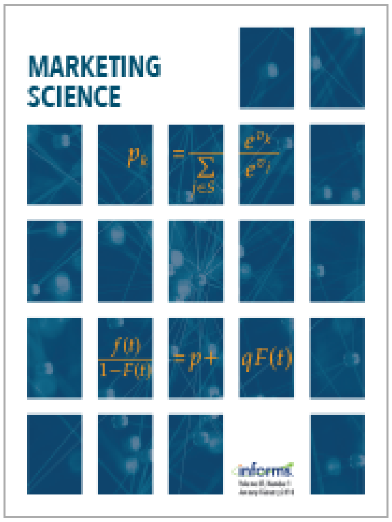



Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# Examining Salesperson Effort Allocation in Teams: A Randomized Field Experiment

Jia Li, Noah Lim, Hua Chen

To cite this article:
Jia Li, Noah Lim, Hua Chen (2020) Examining Salesperson Effort Allocation in Teams: A Randomized Field Experiment. Marketing Science 39(6):1122-1141. https://doi.org/10.1287/mksc.2019.1163

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

MARKETING SCIENCE
Vol. 39, No. 6, November–December 2020, pp. 1122–1141
ISSN 0732-2399 (print), ISSN 1526-548X (online)

# Examining Salesperson Effort Allocation in Teams: A Randomized Field Experiment

**Jia Li,[a] Noah Lim,[b] Hua Chen[c]**

[a] School of Business, Wake Forest University, Winston-Salem, North Carolina 27106; [b] NUS Business School, National University of Singapore, Singapore 119245; [c] Terry College of Business, University of Georgia, Athens, Georgia 30602

**Contact:** lijia@wfu.edu (JL); noahlim@nus.edu.sg, http://orcid.org/0000-0002-3801-8084 (NL); huachen@uga.edu (HC)

**Received:** June 28, 2017
**Revised:** September 12, 2018; January 11, 2019
**Accepted:** January 15, 2019
**Published Online in Articles in Advance:** September 9, 2019

https://doi.org/10.1287/mksc.2019.1163



**Abstract.** When salespeople with heterogeneous sales abilities are assigned into teams, how do they adjust effort as the abilities of their coworkers change? We investigate this question using a field experiment that spans 29 retail booths and 116 salespeople at a major department store in China. Each booth compensates salespeople using either an individual-based commission (IB) or a revenue-sharing (RS) incentive and employs four salespeople, with two salespeople per shift. Our field experiment randomly assigns salespeople to work shifts, thus exogenously varying the ability of a salesperson's coworker. The results show that under the IB incentive, the lower-ability salesperson will strategically decrease effort as the ability of the coworker rises; correspondingly, the higher-ability salesperson reduces effort as the coworker's ability decreases. In contrast, under the RS incentive, the sales pattern suggests that the lower-ability (higher-ability) salesperson increases effort when the coworker's ability increases (decreases). These empirical results provide broad support to our theory that accounts for the effect of social preferences on salespeople's effort decisions. We also examine the revenue implications of team composition on firm performance under different sales incentives.

**History:** This paper has been accepted for the *Marketing Science* Special Issue on Field Experiments.
**Supplemental Material:** Data are available at https://doi.org/10.1287/mksc.2019.1163.

**Keywords:** field experiments • sales force compensation • team composition • peer influence • social preferences • behavioral economics

## 1. Introduction

Across both business-to-business and business-to-consumer organizations, salespeople often work in teams; within the team, they are often heterogeneous in abilities, motivation, and sales effectiveness. While recruiting salespeople with high ability is critical, an equally important challenge for managers is assigning salespeople who have asymmetric abilities to teams. For example, suppose that a brand employs four salespeople who have different sales abilities. The manager must assign the salespeople equally to two teams, across two shifts in a day. Should the manager assign salespeople so that those with similar abilities work together? Or should the manager mix salespeople with more diverse abilities? The answer to this question depends on whether and how salespeople adjust their effort when the ability level of their coworker changes. Moreover, the pattern of effort adjustment may depend on the type of sales incentive used to motivate the salespeople.

The goal of this article is to examine how salespeople adjust effort in response to changes in the ability levels of their coworkers in a sales team. Specifically, we focus on sales teams in a retail context wherein salespeople interact socially and frequently with one another, share work-related information, and are able to observe coworkers' behavior and sales revenues; meanwhile, there is limited interdependence in their selling tasks so that each salesperson's sales revenue can be readily attributed to her own effort. In this paper, we examine behavior across two common types of incentive contracts: the individual-based (IB) incentive, in which salespeople earn commissions based on their individual sales; and the revenue-sharing (RS) incentive, in which salespeople equally divide commissions on the basis of the sales of the entire team. Drawing from the literature on how social preferences may influence salespeople, we propose a theory that predicts that both higher- and lower-ability salespeople will decrease (increase) their effort under the IB (RS) incentive when the ability gap between coworkers enlarges. These predictions also imply that a more homogeneous (heterogeneous) team composition will improve firm performance under the IB (RS) incentive.

We examine these predictions using a large-scale field experiment conducted in a major Chinese department store. Our study involved 29 brands in the women's clothing section of this department store. Each brand has a separate retail booth on the main floor and employs its own sales team to promote and sell its products, while paying the department store a share of its revenues. Typically, four salespeople with different sales abilities work in a retail booth and are assigned to work in one of two shifts in each day. In every shift, two salespeople work together in the same booth. Sixteen of the brands use the IB incentive to compensate salespeople, and the remaining 13 brands use the RS incentive.

During the experimental period, we were able to fully control the staffing assignment of the salespeople across all 29 booths. In Stage 1 of our study, we randomly assigned salespeople in each booth to work shifts, thus exogenously varying the ability level of the coworker that each salesperson was paired with in each day. The experimental results show that under the IB incentive, where salespeople are commissioned on individual sales revenue, sales of the lower-ability salespeople decreased as the ability levels of their coworkers increased. Correspondingly, sales of the higher-ability salespeople declined when they were matched with salespeople of increasingly lower ability. In contrast, for the booths that employed the RS incentive, where salespeople are compensated by group sales revenue, sales of the lower-ability salespeople increased as their coworker's ability increased. Sales of the higher-ability salespeople also increased as their coworker's ability decreased. Hence, these empirical patterns are consistent with the theory that salespeople's effort decisions are driven by both pecuniary considerations and social preferences.[1] We also carefully examine and rule out other potential theories such as learning, having a daily target income, and effort complementarities in teams.

While the randomization of salespeople into teams in Stage 1 of our study effectively eliminates endogeneity concerns and enables us to investigate the effect of staffing composition on salespeople's effort decisions, we also directly examine the impact of team composition on firm performance. To this end, in Stage 2 of our study, we randomly selected 10 booths, 5 from each of the IB and RS booths, and fixed the team pairings throughout the entire stage. Among these 10 booths, we further randomly selected 6 booths, 3 from each of the IB and RS booths, and assigned the sales teams so that the coworkers were heterogeneous in abilities (Heterogeneous treatment). For the remaining 4 booths, the staffing assignment was such that salespeople of similar sales abilities worked in the same team (Homogeneous treatment). The staffing assignment of the remaining 19 booths in Stage 2 was identical to that in Stage 1, wherein we randomly assigned salespeople in each booth to teams every day, which allowed us to confirm whether the findings from Stage 1 remain robust.

The Stage 2 results for the 19 booths corroborate the findings from the Stage 1 data. More importantly, the results indicate that under the IB incentive, assigning salespeople to homogeneous teams (relative to pairing them randomly) improves revenues by 3.7%, while moving to a team with a heterogeneous ability composition decreases revenues by 7.2%. Conversely, under the RS incentive, forming teams with a heterogeneous ability mix increases sales by 4.4%, while switching to a team in which members have homogeneous abilities leads to a revenue drop of 8.9%. Overall, our study provides valuable insights into how the composition of a retail sales team can affect the behavior of salespeople and how managers can optimize the ability mix of a sales team under different incentive contracts.

## 2. Review of Related Literature

This article is related to several streams of research. First, our study adds to the very thin literature that employs field experiments to examine salesperson behavior. To the best of our knowledge, there are only a few such studies in marketing: Lim et al. (2009) investigate the question of optimal contest design by randomizing salespeople into contests with different prize structures. Kishore et al. (2013) use a before–after quasi-experimental design to investigate how salespeople perform when their compensation switches from a lump-sum bonus to an equivalent linear commission rate. Chung and Narayandas (2017) conduct a field experiment to examine how various bonus compensation contracts affect salespeople's performance. They find that while a bonus tied to a sales target has similar impact across salespeople, a bonus not tied to a sales target is more effective on salespeople who have a high base performance. Note that these papers focus on how different types of compensation contracts affect salespeople's performance, but not in a context where salespeople perform their selling tasks alongside coworkers. In contrast, this article keeps the compensation contracts fixed and examines how salespeople strategically adjust their effort decisions when the sales ability of their coworker in a sales team changes.

Second, this article is related to the literature on team-based incentives. Theoretical work in this area has focused on the conditions under which managers should prefer group incentives to individual-based incentives. Specifically, group-based commissions can be optimal when significant task complementarities exist among group members with different expertise (Lazear 1998), when demand shocks across sales territories are negatively correlated so that there are risk-pooling benefits (Caldieraro and Coughlan 2009), and when team members have strong social preferences (Lim and Chen 2014). Empirically, Knez and Simester (2001) demonstrate that a group-based incentive can be efficacious even in a setting in which the number of team members is extremely large (such that it is almost impossible to coordinate effort levels). Hamilton et al. (2003) show that the productivity of garment workers in a U.S. factory increased as production was shifted from an individual-based production system to a team module production method, accompanied by a compensation change from an individual-based to an RS incentive. Hossain and List (2012) conduct field experiments with production workers working in the same factory and show that team-based target bonuses (particularly bonuses that were framed as a deduction or loss, rather than as a gain) can raise productivity more than individual-based target bonuses. Using two laboratory experiments, Chen and Lim (2017) show how heterogeneity in sales ability affects effort in a team-based contest depends on the type of contest metric used to measure team performance—heterogeneity has no effect on team effort under the average metric but reduces team effort under the minimum and maximum metrics. Note that unlike in this paper, their experiments do not manipulate the degree of heterogeneity in sales ability within the team. Overall, we are not aware of any field study that has examined how salespeople working in a team might adjust their behavior when the abilities of their coworkers vary, when salespeople are compensated by an RS incentive.

