Sentiment, and Mechanisms

We now turn to additional evidence on the mechanism behind our results. We investigate three different questions regarding whether agents' sentiment toward the firm or perceptions of fairness can explain our findings. Before the announcement of the commission changes in Division 1, we surveyed agents from all six divisions about their perceptions of firm fairness, their willingness to give referrals, and their future promotion prospects. The exact wording of these questions is provided in Section 3.6. Shortly after the commission changes, we again surveyed agents in Division 1 to see how their answers changed.

How do these responses vary over the performance distribution? Do the highest performers (who eventually leave the firm) also have the most negative responses regarding whether the firm became less fair? Whether they would be less likely to refer others to work at the firm? The first row of Table 4 shows no significant changes in either high or low performers' perceptions of the firm's fairness. Instead, the second row reports that, across all terciles of pre-treatment productivity, agents reduced their reported willingness to refer others to work at the firm, but the decline was greatest among high performers. The 19.8 percentage point decline of high performers indicates a substantial reduction in perceived firm quality and is much larger than the 5.3 percentage point decline among the lowest tercile of agent productivity. The third row shows relatively small changes in agents' perceptions of their own promotion prospects. Based on these changes in survey responses, fairness channels have less support as a mechanism because perceived fairness reductions do not load differentially for the high-performing agents who ultimately leave the firm. Instead, high-performing agents have the largest reduced perception of the quality of their current job, presumably relative to other employment options.

We also investigate whether agents' turnover and effort responses vary with differences in pretreatment survey responses. We use agents' pretreatment survey responses, as we lack data on changes over time for control divisions. Among Division 1 agents where we can measure changes, Table A.5 shows that those who had the most positive responses before the change generally had the largest reductions after the change for each of the survey questions. We separately interact the treatment indicator with indicators for high firm fairness perceptions, high referral likelihood, and a belief that promotion is likely. We continue to find that highly productive workers in Division 1 increased their turnover rates after the commission schedule changes, relative to the average worker in Division 1. We do not, however, find any significant heterogeneity in treatment effects across responses to the three survey questions. The estimates are close to zero and are not statistically significant, as reported in Table A.6. A similar analysis for agent effort in Table A.7 reveals no evidence of statistically significant heterogeneity. Taken together, these results fail to find differential turnover and effort effects that function through ex ante proxies for sentiment.

We caution, however, that these latter tests may be underpowered for ruling out fairness channels. Although the survey evidence finds a limited role for the fairness explanation, there are several key limitations. First, agents may not have internalized the impact of the change at the time of the follow-up survey, because this survey occurred before the agents' first post-treatment paycheck. Second, the response rate of the follow-up survey is 30%, possibly inducing selection bias. Third, we cannot use changes in sentiment as the interactive variable of interest, as only

Table 4. Sentiment Descriptive Statistics

	All	Pre-treatment productivity (z-score) 0%–33%	33%–66%	66%–100%	Difference
	(1)	(2)	(3)	(4)	(4)–(2)
Δ*Fairness Perceptions*	−1.43	−2.74	−3.92	2.48	5.22
	(2.74)	(4.35)	(5.12)	(4.80)	(6.48)
Δ*Referral Likelihood*	−12.51***	−5.30	−11.52**	−19.77***	−14.46**
	(2.90)	(3.32)	(4.85)	(5.93)	(7.04)
Δ*Promotion Prospects*	−0.17**	−0.04	−0.33**	−0.14	−0.10
	(0.07)	(0.10)	(0.16)	(0.12)	(0.16)
Agents	70	23	24	23	

Notes. This table documents the changes in the self-reported sentiment levels of Division 1 agents from before to after the commission schedule changes. We split the data based on terciles of pre-treatment agent productivity. The results of difference-in-means tests between columns 4 and 2 are reported in the far right column. Standard errors of means are reported in parentheses.

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Division 1 agents took the follow-up survey. Finally, the survey questions about fairness encompass many aspects of the job and not just pay considerations. Thus, although workers' fairness concerns do not appear to be driving our results, they cannot be definitively ruled out.

5.5. Effects of the Commission Schedule Changes in Division 2

For expositional ease, we deferred the discussion of the commission schedule changes in Division 2 until now. Division 2 changes allow us to test whether our main findings generalize. We begin our analysis of Division 2 by discussing the trends in commission pay levels for Division 2 agents, relative to control division agents, before and after their commission schedule changes. We then report the results from our difference-in-differences analysis of turnover responses. Finally, we corroborate these findings with graphical evidence that Division 2 workers increased their propensity to quit after the commission schedule changes. We do not consider the effort responses of Division 2 workers, as data limitations prevent us from measuring changes in revenue-per-call for this division.