Third, this article is related to the literature on productivity spillovers, or peer effects, in the workplace. Mas and Moretti (2009) show that supermarket clerks who had high productivity raised the productivity of other clerks working independently at the same time. In their setting, the supermarket clerks were paid fixed hourly wages and there is no incentive pay. Chan et al. (2014) extend this literature by identifying peer effects among salespeople working for cosmetic counters under the IB and RS incentives. Although our work has some similarities to theirs in terms of the sales setting and compensation contracts, it differs both substantively and methodologically. The primary focus of the study by Chan et al. (2014) is on measuring peer effects at the aggregate level (i.e., how a salesperson's contemporaneous set of peers influences his or her temporal productivity, both at the cosmetic counter and across competing counters). In their article, it remains an open question what exactly contributes to the identified peer effects; that is, they do not examine the precise mechanisms behind the productivity outcomes. In contrast, we are primarily interested in understanding how salespeople strategically allocate effort as the abilities of their coworker vary, and we show that changes in salesperson effort allocation can be an important driver of the peer effects that Chan et al. (2014) document. Furthermore, our identification strategy differs from theirs in that they conduct their analysis on archival data; they acknowledge that "it is also important to reiterate the potential endogeneity of the compensation systems and the worker types chosen by firms. . . . Outside of a true experiment, this problem is difficult to resolve" (p. 1981). We answer their call to further examine the drivers of peer effects by using a randomized field experiment. Finally, Stage 2 of our experimental design allows us to directly test the impact of team heterogeneity on team performance, which is unique to our article.

Fourth, this paper builds on recent marketing research on how social preferences affect salespeople's behavior. Lim (2010) and Chen and Lim (2013) show that nonpecuniary factors, such as social comparisons and guilt aversion, can affect salesperson effort in environments where salespeople interact with one another frequently. Chen et al. (2011) examine sales contests where contestants are heterogeneous in sales abilities and find that the higher-ability salespeople are averse to performing worse, while the lower-ability salespeople experience utility gains when they come out ahead in performance. The results of these three papers are based on laboratory experiments. We thus expand this literature by using a field experiment to provide evidence that social preferences can affect salesperson behavior.

The rest of this article proceeds as follows: Section 3 introduces the institutional setting and provides predictions of how salespeople of asymmetric abilities in a sales team adjust effort when their coworker's ability changes and how team composition affects the revenues of a firm, under the IB and RS incentives. Section 4 details the empirical context and the design of the field experiment. In Section 5, we present the empirical analysis of our data in Stage 1. We also examine and rule out other potential explanations for the empirical findings. We then empirically investigate the revenue implications of different team compositions for the IB and RS booths in Section 6, using the data from Stages 1 and 2. Section 7 concludes.

## 3. Expected Salesperson Behavior

### 3.1. Institutional Setting

The field experiment was conducted in a high-end department store located in a major city in China that is among the largest of its kind in the country. Our study involved 29 brands in the women's clothing section of this department store. Each brand has a separate retail booth in the department store. The department store can accurately monitor and track individual sales revenue generated by each salesperson on a daily basis. Sixteen of the 29 booths incentivize salespeople by the IB incentive, in which salespeople earn both fixed salary and commissions based on their individual sales revenues. In the remaining 13 booths, salespeople are compensated by the RS incentive, in which salespeople also receive fixed salary (similar to the IB incentive), and equally divide commissions based on the sales revenue of the entire team.

These 29 booths are located close to one another in the east hall of the department store, which is considered a premium area. Each brand employs its own salespeople to promote and sell its products so that each salesperson works for only one booth. All the salespeople employed by the 29 booths are women. Typically, four salespeople with different sales abilities work in a retail booth and they are assigned to work in one of two shifts in each day. In every shift, two salespeople work together in the same booth. The department store implements a rotated staffing assignment scheme where each salesperson is rotated across different shifts in the day, across different days of the week. Because of this rotation, a salesperson will be paired with one of the other three salespeople in the same booth (whose sales abilities may differ). We will describe the staffing assignment in detail later.

**Salesperson Selling Tasks.** In the context of selling premium women's clothing in a high-end department store, salespeople usually do not expend effort in asking customers to visit their booth; rather, customers choose whether to enter a booth based on factors such as the brand image, clothing design, sales promotion, and so forth.

The daily selling tasks a salesperson carries out can be categorized into two types. When a salesperson is not serving any customer, she will spend time maintaining a good shopping environment in the booth for future customers. The activities mainly include cleaning, restocking, folding clothes that customers had inspected, record keeping, and so forth. Occasionally, a salesperson might provide help to her coworker who is serving customers. We name these activities the *Team Tasks*. These team tasks are split almost equally between coworkers in the same work shift for equity reasons. Once customers enter the booth, salespeople usually take turns to serve customers. Typically, each customer is served by only one salesperson during the entire visit. The salesperson focuses on understanding customers' needs and preferences, showing customers the clothing items that may suit them, providing professional feedback and suggestions, answering customers' questions, and conducting pleasant verbal and nonverbal communications with customers, with the objectives of getting customers to make purchases, increasing the size of their shopping baskets, and earning repeat business in the future.[2] We denote these activities as the *Individual Selling Tasks*.

Note that these two types of tasks are identical across the IB and RS booths. In this paper, we focus on salespeople's individual selling tasks, because they have more direct impact on individual sales revenue, which is our primary outcome variable of interest.

### 3.2. Potential Drivers of Salesperson Behavior

Although sales agents are typically assumed to be self-interested, the behavioral economics literature (e.g., Fehr and Schmidt 1999, Fischbacher and Gächter 2010) has shown that social preferences—such as social comparisons, guilt aversion, and inequality aversion—can affect how people make decisions. Moreover, the strength of these social preferences can depend critically on the social environment in which people make economic choices (Chen et al. 2011).

In our empirical setting, the salespeople within the same booth have been working together for at least 6 months; thus, they are familiar with one another and possess good knowledge about their coworkers' sales abilities. Equally importantly, salespeople can accurately track the sales revenues of their coworkers during the same shift so that they can easily compare their own earnings and sales revenues with their coworker's. Given this social environment, when salespeople receive different amounts of commissions under the IB incentive, they may compare their earnings in a work shift against those of their coworker in the same shift. Particularly, salespeople may have social loss aversion; that is, they will feel worse off when they receive a lower monetary payoff than their coworker (Lim 2010, Chen et al. 2011).

On the other hand, given an RS incentive where salespeople's effort decisions can affect their coworkers' payoffs, they may also care about whether their effort lifts up or pulls down the earnings of their coworker during the work shift. Chen and Lim (2013) show that salespeople working under a team-based incentive have guilt aversion (Charness and Dufwenberg 2006); that is, salespeople suffer disutility if their teammates' payoffs are dragged down as a result of being assigned to work with them.

### 3.3. Theoretical Illustration of Salesperson Behavior

In this section, we use a simple two-person model to formally illustrate how social preferences can affect the way salespeople adjust effort when the ability of their coworker changes (in our experimental context, when their coworker changes due to the staffing assignment in the retail booths) under the IB and RS incentives, respectively. When a coworker's ability changes, the salesperson would adjust effort by considering her coworker's possible effort decision, and vice versa for her coworker. Thus, in our model the change in coworker's sales ability (and consequently, the ability gap between coworkers) is the main exogenous manipulation and salespeople's effort decisions are the outcomes of the strategic interaction between salespeople who work together.

Consider two salespeople who work in the same booth during a shift. The individual sales revenue of salesperson $j$ ($j$ = $H$, $L$) is $y_j = \alpha_j e_j + \varepsilon_j$, where $\alpha_j > 0$ captures the sales ability of the salesperson and $e_j$ is the effort level expended to serve the customer. Thus, the ability of a salesperson is defined as the effectiveness of her effort in generating sales. Since a salesperson's ability is unlikely to change within a short period of time, our theoretical analysis treats sales ability as an exogenous parameter and focuses on examining the change in salesperson effort due to a change in her coworker's ability.[3] The term $\varepsilon_j$ accounts for the demand shocks that are beyond the salesperson's control, and we assume that $\varepsilon_j \sim U[-V,V]$ (where $V > 0$) is i.i.d. across salespeople. Without loss of generality, we assume that salesperson $H$ has a higher ability than salesperson $L$; that is, $\alpha_H > \alpha_L > 0$. Expending sales effort is costly for salespeople, and the cost of effort function of salesperson $j$ is given by $c(e_j) = e_j^2$. Note that the individual sales output function is common to salespeople in both the IB and RS booths.

For the 16 brands that incentivize salespeople by the IB incentive, the compensation contract offered is a commission based on salespeople's own sales, which is $\pi_j = m * y_j$, where $m$ is the commission rate. For the 13 booths in which salespeople are compensated by the RS incentive, salespeople receive commissions based on the average sales of the team, i.e., $\pi_j = m * (y_i + y_{-j})/2$, where $-j$ denotes the other salesperson in the shift.

**Salesperson Utility.** Based on the discussion in Section 3.2, we propose that the utility of salesperson $j$ includes both pecuniary payoff and psychological components as follows:

$$U_j = \pi_j - \gamma_j * \mathrm{Prob}(\pi_j < \pi_{-j}) - \psi_j * \mathrm{Prob}(\pi_{-j} < \bar{\pi}_{-j}) - e_j^2; \tag{1}$$

$\pi_j$ and $e_j^2$ represent the pecuniary payment and cost of effort, respectively, as shown before. $\gamma_j > 0$ captures the degree of disutility salesperson $j$ suffers if she receives a lower payment than her coworker (i.e., salesperson $-j$) and $\mathrm{Prob}(\pi_j < \pi_{-j})$ is the probability that this disutility occurs. The second nonpecuniary component in Equation (1) indicates the utility loss (captured by $\psi_j > 0$) experienced by salesperson $j$ when her coworker's actual payoff $\pi_{-j}$ is lower than the potential commission coworker $-j$ could earn without working with salesperson $j$. This potential commission is represented by $\bar{\pi}_{-j} = m * y_{-j}$. Again, $\mathrm{Prob}(\pi_{-j} < \bar{\pi}_{-j})$ represents the probability that salesperson $j$ experiences this type of guilt.

**Effort Allocation Under the IB Incentive.** Under the IB incentive, salespeople are compensated purely on the basis of individual sales revenue, so $\pi_{-j} = \bar{\pi}_{-j} = m * y_{-j}$. As a result, $\mathrm{Prob}(\pi_{-j} < \bar{\pi}_{-j}) = 0$. Thus, the utility of salesperson $j$ under the IB incentive reduces to

$$U_j = \pi_j - \gamma_j * \mathrm{Prob}(\pi_j < \pi_{-j}) - e_j^2. \tag{2}$$

Taking the derivative of Equation (2) with respect to effort, the equilibrium effort of Salespeople $H$ and $L$ are $e_H^* = \alpha_H\{4V\gamma_H + m[8V^2 + \alpha_L^2(\gamma_H - \gamma_L)]\}/[2(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)]$ and $e_L^* = \alpha_L\{4V\gamma_L + m[8V^2 + \alpha_H^2(\gamma_H - \gamma_L)]\}/[2(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)]$. Recall that $\alpha_H$ and $\alpha_L$ represent the abilities of Salespeople $H$ and $L$, respectively, where $\alpha_H > \alpha_L$.

The main research question of this paper is how a salesperson adjusts effort when her coworker's ability varies (since she is paired with different coworkers across different shifts). To answer this, our key interest lies in examining the sign of $\partial e_j^*/\partial\alpha_{-j}$. Taking the derivative of equilibrium effort with respect to the coworker's ability, we obtain $\partial e_H^*/\partial\alpha_L = \alpha_H\alpha_L\gamma_H\{4V\gamma_L + m[8V^2 + \alpha_H^2(\gamma_H - \gamma_L)]\}/(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)^2$ and $\partial e_L^*/\partial\alpha_H = -\alpha_H\alpha_L\gamma_H\{4V\gamma_H + m[8V^2 + \alpha_L^2(\gamma_H - \gamma_L)]\}/(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)^2$. We can show that as long as $\gamma_H > max\{(\alpha_L^2\gamma_L - 8V^2)/\alpha_H^2,\ [(\alpha_H^2 m - 4V)\gamma_L - 8mV^2]/(\alpha_H^2 m),\ m(\alpha_L^2\gamma_L - 8V^2)/(\alpha_L^2 m + 4V)\}$,[4] we will have $\partial e_H^*/\partial\alpha_L > 0$ and $\partial e_L^*/\partial\alpha_H < 0$. Thus, we have the following prediction:

**Prediction 1.** *Under the IB incentive, when social loss aversion is sufficiently strong, the effort expended by both higher- and lower-ability salespeople decreases when the difference in ability between the coworkers increases.*

The intuition can be explained as follows. We start with the focal salesperson's effort adjustment due to the change in her coworker's ability, and then explain how this adjustment affects her coworker's corresponding effort decision. As the ability gap between the coworkers expands, increasing effort is not optimal for the lower-ability salesperson. This is because any effort increase by the lower-ability salesperson will trigger a substantial increase in effort by the higher-ability coworker, if the latter is sufficiently averse to earning less. This effort increase by the higher-ability salesperson, coupled with the increased sales ability gap, will further raise the probability that the lower-ability salesperson suffers psychological disutility from earning relatively less than her coworker. In other words, increasing effort would eventually worsen the status of the lower-ability salesperson in the social comparison against her higher-ability coworker. Hence, it would be better for the lower-ability salesperson to decrease effort and save on effort costs. Knowing this, the higher-ability salesperson would best respond by reducing effort correspondingly.[5]

**Effort Allocation Under the RS Incentive.** Under the RS incentive, because both salespeople receive identical amounts of commissions, i.e. $\pi_j = \pi_{-j}$, we have $\text{Prob}(\pi_j < \pi_{-j}) = 0$. Thus, the utility of salesperson $j$ under the RS incentive simplifies to

$$U_j = \pi_j - \psi_j * \text{Prob}(\pi_{-j} < \bar{\pi}_{-j}) - e_j^2. \qquad (3)$$

From Equation (3), the equilibrium effort of Salespeople $H$ and $L$ are $e_H^* = \alpha_H\{8V\psi_H + m[8V^2 - \alpha_L^2(\psi_L - \psi_H)]\}/[4(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L)]$ and $e_L^* = \alpha_L\{8V\psi_L + m[8V^2 - \alpha_H^2(\psi_L - \psi_H)]\}/[4(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L)]$, respectively. Taking the derivative of equilibrium effort with respect to the coworker's ability, we obtain $\partial e_H^*/\partial\alpha_L = \alpha_H\alpha_L\psi_H\{8V\psi_L + m[8V^2 - \alpha_H^2(\psi_L - \psi_H)]\}/[2(8V^2 + \alpha_H^2 \cdot \psi_H - \alpha_L^2\psi_L)^2]$ and $\partial e_L^*/\partial\alpha_H = -\alpha_H\alpha_L\psi_H\{8V\psi_H + m \cdot [8V^2 - \alpha_L^2(\psi_L - \psi_H)]\}/[2(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L)^2]$. It is straightforward to show that $\partial e_H^*/\partial\alpha_L < 0$ and $\partial e_L^*/\partial\alpha_H > 0$ when $\psi_L > max\{(\alpha_H^2\psi_H + 8V^2)/\alpha_L^2, [8\psi_H V + m(\alpha_L^2\psi_H + 8V^2)]/(\alpha_L^2 m), m(\alpha_H^2\psi_H + 8V^2)/(\alpha_H^2 m - 8V)\}$.[6]

**Prediction 2.** *Under the RS incentive, when guilt aversion is sufficiently strong, the effort expended by both higher- and lower-ability salespeople increases when the difference in ability between the coworkers increases.*

The intuition for Prediction 2 is as follows. Given the RS contract, when the guilt aversion of the lower-ability salesperson is sufficiently strong, she increases effort as the ability gap between the coworkers widens. There are two benefits from doing so. First, it makes her less likely to feel guilty for pulling down the earnings of her higher-ability coworker. Second, she would benefit economically from the response of her coworker, which would be to increase effort. Turning next to the higher-ability salesperson, as her lower-ability coworker increases effort, the probability that she would now suffer disutility from holding down her coworker's earnings increases, so her best response is to raise effort as well.

Overall, it is interesting to note that our theory of social preferences predicts that salespeople would respond differently to changes in their coworker's ability, across the IB and RS contracts.[7] Also note that if there are no social preferences (i.e., $\gamma_j = \psi_j = 0$), then salespeople would not change their effort when their coworkers' abilities vary, in both the IB and RS conditions. Finally, at the firm level, Predictions 1 and 2 imply the following:

**Prediction 3.** *Under the IB incentive, a sales team with a more homogeneous ability composition improves firm revenue. In contrast, under the RS incentive, the firm benefits more when a sales team has a more heterogeneous ability mix.*

## 4. Empirical Context and Experimental Design

To test the predictions above, we conducted a field experiment in which we manipulated the staffing composition of salespeople working in the women's clothing section of a leading high-end department store in China. We first detail the empirical context of our study and then explain the experimental design.

### 4.1. Details of Empirical Context

Figure 1 displays the floor plan and locations of the 29 booths in the east hall of the women's clothing section of the department store. This location within the department store is particularly ideal for our experiment for several reasons. First, the booths consist of different brands that employ both the individual- and group-based incentive contracts. The booths that employ the former compensate a salesperson using a fixed salary between U.S. \$140 and \$230 per month (\$160 on average), plus 2% of the sales revenue generated by that salesperson during that month. The booths using the RS incentive typically pay each salesperson a fixed salary around \$140 to \$200 (\$155 on average), plus 0.5% of the total monthly sales revenue generated by all salespeople working in that booth. A Wilcoxon rank-sum test shows that there is no difference between the fixed salaries across the IB and RS booths ($p = 0.40$). Note that while there is slight variation in fixed salary across the booths under the same incentive contract, it does not vary within a booth; that is, all salespeople working for the same booth receive an identical fixed salary so that differences in total compensation within a booth are driven by differences in incentive compensation. The left panel of Table 1 provides the average brand daily sales revenues for both