Figure A.5(a) shows the evolution of average commission levels in Division 2 and the control divisions before and after the changes. As Division 2 agents sell products to small businesses, their sales and commissions are highly cyclical. Taking the average of Division 2 commissions before and after the change, commission pay fell from $392 to $309 per week. In Figure A.5(b), we again overlay a plot of call volume to show that the observed drop in commissions is not driven by a decrease in the number of calls.

The commissions change in Division 2 caused the average turnover rate to double, from a baseline of 0.83% per week to 2% per week (see Table A.8 with a coefficient of 0.013). Figure A.6, (a) and (b), corroborates the estimates by showing the Kaplan-Meier survival rates for agents in Division 2, relative to the control divisions. We do not detect within-division heterogeneous turnover responses in Division 2 but note that these agents are highly productive and highly compensated relative to the rest of the firm. Division 2 agents come from the right tail of the firm-wide productivity distribution, resembling the best agents in Division 1. Highly productive agents in Division 1 and all agents in Division 2 would have had similar outside options and similar incentives to search for other employment after their commission schedules changed, effectively reducing their take-home pay. This may explain why we observe a heterogeneous turnover effect in Division 1 and an overall increase in turnover in Division 2.

Our relatively limited data for Division 2 also prevent us from performing the same NPV calculation for Division 2 as we performed for Division 1 in Section 5.3. The Division 2 compensation changes were also likely NPV negative in the long-term given the results from Division 1 and the large increase in attrition of Division 2 workers.

6. Concluding Discussion

The strategy and management literature has long studied the importance of incentive compensation in attracting and retaining workers, with a focus on top-talent (Zenger 1992, Campbell et al. 2012, Carnahan et al. 2012). We extend this literature by showing how workers of varying ability respond differently to incentive compensation *changes*. Specifically, we study effort and retention effects associated with a compensation change that reduced overall pay by 7% in an inbound sales call center.

We find that the most productive workers, those with pretreatment productivity one standard deviation above the mean, increased their turnover likelihood by between 40% and 56%. By contrast, other employees had minimal responses to the compensation changes. Apart from the attrition of high performing workers, we find limited changes in effort, sales revenue, and other on-the-job performance measures. The increased attrition of the most productive employees took several weeks to manifest, and we find no evidence that ex ante worker sentiment about their jobs or the firm drove the increase in turnover. Instead, highly productive workers likely left the firm on securing more attractive outside options, whereas their less productive peers were unable, or unwilling, to do so.

We find that the compensation change was ultimately unprofitable over the long-term. The change allowed the firm to reduce their payroll expenses, but revenues decreased because of the subsequent departure of highly skilled workers who were replaced by less experienced and less productive workers. Given that measured effort and productivity for retained workers did not change—although separation of highly skilled workers occurred with a lag—analysis of average turnover rates alone, or looking at a relatively narrow time window, would have incorrectly concluded that the changes were profitable. Combined, these results highlight the importance of (patiently) measuring responses across the productivity distribution when one evaluates the net effects associated with incentive changes and—more generally—human resource management changes.

Although our study firm provides an ideal setting to estimate heterogeneous worker responses to a compensation change, there remain important questions that we are unable to answer. First, our data are only from one firm, which limits our ability to determine whether employees who leave are making optimal decisions.

This is especially important in light of the firm-specific comparative advantage that employees would lose by leaving (Groysberg et al. 2008). Second, there are open questions about the role of performance pay in limiting employee effort responses. The extant literature on responses to compensation changes has largely focused on fixed or hourly wages (Fehr and Falk 1999, Dickson and Fongoni 2019) and generally finds that workers reduce output after compensation reductions. By contrast, workers in our setting incurred an adverse change to their performance pay, which potentially explains the limited effort response in our setting. We conjecture that the lack of observed effort changes stems from a combination of income effects, income targeting, or reference points (Mas 2006). Although we are unable to precisely isolate these mechanisms with the variation available in our setting, better understanding these mechanisms is important because of the growing share of workers receiving performance pay (Lemieux et al. 2009).

Our work also addresses the microfoundations of downward nominal wage rigidity. By linking compensation changes with heterogeneous worker responses, our findings validate earlier surveys and interviews wherein managers report that the fear of losing top talent constrains them from adjusting compensation downward (Campbell and Kamlani 1997, Kahn 1997, Bewley 1998).