**Figure 1.** Floor Plan and Locations of the 29 Booths

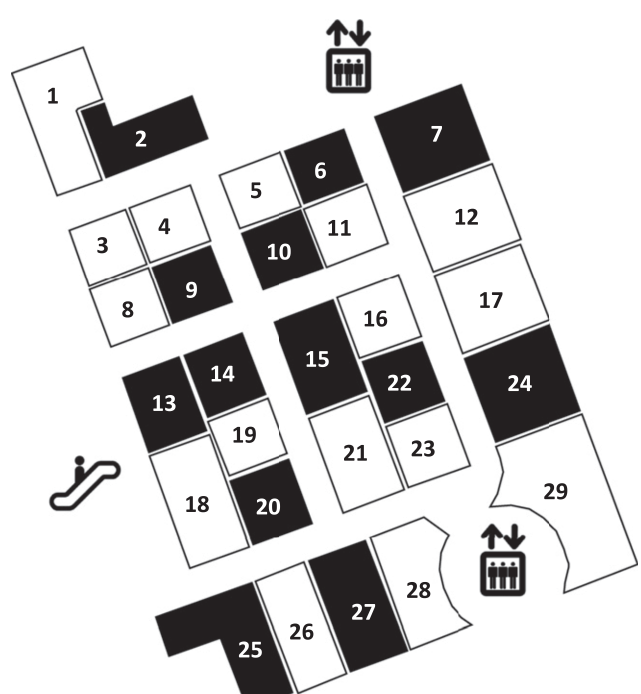


**Table 1.** Summary Statistics of the Booths and Salespeople in the Experiment

| | Booths | | | Salespeople | | | |
|---|---|---|---|---|---|---|---|
| | Number of booths | Average sales force size | Average booth daily sales | Average individual daily sales | Maximum individual daily sales | Minimum individual daily sales | C.V. of individual daily sales |
| All IB booths | 16 | 4.06 | 9,124.29 | 2,264.26 | | | 0.62 |
| IB booths: Stage 2 Random treatment | 11 | 4.09 | 9,231.22 | 2,257.02 | 7,139 | 0 | 0.64 |
| IB booths: Stage 2 Heterogeneous treatment | 3 | 4.00 | 9,226.13 | 2,306.53 | 6,208 | 0 | 0.57 |
| IB Booths: Stage 2 Homogeneous treatment | 2 | 4.00 | 8,916.93 | 2,229.23 | 5,718 | 0 | 0.58 |
| All RS booths | 13 | 3.92 | 9,171.72 | 2,316.01 | | | 0.61 |
| RS booths: Stage 2 Random treatment | 8 | 3.88 | 8,967.68 | 2,311.25 | 6,840 | 0 | 0.59 |
| RS booths: Stage 2 Heterogeneous treatment | 3 | 4.00 | 9,316.10 | 2,329.03 | 7,020 | 0 | 0.63 |
| RS booths: Stage 2 Homogeneous treatment | 2 | 4.00 | 9,231.04 | 2,307.76 | 6,609 | 0 | 0.61 |
| Overall | 29 | 4.00 | 9,144.39 | 2,286.10 | | | 0.62 |

*Notes.* Daily sales (at the booth or individual level) is calculated based on the sales of the 180 pre-experiment days. Daily sales are in Chinese Renminbi (¥). "C.V. of individual daily sales" is the coefficient of variation in sales. The sales force size is four for most of the booths (26 of 29). Two IB booths have five salespeople, and one RS booth has three salespeople. Statistics tests confirm that there is no difference between the IB and RS brands in terms of daily sales.

the IB and RS booths, for a period of 180 days preceding the experiment. Statistical tests show that there is no difference in the average daily sales revenues between the 16 booths that employ the IB incentive and the 13 booths that use the RS incentive ($p = 0.59$). In addition, according to the store management, customers perceive all 29 brands in the east hall as having similar brand prestige and price points.

Second, as the left panel of Table 1 shows, all 29 booths in the area are also almost identical in terms of the number of salespeople employed. Specifically, all but three booths hire four salespeople each. Moreover, the 29 booths are similar in the way they assign salespeople to work in the booths. That is, all of the booths with four salespeople employ an identical two-shift system for staffing salespeople: In a typical day, two salespeople work the first (morning) shift from 9 a.m. to 3 p.m. (6 hours), and the other two salespeople work the second (afternoon) 6-hour shift from 3 p.m. to 9 p.m.[8] In this way, this empirical setting naturally controls for the potential differences in sales revenues due to different team sizes and length of working hours. Furthermore, because there are only two salespeople working together at any time in a booth, we can cleanly attribute a salesperson's effort adjustment (if any) to the varying abilities of coworkers (we discuss this in more detail subsequently). The right panel of Table 1 provides summary statistics of the daily sales revenues of the salespeople in our sample, still for a period of 180 days preceding the experiment. Although there is substantial variation in sales revenue generated by the salespeople, which in turn suggests the presence of heterogeneity in sales abilities, there is no systematic difference in the variation in sales (as measured by the coefficient of variation) across the booths that employ the IB versus RS incentives ($p = 0.62$).

The empirical setting also provides additional advantages for our experimental study. First, the east hall is considered a premium area of the department store, located on one of the busiest floors of the store and anchored by the main elevators to that floor. Because the floor traffic is sufficiently busy, any reallocation of effort by salespeople can translate readily into significant differences in their sales revenues. Second, because all 29 booths are located close to one another in the same hall, if any exogenous shocks were to occur (e.g., changes in floor traffic), they would likely affect all the booths in a similar way. Third, there was no employee turnover or change in compensation contract for all 29 booths during the experimental period.

### 4.2. Experimental Design

Our experiment took place from July 23 to August 27 (36 consecutive days) in 2013. We chose this period because there were no major holidays or storewide events scheduled. Before the experiment began, for each of the 29 booths, we ranked all the salespeople according to their long-term sales productivities, which we defined as their average daily sales in the past 180 days. We used this ranking as a proxy for a salesperson's relative sales ability in the booth.[9] In each booth, we labeled the salesperson with the highest rank as "Salesperson 1," the one with the second-highest rank as "Salesperson 2," and so on.[10] For example, if the average daily sales of the four salespeople in a booth were ¥1,800, ¥3,000, ¥2,400, and ¥1,200, we labeled them as Salesperson 3, Salesperson 1, Salesperson 2, and Salesperson 4, respectively.[11]

Across the entire experiment, the only factor we manipulated was the staffing composition of the salespeople (i.e., which salespeople worked together in a given shift). Except for the experimental period, the staffing assignment was decided by the floor managers of the department store. Salespeople in each booth are typically rotated across days (i.e., day of week), shifts, and coworkers based on equity reasons (for example, they would naturally choose those shifts where customer traffic is heavier to maximize their incentive pay). During the experimental period, the floor managers were directed by the senior management to help us implement our prescribed staffing assignment.

Our experiment consists of two stages: Stage 1 took place from July 23 to August 6 (15 consecutive days), and Stage 2 took place from August 7 to August 27 (21 consecutive days). The key objective of Stage 1 was to directly test Predictions 1 and 2 on how salespeople adjust effort when their coworker's ability varies. In Stage 1, we manipulated the staffing assignment for each of the 29 booths in the following way: We first *randomly* divided the four salespeople working in the booth into two pairs and then *randomly* assigned the two pairs across the two shifts in a day.[12] As a result, each salesperson had the opportunity to be paired with each of her three coworkers (with varying sales abilities) during Stage 1. In other words, we operationalized the change in coworker's ability by replacing a salesperson's coworker with a new one. Thus, our study exogenously varies the sales ability of the coworker with whom each salesperson is assigned to work using a full randomization design, which allows us to draw causal inferences about how salespeople might adjust their efforts in response to the change in their coworker's ability.

Note that this staffing assignment manipulation leads to a total of three possible *staffing conditions* for a booth in any single day. Each of the three conditions consists of two *staffing compositions*, one for the morning shift and one for the afternoon shift. We label these three conditions as "Condition A," "Condition B," and "Condition C," respectively, and describe them in Table 2. For example, Condition A describes the staffing condition in which Salesperson 1 is paired with Salesperson 4 in one shift (either the morning or afternoon shift) and Salesperson 2 is paired with Salesperson 3 in the other shift. In this way, our Stage 1 experiment can also be considered a random assignment of a booth to one of the three conditions on each day. Note also that in each condition, the shift in which a staffing composition is allocated is randomly determined. Taking Condition A as an example, Salespeople 1 and 4 work the morning shift on some days and the afternoon shift on other days, whenever Condition A is selected. Overall, the randomized staffing assignment in Stage 1 leads to an almost equal frequency of the three conditions and, within each condition, an almost equal frequency of the two compositions across the two shifts in a day.[13]

The salespeople were not informed of the staffing manipulation in Stage 1. Moreover, even though there was a change from a nonrandom to a randomized staffing assignment, it is unlikely that salespeople would have perceived a significant change in the way their work shifts were determined. This is because in Stage 1, they continued to be assigned to different shifts within a day and different coworkers in their booth, similar to their typical assignments.

Stage 2 of our experiment began immediately after Stage 1. We designed Stage 2 to achieve two goals simultaneously: (1) to provide additional evidence to validate the empirical observations in Stage 1 and (2) to directly test Prediction 3 on the impact of staffing composition on firm-level performance. To accomplish these goals, in Stage 2, we randomly assigned each of the 29 booths to one of three treatments, which we label the "Heterogeneous," "Homogeneous," and "Random" treatments. Specifically, we first randomly selected three of the 16 IB booths and three of the 13 RS booths and assigned them to the Heterogeneous treatment. Similarly, for the Homogeneous treatment, we randomly selected four booths, two each from the IB and RS booths. The remaining 19 booths, consisting of 11 IB booths and eight RS booths, formed the Random treatment.[14]

Note that for the booths in both the Heterogeneous and Homogeneous treatments, we implemented a *fixed* staffing assignment throughout the 21 days in Stage 2. To begin, we classified the four salespeople employed by a booth as "high ability" and "low ability" using a median split of the salespeople's long-term productivities, as previously defined. Thus, in each booth, Salespeople 1 and 2 are classified as high-ability salespeople while Salespeople 3 and 4 as low-ability salespeople.

**Table 2.** Conditions and Corresponding Staffing Compositions in Stage 1

| Condition | Staffing composition 1 | Staffing composition 2 |
|---|---|---|
| A | Salesperson 1 and Salesperson 4 | Salesperson 2 and Salesperson 3 |
| B | Salesperson 1 and Salesperson 3 | Salesperson 2 and Salesperson 4 |
| C | Salesperson 1 and Salesperson 2 | Salesperson 3 and Salesperson 4 |

*Note.* During the experiment, the two compositions in each of the three conditions are equally assigned across the morning and afternoon shifts.

In the Heterogeneous treatment, the high-ability salesperson in a booth was assigned to work with the *same* low-ability salesperson every day during Stage 2. Note, however, that which high-ability salesperson and low-ability salesperson were paired was *randomly* decided.[15] In the Homogeneous treatment, the high-ability salesperson was paired with the other high-ability salesperson (i.e., Salespeople 1 and 2), and the low-ability salesperson was paired with the other low-ability salesperson (i.e., Salespeople 3 and 4), every day throughout Stage 2.

In both the Heterogeneous and Homogeneous treatments, the fixed pairs of salespeople rotated shifts every day. For example, in the Heterogeneous treatment, if Salespeople 1 and 3 worked together in the morning shift on the first day of Stage 2, they were assigned to the afternoon shift on the second day, and so on. Finally, in the Random treatment, the staffing assignment in each booth followed the same staffing scheme as in Stage 1, in which salespeople were randomly paired to work with each other. Thus, the 19 booths in the Random treatment allow us to continue testing the predictions of salespeople's individual behavior in the same way as in Stage 1. Meanwhile, the comparison between the Heterogeneous and Homogeneous treatments in relation to the Random treatment in Stage 2 allows us to empirically investigate how team ability composition affects firm performance.

Next, because we were only able to assign a limited number of booths to the Heterogeneous and Homogeneous treatments in Stage 2, it is important to ensure that there are no systematic differences in sales revenues between these booths and the booths in the Random treatment. The left panel of Table 1 displays the average daily booth revenues, for the period of 180 days preceding the experiment, across the three treatment conditions in Stage 2. Our statistical tests indicate that there are no significant differences in sales across the booths in the three treatments ($p = 0.21$).

Finally, we explain how the experimental designs in Stages 1 and 2 relate to each other. Recall that Stage 1 has three possible staffing conditions (see Table 2). Conditions A and B in Stage 1 correspond to the Heterogeneous treatment in Stage 2, and Condition C corresponds to the Homogeneous treatment. Next, all 29 booths in Stage 1 are assigned to the Stage 2 Random treatment. Table 3 summarizes the staffing assignment across both stages of our experiment.

## 5. Empirical Analysis on Stage 1 Data

Across the two experimental stages, we manipulated the staffing composition of the salespeople in a booth, across all the 29 booths, for 36 days and two shifts in a day. The department store also provided us with the daily sales revenue of each salesperson. We conduct our empirical analysis based on these data. In this section, we present our empirical analysis on the Stage 1 data to examine Predictions 1 and 2.

### 5.1. Initial Model-Free Evidence

Before conducting the formal empirical analysis, we graphically describe the patterns in the raw data. As mentioned previously, we ranked the salespeople in each booth by their sales ability and labeled them as Salespersons 1, 2, 3, and 4, respectively. Figure 2 displays the average daily sales revenue of each of these four salespeople, depending on which of the other coworkers the salesperson is paired with. The average

**Table 3.** Booth Staffing Assignment and Composition Details

| IB booths | | | RS booths | | |
|---|---|---|---|---|---|
| Booth ID | Stage 1 composition | Stage 2 composition | Booth ID | Stage 1 composition | Stage 2 composition |
| 1 | Random | Heterogeneous (1-4/2-3) | 2 | Random | Random |
| 3 | Random | Random | 6 | Random | Random |
| 4 | Random | Random | 7 | Random | Heterogeneous (1-4/2-3) |
| 5 | Random | Random | 9 | Random | Random |
| 8 | Random | Homogeneous (1-2/3-4) | 10 | Random | Homogeneous (1-2/3-4) |
| 11 | Random | Random | 13 | Random | Random |
| 12 | Random | Heterogeneous (1-3/2-4) | 14 | Random | Heterogeneous (1-4/2-3) |
| 16 | Random | Homogeneous (1-2/3-4) | 15 | Random | Random |
| 17 | Random | Random | 20 | Random | Random |
| 18 | Random | Random | 22 | Random | Random |
| 19 | Random | Heterogeneous (1-3/2-4) | 24 | Random | Heterogeneous (1-3/2-4) |
| 21 | Random | Random | 25 | Random | Random |
| 23 | Random | Random | 27 | Random | Homogeneous (1-2/3-4) |
| 26 | Random | Random | | | |
| 28 | Random | Random | | | |
| 29 | Random | Random | | | |

*Notes.* "1-4/2-3" stands for Salesperson 1 working with Salesperson 4 in one shift and Salesperson 2 working with Salesperson 3 in the other shift. Both "1-4/2-3" and "1-3/2-4" belong to the Heterogeneous treatment in Stage 2 of the experiment, and "1-2/3-4" describes the Homogeneous treatment.

sales revenue is based on the entire experimental period and includes the data from all booths employing the same incentive contract.

As the salesperson with the highest ability in a booth, Salesperson 1 always works with a lower-ability coworker, which allows us to discern changes in her sales because the ability of her coworker is progressively lower. Panel I of Figure 2 illustrates that in the IB booth, Salesperson 1's sales are lower when she is paired with Salesperson 3 than when she is paired with Salesperson 2. Salesperson 1's sales decrease further when she is paired with Salesperson 4, the salesperson with the lowest sales ability in the booth. This pattern suggests that under the IB incentive, the higher-ability salesperson reduces effort as the ability of the lower-ability coworker decreases. The pattern of the change in Salesperson 1's sales is opposite in the RS booths; here, compared with when she is paired with Salesperson 2, her sales *increase* when she is paired with Salesperson 3 and increase still further when she is paired with Salesperson 4.

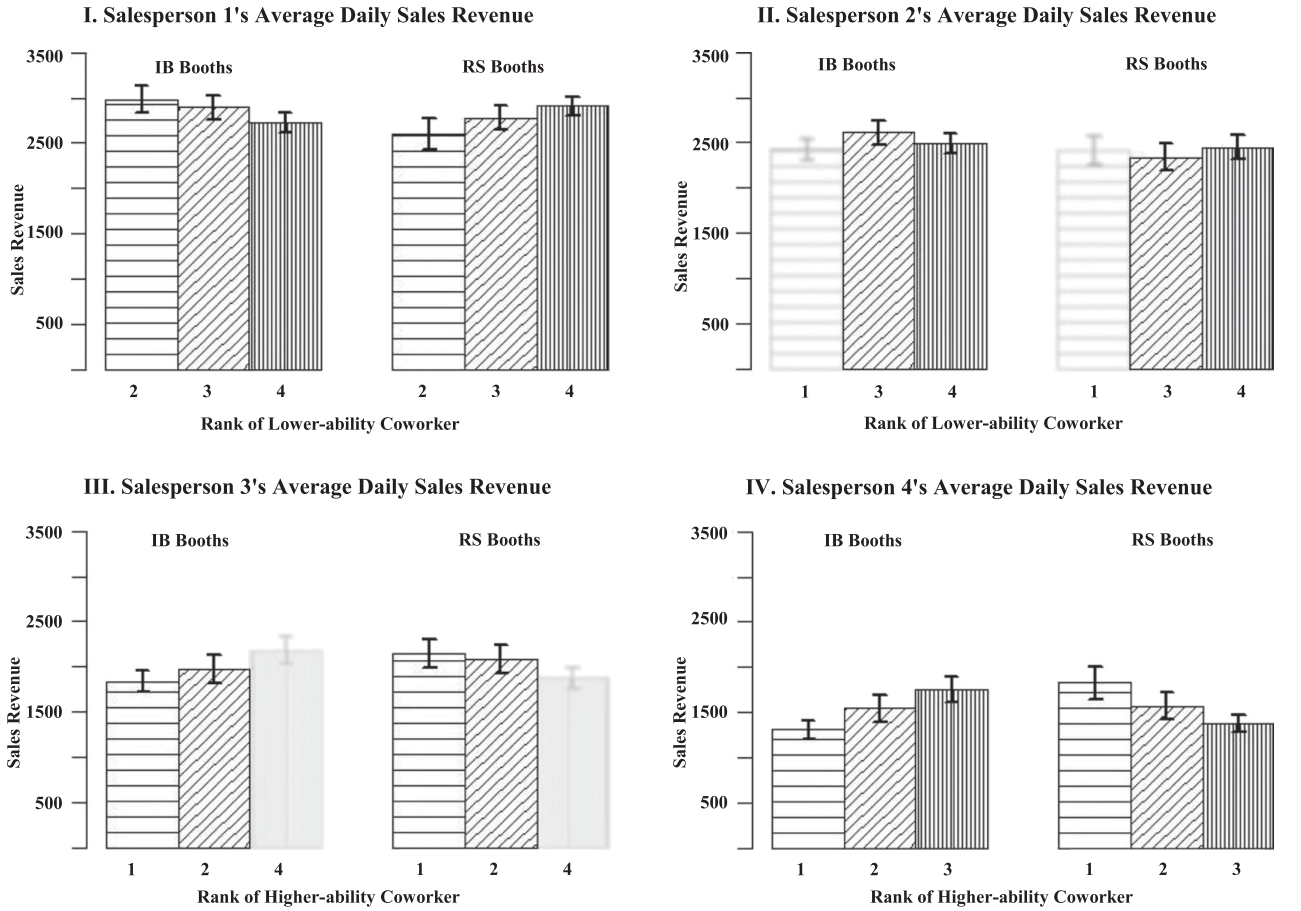
Next, we turn to panel IV of Figure 2, which describes the change in sales for Salesperson 4, who has the lowest sales ability in the booth. The graph shows what happens when Salesperson 4 is paired with a coworker with progressively higher ability. In the IB booths, Salesperson 4's sales decrease when her coworker's ability increases. In contrast, under the RS incentive, Salesperson 4's sales increase when her coworker's ability increases.

In contrast to Salespersons 1 and 4, the ability levels of the coworkers that Salespersons 2 and 3 are paired with can be either higher or lower. As shown in panel II of Figure 2, when Salesperson 2 is paired with a lower-ability coworker, the changes in sales appear similar to what we find for Salesperson 1, under both the IB and RS incentives. When Salesperson 3 is paired with a higher-ability salesperson (as indicated in panel III of Figure 2), the sales changes are again similar to what we observe for Salesperson 4. We proceed to examine the data using formal statistical tests.