Prior research has shown that contextual framing and communication matters for how individuals respond to changes in their environment (Kahneman et al. 1986, Chen and Horton 2016, Englmaier et al. 2017). Future studies might consider how workers' responses to externally motivated events might vary compared with the within-firm motivations examined here (e.g., pandemics or business cycle shocks, as examined in Mascarenhas and Aaker 1989). A second area of future study is to apply our results more generally to the study of monopsony in the labor market. Existing studies tend to take the existence of a limited labor supply elasticity facing an individual firm as evidence that the firm can exercise labor market power. Given our results on worker heterogeneity, more work is needed to understand the conditions under which firms may exercise labor market power, or the contracts they would need to write with different workers in order to do so.

Our findings have two direct implications for managers. First, high performing employees are the most sensitive subgroup to adverse compensation changes. Therefore, although compensation changes may reduce payroll costs across all impacted employees, retention risks—and the subsequent costs—are greatest among top performers. This finding suggests insulating the most productive employees from adverse pay changes may be beneficial, albeit more research is required to understand the potential adverse effects of workplace inequality. Second, the presence of performance pay may limit negative, on-the-job responses to adverse compensation changes via income effects, suggesting a potential pathway to avoid previously observed forms of behavioral and sentiment-driven reactions to adverse compensation changes.

Acknowledgments
The authors thank Jen Brown, Lauren Cohen, Alain Cohn, Jeff Coles, Guido Friebel, Peter Kuhn, Bentley Macleod, Ramana Nanda, Paige Ouimet, Luke Stein, Ed Van Wesep, and seminar participants at the Arizona State Meeting of the Labor and Finance Group, Harvard Business School, 2018 Society for Institutional & Organizational Economics (SIOE) Conference, and University of California Santa Barbara for helpful feedback.

Endnotes

[1] Upstream service providers pay the firm for every sale in accordance with set contracts, which leads to the top-line revenue generated for the firm. All use of the term *revenue* in this paper refers to the transfer prices the firm uses to incentivize agents.

[2] Every agent has a fixed number of calls audited each week. If any conduct violations are identified, the agent's weekly commission rate is reduced. Selling efficiency is based on revenue-per-call (RPC) and revenue-per-hour (RPH). Being in higher quintiles of RPC and RPH increases an agent's commission rate.

[3] Although schedule adherence is tracked by the firm, agents are only penalized if their adherence level dips below a threshold of 80%, and the average pre-treatment adherence level among Division 1 agents was 83%.

[4] We cannot separate changes in sales from changes in effort in Division 2 because we lack product-level data with revenue transfer prices before and after treatment.

[5] Our results are little changed when using lower order polynomials, as shown in Table A.2.

[6] There is now a significant amount of literature that addresses these issues, and applied papers have generally used some version of the wild cluster bootstrap to get valid confidence regions. See, for example, Lazear et al. (2016). This estimator has been shown to perform well in simulations and avoids problems of over-rejection that are often endemic when there are few clusters.

[7] Heterogeneous responses can be estimated using division-by-time fixed effects without appealing to common trends across divisions. The maintained assumption here is common trends between different groups within each division. Figure A.1, (a) and (b), plots the evolution of within-division differences in performance by worker pre-treatment productivity, suggesting the validity of common trends within division.

[8] The details of this reweighting procedure are provided in Online Appendix C.1. Figure A.3, (a) and (b), displays the weighted and unweighted measures of log commissions-per-call and log commissions, respectively.

[9] Our estimation of changes in worker effort are conditional on the worker remaining at the firm. As turnover takes time to happen, however, we observe almost all treated workers with at least some sales data in the post-treatment period. The inclusion of agent fixed effects also partially addresses the concern that attrition could affect our measures of employees' effort responses. Importantly, the

turnover responses that we discussed in Section 5.1 emerge several weeks (at least six) after the commission schedule changes occurred. We would expect effort responses to manifest much earlier, so it is unlikely that our estimates of effort responses are driven by abnormal attrition.

[10] Our results are qualitatively unchanged when we measure productivity as revenue-per-hour rather than revenue-per-call. The call-based metric is the firm's focal measure and is more salient to sales agents and their direct supervisors, which is why we focus on it. The time-based metric provides an interesting complement to this measure. Panels A and B of Table A.4 use the logarithm of revenue-per-hour as the dependent variable and report similar estimates as those in Table 3, suggesting a limited time-spent-per-call response to the compensation changes. This evidence aligns with the fact that agents have a limited capacity to control the total number of calls received each week. Panels C and D of Table A.4 use level RPC as the dependent variable.

[11] Similarly, we may expect that productivity would immediately decline as people learn to allocate effort and game new incentive plans (Obloj and Sengul 2012). In our setting, however, it is difficult to disentangle learning and other time trends because the compensation changes impacted all eligible agents simultaneously.