### 5.2. Model Specification for Stage 1 Data

The key objective of Stage 1 in our experiment is to examine how salespeople might adjust effort when paired with coworkers of varying sales abilities, under the IB and RS incentives. Recall that the staffing assignment of Salespeople 1, 2, 3, and 4 in a booth reflects one of three possible *staffing conditions*, A, B, or C, as described in Table 2. In constructing our

**Figure 2.** "Model-Free" Sales Patterns for Salespeople 1–4



model to test these predictions, we use these three conditions to identify how salespeople might respond to coworkers of different abilities. Formally, the specification of Model 1 is as follows:

$$s_{ijt} = \theta_j + \sum_{p=1}^{2}\sum_{r=1}^{4}\sum_{d=1}^{3}\beta_{p,r,d}\ 1\{Incentive(i) = p\} \cdot 1\{Rank(j) = r\} \cdot 1\{Condition(i,t) = d\} + \mathbf{Z}_t\rho + \varepsilon_{ijt}. \tag{4}$$

In Model 1, the dependent variable $s_{ijt}$ is the daily sales revenue of salesperson $j$ working in booth $i$ on day $t$. On the right-hand side, we include a salesperson-level fixed effect $\theta_j$ ($j$ = 1, 2, ..., 116).[16] We also include $\mathbf{Z}_t$, a vector of control variables that may affect sales, such as day of the week (Monday through Saturday), the shift in a day (morning or afternoon), and the day's weather (e.g., rain or shine, average temperature).[17] Finally, $\varepsilon_{ijt}$ is the error term.

Regarding the key variables of interest in the model, the first indicator function, $1\{Incentive(i) = p\}$, describes whether booth $i$ pays its salespeople using the IB or RS incentive. We index this using $p = 1$ and $p = 2$ for the IB and RS booths, respectively. This indicator variable allows us to separately identify the sales effects under the two incentives. The second indicator function, $1\{Rank(j) = r\}$, captures the rank of salesperson $j$ in booth $i$.[18] We index this by $r$ = 1, 2, 3, and 4, to indicate whether the salesperson is Salesperson 1, 2, 3, or 4. For example, $Rank(j) = 3$ indicates that Salesperson $j$ is ranked third in her booth (i.e., Salesperson 3). The third indicator function, $1\{Condition(i,t) = d\}$, captures which of the three staffing conditions, A, B, or C, booth $i$ is assigned to on day $t$. The index $d$ = 1, 2, and 3 represents Condition A, Condition B, and Condition C, respectively.

Given this setup, we have a total of $2 \times 4 \times 3 = 24$ possible $\beta_{p,r,d}$ coefficients in the model, one for each combination of $p$, $r$, and $d$. In the estimation, however, we use eight of these combinations as the reference category, so only 16 of the $\beta_{p,r,d}$ coefficients can be identified in the estimation. Specifically, note that 12 $\beta$ parameters (out of 24) are associated with the IB booths (i.e., when $p$ = 1). In this case, for each salesperson $r$ = 1, 2, 3, or 4, we need to use one of the three staffing conditions (i.e., Condition A, B, or C, denoted by $d$) as the reference category. Consequently, we treat four $\beta$s as the reference category. The remaining eight $\beta_{p,r,d}$ coefficients measure the impact of the various staffing conditions on the sales of the salespeople of different ranks in the IB booths, *with respect to* the selected reference categories. For example, if we choose Condition C ($d$ = 3) as the reference for Salesperson 1 ($r$ = 1), the estimated $\beta_{1,1,1}$ measures how the sales of the highest-ranked salesperson in an IB booth change in staffing Condition A compared with staffing Condition C. The $\beta$ parameters for the RS booths can be similarly described.

We can use Model 1 to test the predictions of salespeople's behavior. For example, to examine how Salesperson 1 might adjust her effort as the ability level of her lower-ability coworker decreases, we can compare how her sales change when her coworker is switched from Salesperson 2 (Condition C) to Salesperson 3 (Condition B) and then to Salesperson 4 (Condition A). We estimate three versions of Model 1, in which we choose Condition A, B, or C as the reference category for each of the four salespeople. The different reference categories in the three versions allow us to interpret the results more easily. Panels I and II of Table 4 present the key estimates of Model 1 on the Stage 1 data, which we collate from all three versions. We report the full set of estimates in Table B.1 of Appendix B. Note that all the standard errors are clustered at the booth level.[19]

**Results for the IB Booths.** We begin by discussing the results for the IB booths. We start with the sales patterns for Salesperson 1. As Panel I of Table 4 shows, compared with when she is paired with Salesperson 2 (the reference category), Salesperson 1's average daily sales exhibit a directional (though not significant) decrease of ¥164.81 when she is paired with Salesperson 3, whose sales ability is lower than that of Salesperson 2. When Salesperson 1 is paired with Salesperson 4, the salesperson with the lowest sales ability in the booth, her average daily sales show a statistically significant drop of ¥343.03, compared with when she is paired with Salesperson 2.[20] Panel I of Table 4 shows that Salesperson 2's average daily revenue is significantly lower by ¥179.17 when she is paired with Salesperson 4 than when she is paired with Salesperson 3 (the reference category).

Next, panel I of Table 4 shows that compared with when she is paired with Salesperson 3 (the reference category), Salesperson 4's average daily sales decrease by ¥307.2 when she is paired with Salesperson 2. When she is paired with Salesperson 1, Salesperson 4's sales drop even further, by ¥626.19. For Salesperson 3, Panel I of Table 4 shows that compared with when she is paired with Salesperson 2 (the reference category), her average daily sales decrease by ¥203.13 when she is paired with Salesperson 1. Together, these results are consistent with Prediction 1, which states that given an IB contract, when salespeople make social comparisons about their relative earnings, the higher-ability salesperson will decrease effort as the ability of the salesperson she is paired with decreases, while the lower-ability salesperson will decrease effort when the ability level of her (higher-ability) coworker increases.

**Results for the RS Booths.** We now turn to the RS booths. We first report the results for the higher-ability

**Table 4.** Key Estimation Results of Model 1

| | | | I: Stage 1 | II: Stage 1 | III: Stage 2 | IV: Stage 2 |
|---|---|---|---|---|---|---|
| Salesperson | Condition | Staffing composition | IB booths | RS booths | IB booths | RS booths |
| Salesperson 1 | A | 1-4 | –343.03***[a] (124.45) | 388.36***[a] (144.57) | –318.02***[a] (116.55) | 322.74**[a] (164.45) |
| | B | 1-3 | –164.81[a] (128.89) | 181.23[a] (137.29) | –179.46[a] (227.75) | 135.38[a] (140.51) |
| | C | 1-2 | Reference | Reference | Reference | Reference |
| Salesperson 2 | B | 2-4 | –179.17*[a] (107.84) | 218.09[a] (169.83) | –204.12*[a] (114.88) | 273.64*[a] (161.15) |
| | A | 2-3 | Reference | Reference | Reference | Reference |
| Salesperson 3 | B | 3-1 | –203.13*[a] (107.47) | 96.85[a] (172.59) | –185.18*[a] (98.49) | 89.33[a] (136.25) |
| | A | 3-2 | Reference | Reference | Reference | Reference |
| Salesperson 4 | A | 4-1 | –626.19***[a] (132.76) | 612.13***[a] (154.22) | –626.61***[a] (137.37) | 553.07***[a] (171.58) |
| | B | 4-2 | –307.20***[a] (137.49) | 272.44**[a] (146.43) | –381.57***[a] (136.24) | 231.98**[a] (109.22) |
| | C | 4-3 | Reference | Reference | Reference | Reference |

*Notes.* Three versions of Model 1 are estimated, in which one of the Conditions A, B, or C serves as the reference category. This table collates the key estimates from all three versions. Panel I and II (Panel III and IV) present the key results of estimating Model 1 on Stage 1 (Stage 2) data. The dependent variable is individual daily sales (in Chinese Renminbi). Standard errors are clustered at the booth level. $R^2$ is 0.42 and 0.39 for Model 1 estimation on Stage 1 data and Stage 2 data, respectively. Under the column "Staffing composition," the first number represents the rank of the focal salesperson, and the second number represents the rank of the coworker.

[a]Estimates are significantly different across the IB and RS incentives for that staffing composition, at the 5% level.

*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

salespeople (Salespeople 1 and 2), beginning with Salesperson 1. Panel II of Table 4 shows that compared with when Salesperson 1 is paired with Salesperson 2 (the reference category), when she is paired with Salesperson 3, her sales show a directional increase of ¥181.23, but this is not statistically significant. When Salesperson 1 is paired with Salesperson 4, however, her sales show a significant increase of ¥388.36. For Salesperson 2, the findings provide mild support for our prediction: sales are directionally higher by ¥218.09 when she is paired with Salesperson 4 than when she is paired with Salesperson 3, although the change is not statistically significant.