[12] Therefore, the commission adjustments likely did not alter the firm's per-call unit economics, because of a substantial change in the composition of products sold. Additionally, these results are not driven by spillovers or reactions by agents in the control division, as discussed in the Online Appendix C.2.

[13] The heterogeneous treatment effects are based on standardized measures, so the average worker will have an effect that is captured by Treated $\times$ Post because the productivity average is zero. The interpretation for other workers requires multiplying by their productivity, which has a mean of zero and standard deviation 1, so the coefficients on these interactions reflect the effect of a standard deviation change around the mean.

[14] The relative reduction in the conversion of high performers is unlikely to be driven by mean reversion. Adjusted worker fixed effects—used to distinguish between high and low performers—are established using pre-treatment data up to four weeks before the changes occurred. The average conversion of agents in each of the three terciles of adjusted worker fixed effects increased from the weeks before this cutoff to the weeks after, suggesting mean reversion is not a likely cause of our findings. For example, top tercile agents in Division 1 had average conversion in September 2016 of 35%. In October, after the adjusted worker fixed effects had already been measured, this average increased to 37%. Across this same time horizon, bottom tercile agents maintained an average conversion of 29% and middle tercile agents increased their conversion from 32% to 33%. We cannot, however, disentangle whether changes in RPC and conversion rates are because of decreased effort or because of highly productive agents losing sales as the result of aggressively trying to upsell.

References

Aldatmaz S, Ouimet P, Van Wesep E (2018) The option to quit: The effect of employee stock options on turnover *J. Financial Econom.* 127(1):136–151.

Ashenfelter O, Heckman J (1974) The estimation of income and substitution effects in a model of family labor supply. *Econometrica* 42(1):73–85.

Bartel AP, Borjas GJ (1981) *Wage Growth and Job Turnover: An Empirical Analysis. Studies in Labor Markets* (University of Chicago Press, Chicago).

Bartel A, Ichniowski C, Shaw K (2004) Using "insider econometrics" to study productivity. *Amer. Econom. Rev.* 94(2):217–223.

Bewley TF (1998) Why not cut pay? *Eur. Econom. Rev.* 42(3-5):459–490.

Burdett K, Mortensen D (1998) Wage differentials, employer size, and unemployment. *Internat. Econom. Rev.* 39(2):257–273.

Campbell CM III, Kamlani KS (1997) The reasons for wage rigidity: Evidence from a survey of firms. *Quart. J. Econom.* 112(3): 759–789.

Campbell BA, Ganco M, Franco AM, Agarwal R (2012) Who leaves, where to, and why worry? Employee mobility, entrepreneurship and effects on source firm performance. *Strategic Management J.* 33(1):65–87.

Carnahan S, Agarwal R, Campbell BA (2012) Heterogeneity in turnover: The effect of relative compensation dispersion of firms on the mobility and entrepreneurship of extreme performers. *Strategic Management J.* 33(12):1411–1430.

Cengiz D, Dube A, Lindner A, Zipperer B (2019) The effect of minimum wages on low-wage jobs. *Quart. J. Econom.* 134(3): 1405–1454.

Chen DL, Horton JJ (2016) Research note–are online labor markets spot markets for tasks? A field experiment on the behavioral response to wage cuts. *Inform. Systems Res.* 27(2):403–423.

Cohn A, Fehr E, Goette L (2014a) Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Sci.* 61(8):1777–1794.

Cohn A, Fehr E, Herrmann B, Schneider F (2014b) Social comparison and effort provision: Evidence from a field experiment. *J. Eur. Econom. Assoc.* 12(4):877–898.

DellaVigna S, List JA, Malmendier U, Rao G (2016) *Estimating social preferences and gift exchange at work.* Technical report, National Bureau of Economic Research, Cambridge, MA.

Dickson A, Fongoni M (2019) Asymmetric reference-dependent reciprocity, downward wage rigidity, and the employment contract *J. Econom. Behav. Organ.* 163:409–429.

Dube A, Giuliano L, Leonard J (2018) *Fairness and frictions: The impact of unequal raises on quit behavior.* Technical report, National Bureau of Economic Research, Cambridge, MA.

Englmaier F, Roider A, Sunde U (2017) The role of communication of performance schemes: Evidence from a field experiment. *Management Sci.* 63(12):4061–4080.

Esteves-Sorenson C (2018) Gift exchange in the workplace: Addressing the conflicting evidence with a careful test. *Management Sci.* 64(9):4365–4388

Fehr E, Falk A (1999) Wage rigidity in a competitive incomplete contract market. *J. Political Econom.* 107(1):106–134.