Next, we examine the results for Salesperson 4, the salesperson with the lowest sales ability in the booth. As panel II of Table 4 shows, Salesperson 4's average daily sales increase by ¥272.44 when she is paired with Salesperson 2, compared with when she is paired with Salesperson 3 (the reference category). When she is paired with Salesperson 1, her sales increase even more, by ¥612.13. For Salesperson 3, there is a directional increase of ¥96.85 in sales when she is paired with Salesperson 1, compared with when she is paired with Salesperson 2 (the reference category); however, the difference is not statistically significant. Taken together, the empirical results indicate support for Prediction 2, which states that under an RS incentive, if salespeople are averse to "pulling down" the earnings of their coworker in the shift, the higher-ability salesperson will increase effort as the ability of the salesperson she is paired with decreases, while the lower-ability salesperson will exert greater effort as the ability of her coworker rises.[21]

**Comparison of Estimates Across the IB and RS Booths.** Predictions 1 and 2 together suggest that salespeople with the same ability rank would adjust effort in opposite directions across the IB and RS incentives, when their coworker's ability changes. Specifically, when their coworker's ability decreases, the higher-ability salesperson would lower effort under the IB incentive but raise effort under the RS incentive, while the lower-ability salesperson would increase effort under the IB incentive but decrease effort under the RS incentive. To examine this, for each staffing composition, we compare the estimates across the IB and RS incentives. The results (shown in Table 4, denoted by note "a") demonstrate support for this theory: for every staffing composition, the estimate of the IB incentive is significantly different from that of the RS incentive, at the 5% level.[22]

**Alternative Model Specification.** In the empirical model described previously (Model 1), we test the theoretical predictions using the *ability rank* of the salespeople in a booth, which is a discrete variable. In Model 2 below, we now capture differences in coworkers' abilities using a *continuous* variable:

$$
\begin{aligned}
s_{ijt} = \theta_j + \sum_{p=1}^{2} 1\{Incentive(i) = p\} \cdot \\
&[\tau_{p,1} \cdot 1\{Rank(j) = 1\} \cdot 1\{Rank(j) > Rank(-j)\} \\
&+\tau_{p,2} \cdot 1\{Rank(j) = 2\} \cdot 1\{Rank(j) > Rank(-j)\} \\
&+\tau_{p,3} \cdot 1\{Rank(j) = 3\} \cdot 1\{Rank(j) < Rank(-j)\} \\
&+\tau_{p,4} \cdot 1\{Rank(j) = 4\} \cdot 1\{Rank(j) < Rank(-j)\}] \cdot \delta_{i,t}^{d} \\
&+\mathbf{Z}_t \rho + \varepsilon_{ijt}\,, \qquad (5)
\end{aligned}
$$

where $Rank(j) > Rank(-j)$ and $Rank(j) < Rank(-j)$ indicate that the focal salesperson $j$'s ability is higher and lower than that of the coworker (salesperson $-j$) whom she is paired with on day $t$, respectively. The continuous variable $\delta_{i,t}^{d}$ captures the ability difference between salesperson $j$ and the coworker, which depends on the staffing condition $d$ in booth $i$ on day $t$.

To construct the $\delta$ variables, we again use a salesperson's average daily sales in the past 180 days before the experiment as a proxy for sales ability. Using the example in Section 4.2, suppose that the average daily sales of Salespeople 1, 2, 3, and 4 are ¥3,000, ¥2,400, ¥1,800, and ¥1,200, respectively. If Salesperson 2 is assigned to staffing Condition A on day $t$ (meaning that she is paired with Salesperson 3), then $\delta_{2,t}^{A}$ = ¥2,400 − ¥1,800 = ¥600. Similarly, for staffing Conditions B and C, we have $\delta_{2,t}^{B}$ = ¥1,200 and $\delta_{2,t}^{C}$ = − ¥600, respectively. Thus, $\delta_{i,t}^{d}$ can be either positive or negative.

In Model 2, the coefficients to be estimated are represented by $\tau_{p,r}\,(r = 1, 2, 3, 4)$, which measures the impact of the ability difference between the focal salesperson and the coworker on the sales of the former, relative to when the ability difference is zero. Three points are worth noting here. First, as stated in Section 4.2, we operationalize the change in the coworker's ability by replacing a salesperson's coworker with another one via the randomized staffing assignment. Since there are four salespeople in each booth, Model 2 only estimates how Salesperson 1 and Salesperson 2 adjust effort when their lower-ability coworker's ability varies, and how Salesperson 3 and Salesperson 4 respond to the change in their (higher-ability) coworker's ability.

Second, the interpretation of $\tau$ depends on whether the salesperson has higher or lower ability than her coworker. For example, because the ability of Salesperson 1's coworker is always lower than hers, a negative $\tau$ implies that Salesperson 1 reduces effort as the ability difference between her and her coworker increases. Conversely, a negative $\tau$ for Salesperson 4 (where all her coworkers have higher abilities) implies that she increases effort as the ability gap widens.

Third, recall that our theoretical predictions (Predictions 1 and 2) and the model-free results of Stage 1 data suggest that salespeople in the IB (RS) booths will decrease (increase) effort when the difference in sales ability enlarges. Therefore, we should expect the estimates for the higher-ability salespeople (Salesperson 1 and Salesperson 2) of Model 2 to be negative (positive) under the IB (RS) incentive, while that for the lower-ability salespeople (Salesperson 3 and Salesperson 4) to be positive (negative) under the IB (RS) incentive.

Table 5 reports the estimation results of Model 2. The pattern of sales changes is consistent with Predictions 1 and 2 and what we have found from the discrete model (Model 1). For example, for the IB booths, the estimate of –0.240 for Salesperson 1 implies that her sales and effort decline when the ability difference between her and her coworker increases. In the RS booths, the estimated coefficient for Salesperson 4 is –1.293, which suggests that the lower-ability salesperson increases effort as the ability gap with her higher-ability coworker increases.

## 5.3. Examining Alternative Theories of the Empirical Findings in Stage 1

The empirical analyses of the Stage 1 data suggest that our theory that incorporates social preferences (namely, social comparisons and guilt aversion in the IB and RS contracts, respectively) provides a plausible explanation for salespeople's behavior. We now proceed to discuss and rule out several alternative explanations.

**Learning from Coworkers.** During the same shift, salespeople can observe their coworker's selling techniques. In particular, the lower-ability salespeople might learn more from being paired with a higher-ability coworker and thus increase her revenues. This is consistent with our finding for the lower-ability

**Table 5.** Estimation Results of Model 2

| Salesperson's rank | Coworkers' ranks | Estimated $\tau$: IB booths | Estimated $\tau$: RS booths |
|---|---|---|---|
| 1 | 2, 3, 4 | –0.240*** (0.104) | 0.278*** (0.129) |
| 2 | 3, 4 | –0.221** (0.105) | 0.256** (0.103) |
| 3 | 1, 2 | 0.234*** (0.102) | –0.746*** (0.230) |
| 4 | 1, 2, 3 | 0.798*** (0.136) | –1.293*** (0.337) |

*Notes.* The dependent variable is individual daily sales (in Chinese Renminbi). Standard errors are clustered at the booth level.

*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. $R^2$ is 0.41.

salesperson, but only in the RS booths. In the IB booths, however, we find that the sales revenue of the lower-ability salesperson *decreases* as her coworker's ability increases. If learning were a primary explanation, then we should expect the sales of the lower-ability salesperson to increase as she learns from better coworkers. Moreover, learning cannot explain why the sales of the higher-ability salesperson would increase when her coworker's sales ability decreases under the RS incentive. Finally, the salespeople in the same booth have been paired frequently for at least 6 months prior to the experiment; hence, most of the learning should already have taken place.[23]

**Effort Adjustment Based on Daily Target Income.** Another theory could be that salespeople have a daily target income and might adjust effort to reach this target. Specifically, to reach her daily target income, the higher-ability salesperson in the RS booth would exert higher effort to compensate for the lower sales contribution by her lower-ability coworker. This is consistent with our empirical finding for the higher-ability salespeople in the RS booths. Next, based on the target income hypothesis, note that the lower-ability salesperson in the RS booth should *not* increase effort when her coworker's ability increases, because it is now easier to reach her target due to her coworker's greater higher sales contribution. However, we find the opposite in the data: sales of the lower-ability salesperson in the RS booth increase when her coworker's ability rises. Furthermore, in the IB booths, the target income hypothesis predicts that salesperson effort would be unaffected by their coworker's abilities, since their earnings do not depend on their coworker's sales revenue. This is also inconsistent with our findings from the IB booths.

**Effort Complementarities.** There could be effort complementarities between coworkers under a group-based incentive (either as an antecedent or a consequence of the group incentive), particularly in a factory production line (Hamilton et al. 2003, Hossain and List 2012). To begin, it is important to note that, unlike in a factory production line, the salesperson's output in our setting (i.e., the dollar amount that the customer served by the salesperson spends) is relatively unaffected by her coworker's effort: as mentioned earlier, in the context of retailing women's clothing in a high-end department store, customers are typically served by one salesperson during each visit. Nevertheless, we analytically examine how effort complementarities (if they exist in our context) would affect salespeople's effort when their coworker's ability changes.

One way to model production complementarities is to assume that the sales of salesperson $j$ ($j = H, L$) are given by $y_j = \alpha_j e_j(1 + \alpha_{-j}e_{-j})$. In this way, a higher level of effort expended by coworker $-j$ raises salesperson $j$'s output. Recall that a salesperson's utility in the RS booth is $U_j = m(y_j + y_{-j})/2 - e_j^2$. The first-order condition is thus

$$\frac{\partial U_j}{\partial e_j} = \frac{m}{2}[\alpha_j\alpha_{-j}e_{-j} + \alpha_j(1 + \alpha_{-j}e_{-j})] - 2e_j = 0. \quad (6)$$

Solving for effort, we obtain $e_j^* = (2m\alpha_j + m^2\alpha_j\alpha_{-j}^2)/[2(4 - m^2\alpha_j^2\alpha_{-j}^2)]$. More importantly, we have $\partial e_j^*/\partial\alpha_{-j} = 2m^2\alpha_j(2 + m\alpha_j^2)\,\alpha_{-j}/(m^2\alpha_j^2\alpha_{-j}^2 - 4)^2 > 0$, i.e., the salesperson's effort increases in the ability level of her coworker.[24] This result is intuitive because when the ability of the coworker increases, the returns to effort are higher, so it pays to expend more effort. While this theory can explain why the lower-ability salesperson in the RS booths increases effort when paired with coworkers of increasing ability, it is unable to explain why effort of the higher-ability salesperson increases as the ability of her lower-ability coworker decreases.

## 6. Empirical Analysis on Stage 2 Data: Implications for Firm Performance

So far, we have empirically shown how salespeople strategically allocate their effort when the ability of their coworker in the team changes. We now use the data from both Stages 1 and 2 to examine Prediction 3, which is about the revenue implications of team composition.