Fehr E, Goette L, Zehnder C (2009) A behavioral account of the labor market: The role of fairness concerns. *Annual Rev. Econom.* 1(1):355–384.

Fowlie M, Greenstone M, Wolfram C (2018) Do energy efficiency investments deliver? Evidence from the weatherization assistance program. *Quart. J. Econom.* 133(3):1597–1644.

Giacalone RA, Greenberg J (1997) *Antisocial Behavior in Organizations* (Sage, New York).

Greenberg J (1990) Employee theft as a reaction to underpayment inequity: The hidden cost of pay cuts. *J. Appl. Psychol.* 75(5):561.

Groysberg B, Lee LE, Nanda A (2008) Can they take it with them? The portability of star knowledge workers' performance. *Management Sci.* 54(7):1213–1230.

Gubler T, Larkin I, Pierce L (2018) Doing well by making well: The impact of corporate wellness programs on employee productivity. *Management Sci.* 64(11):4967–4987.

Hölmstrom B (1979) Moral hazard and observability. *Bell J. Econom.* 10(1):74–91.

Jensen MC, Meckling WH (1976) Theory of the firm: Managerial behavior, agency costs and ownership structure. *J. Financial Econom.* 3(4):305–360.

Kahn S (1997) Evidence of nominal wage stickiness from microdata. *Amer. Econom. Rev.* 87(5):993–1008.

Kahneman D, Knetsch JL, Thaler R (1986) Fairness as a constraint on profit seeking: Entitlements in the market. *Amer. Econom. Rev.* 76(4):728–741.

Krueger M, Friebel G (2018) A pay change and its long-term consequences. Technical report, IZA Institute of Labor Economics, Bonn.

Krueger AB, Mas A (2004) Strikes, scabs, and tread separations: Labor strife and the production of defective bridgestone/firestone tires. *J. Political Econom.* 112(2):253–289.

Kube S, Maréchal MA, Puppe C (2013) Do wage cuts damage work morale? Evidence from a natural field experiment. *J. Eur. Econom. Assoc.* 11(4):853–870.

Larkin I, Leider S (2012) Incentive schemes, sorting, and behavioral biases of employees: Experimental evidence. *Amer. Econom. J. Microeconom.* 4(2):184–214.

Larkin I, Pierce L, Gino F (2012) The psychological costs of pay-for-performance: Implications for the strategic compensation of employees. *Strategic Management J.* 33(10):1194–1214.

Lazear EP (2000) Performance pay and productivity. *Amer. Econom. Rev.* 90(5):1346–1361.

Lazear EP, Shaw KL, Stanton CT (2015) The value of bosses. *J. Labor Econom.* 33(4):823–861.

Lazear EP, Shaw KL, Stanton C (2016) Making do with less: Working harder during recessions. *J. Labor Econom.* 34(S1):S333–S360.

Lemieux T, MacLeod WB, Parent D (2009) Performance pay and wage inequality. *Quart. J. Econom.* 124(1):1–49.

MacKinnon JG, Webb MD (2018) The wild bootstrap for few (treated) clusters. *Econometrics J.* 21(2):114–135.

Manning A (2003) *Monopsony in Motion: Imperfect Competition in Labor Markets* (Princeton University Press, Princeton, NJ).

Mas A (2006) Pay, reference points, and police performance. *Quart. J. Econom.* 121(3):783–821.

Mas A, Moretti E (2009) Peers at work. *Amer. Econom. Rev.* 99(1):112–145.

Mascarenhas B, Aaker DA (1989) Strategy over the business cycle. *Strategic Management J.* 10(3):199–210.

Obloj T, Sengul M (2012) Incentive life-cycles: Learning and the division of value in firms. *Admin. Sci. Quart.* 57(2):305–347.

Obloj T, Zenger T (2017) Organization design, proximity, and productivity responses to upward social comparison. *Organ. Sci.* 28(1):1–18.

Oyer P (2004) Why do firms use incentives that have no incentive effects? *J. Finance* 59(4):1619–1650.

Oyer P, Schaefer S (2005) Why do some firms give stock options to all employees?: An empirical examination of alternative theories. *J. Financial Econom.* 76(1):99–133.

Sandvik JJ, Saouma RE, Seegert NT, Stanton CT (2020) Workplace knowledge flows. *Quart. J. Econom.* 135(3):1635–1680.

Stafford TM (2015) What do fishermen tell us that taxi drivers do not? An empirical investigation of labor supply. *J. Labor Econom.* 33(3):683–710.

Zenger TR (1992) Why do employers only reward extreme performance? Examining the relationships among performance, pay, and turnover. *Admin. Sci. Quart.* 37(2):198–219.