Recall that in Stage 2 of our experiment, we randomly assigned each of the 29 booths to one of three treatments. First, we randomly selected three IB brands and three RS brands and assigned them to the Heterogeneous treatment, in which a high-ability salesperson was paired with a low-ability salesperson. Second, we randomly selected two IB booths and two RS booths and assigned them to the Homogeneous treatment, in which salespeople of similar ability levels were paired. In these two treatments, we fixed the staffing assignment so that salespeople who are paired work together for 21 consecutive days. Third, the remaining 19 booths, which consist of 11 IB booths and 8 RS booths, formed the Random treatment, in which the staffing assignment was identical to that in Stage 1 (i.e., randomized staffing assignment every day).

We begin the analysis by running Model 1 (Equation (4)) on the data for the 19 booths in the Random treatment. We report the results in panels III (for the IB booths) and IV (for the RS booths) of Table 4. The estimates and patterns of sales changes are almost identical to the findings from the Stage 1 data. For example, when Salesperson 1 in an IB booth is paired with Salesperson 4, her average daily sales decrease by ¥318.02, compared with when she is paired with

Salesperson 2. As in Stage 1, for Salesperson 4 in an RS booth, her sales increase by ¥553.07 when she is paired with Salesperson 1, compared with when she is paired with Salesperson 3. The only difference in the results between Stages 1 and 2 is for Salesperson 2 in the RS booths (panel IV): Her sales increase by ¥273.64 when she is paired with Salesperson 4, compared with when she is paired with Salesperson 3 (the reference category). These findings lend further support to Predictions 1 and 2.

Now, we investigate how varying the team composition affects firm performance. To examine Prediction 3, we use Model 3 below to examine the data across Stages 1 and 2:

$$S_{it} = \omega_i + \chi_1 Hete_{it}^{IB} + \chi_2 Homo_{it}^{IB} + \chi_3 Hete_{it}^{RS} + \chi_4 Homo_{it}^{RS} + \xi_t + \boldsymbol{Z}_t\theta + \varepsilon_{it}, \quad (7)$$

where $S_{it}$ is the total daily sales generated in booth $i$ on day $t$ and $\omega_i$ is a fixed effect that captures the brand equity of booth $i$. The four variables $Hete_{it}^{IB}$, $Homo_{it}^{IB}$, $Hete_{it}^{RS}$, and $Homo_{it}^{RS}$ are indicators: $Hete_{it}^{IB}$ ($Hete_{it}^{RS}$) is equal to 1 if booth $i$ employs an IB (RS) incentive and is assigned to the Stage 2 Heterogeneous treatment on day $t$, and 0 otherwise. The coefficient $\chi_1(\chi_3)$ measures the impact of the Heterogeneous treatment on booth revenues, relative to when the staffing assignment is randomized (during Stage 1), under the IB (RS) incentive. Similarly, $Homo_{it}^{IB}$ ($Homo_{it}^{RS}$) is equal to 1 if booth $i$ uses an IB (RS) incentive and is assigned to the Homogeneous treatment on day $t$, and 0 otherwise. Thus, $\chi_2(\chi_4)$ measures the impact of the Homogeneous treatment on booth revenues, relative to when salespeople in the booth are randomly assigned, under the IB (RS) incentive. $\xi_t$ is a dummy variable that indicates whether day $t$ belongs to Stage 1 or Stage 2, $\boldsymbol{Z}_t$ is a vector of control variables as defined in Model 1, and $\varepsilon_{it}$ is the error term.

Model 3 is essentially a difference-in-differences model that controls for time-varying differences across Stages 1 and 2. Table 6 reports the estimates of Model 3. The estimate of $\chi_1$ is −654.25, which indicates that under an IB incentive, when salespeople of different abilities are always paired together, booth revenues will decline by ¥654, compared with when the staffing assignment is randomly determined. The estimate is statistically significant. Given that the average daily sales of an IB booth are ¥9124 (Table 1), this drop represents a 7.16% reduction in revenues, which is economically significant. However, if an IB booth were always to assign salespeople of similar abilities to work together (i.e., Homogeneous treatment), our estimate of suggests that booth revenues will rise by ¥336 a day, a 3.68% increase, compared with when salespeople are randomly assigned. These findings are consistent with Prediction 3 for the IB booths.

**Table 6.** Estimation Results of Model 3 (Stages 1 and 2 Data)

| $\chi_1$ | $\chi_2$ | $\chi_3$ | $\chi_4$ |
|---|---|---|---|
| −654.25*** | 336.41* | 814.85*** | −407.37** |
| (180.96) | (177.63) | (291.17) | (178.91) |

*Notes.* The dependent variable is booth daily sales (in Chinese Renminbi). Standard errors are clustered at the booth level.
*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. $R^2$ is 0.53.

The results for the RS booths suggest that the optimal ability mix in a sales team differs from that under the IB incentive. The estimates of and show the following: If the ability mix in the team is heterogeneous, revenues will increase by ¥815 (8.89%), compared with when salespeople are randomly paired. Furthermore, forming a sales team with salespeople who are homogeneous in ability will decrease daily revenues in an RS booth by ¥407 (4.43%). Thus, the data are supportive of Prediction 3 for the RS booths as well.[25]

## 7. Conclusion

In this article, we investigate how salespeople in a team allocate their effort (under the IB and RS incentives) when the ability of their coworker changes, using a field experiment in which we were able to randomize the staffing assignment of salespeople working for various booths in the women's clothing section of a major Chinese department store. The data suggests that as the ability gap between the coworkers widens, salesperson effort decreases (increases) under the IB (RS) incentive. These empirical findings broadly support our theory that incorporates social preferences in explaining salespeople's effort decisions. Several alternative theories that might also be able to explain salespeople's behavior are carefully examined and ruled out.

We also examine the revenue implications of team composition on firm performance and find that while a firm that uses an IB incentive should have more homogeneous sales teams, a firm that employs an RS incentive might benefit more if the ability levels in the team are more heterogeneous. Our results suggest that firms can achieve significant economic gains simply by optimizing the ability mix of a retail sales team.

**Discussion and Future Research.** We conclude with a discussion on the generalizability of our empirical observations and related managerial implications, which can also serve as directions for future research.

First, the empirical findings in this article must be interpreted with the caveat that salespeople in the retail teams we examine have minimum task interdependence, i.e., each salesperson's sales revenue can be readily attributed to her individual effort. Therefore, for retail settings where salespeople usually carry out their daily selling tasks individually, such as salespeople at car dealerships and agents of insurance companies, our findings will very likely apply. Meanwhile, task interdependence may exist in other types of teams and can influence team performance (Wageman 1995, Van der Vegt and Janssen 2003). For settings where task interdependence or complementarities play an important role (e.g., team selling is critical in customer acquisition, or different salespeople in a team perform different specialized functions) so that it is difficult to assign individual credit given a sales outcome, our findings might not apply.

Second, the management literature has documented various factors that can moderate the relationship between the heterogeneity/diversity of a team and team performance, such as the nature of tasks (Bantel and Jackson 1989) and organizational culture (Brickson 2000, Ely and Thomas 2001). These factors can also affect the generalizability of our empirical findings and should be examined closely. For instance, when a sales team is implementing routine sales tasks, as in our experimental setting, then we should expect our findings to apply. However, if a sales team is selling a complex product or service so that the selling task requires more creativity and innovation, how heterogeneity affects team performance deserves a closer examination. The extant literature (e.g., Jackson et al. 1995, Guzzo and Dickson 1996) suggests that increased diversity and heterogeneity of a team can encourage creativity/innovation. However, if the ability gap among coworkers becomes too large, the lower-ability or less experienced team members may feel intimidated by their higher-ability coworkers, and thus may not fully express their opinions, leading to a narrower set and possibly less efficient solutions.

Third, the size of the team is another important factor that can affect team performance (Cunningham and Chelladurai 2004, Stewart 2006). In this paper, our results derive from a setup of a two-person team. When the size of a team increases (e.g., from two to three salespeople), how will the sales ability of the third salesperson affect the effort decisions of the other two salespeople, compared with the case where there are only two team members? The theoretical analysis would be much more complicated, but our fundamental intuition would still apply. For instance, assume that the ability of the third salesperson is in between the highest- and lowest-ability salespeople. We can show that under the IB incentive, as the ability of the third salesperson decreases, the lowest-ability salesperson will increase effort. This is because the lowest-ability salesperson now has a higher chance to avoid coming last in the three-person team. In response to the increased effort by the lowest-ability salesperson, the highest-ability salesperson would also raise effort. In contrast, under the RS incentive, when the sales ability of the third salesperson becomes closer to that of the lowest-ability salesperson, the latter will expend lower effort than before, because the chance that she would be the member that pulls down the earnings of her other team members diminishes. Correspondingly, the highest-ability salesperson would best-respond by dropping her effort as well.

Fourth, our data suggests that managers should strategically design or manipulate the social environment in which salespeople interact with one another. For example, when managers disclose salespeople's performance or reward (punish) the top (low) performers in the sales team, social comparisons would become more salient (Chen et al. 2011). Therefore, when the ability mix is large, managers should strategically avoid performance disclosure under the IB incentive; whereas under the RS incentive, managers might choose to highlight performance to induce higher effort from team members who are averse to dragging down their coworkers' earnings. Moreover, given an RS contract, managers could invest in building and strengthening the social bonds among salespeople (e.g., organizing workshops or retreats for team members, encouraging communication between coworkers), to motivate greater effort as team members care more about one another's earnings.

Lastly, our findings can be helpful to managers who make decisions on the type of compensation contracts to offer to their sales force. Our paper provides field evidence that different nonpecuniary drivers of behavior can arise under different contracts. For instance, to motivate relatively homogeneous sales teams, managers could consider using individual- or relative-based incentives such as sales contests, because such contracts can induce greater social comparisons about relative status within the sales team. On the other hand, we conjecture that other compensation contracts such the quota-bonus may reduce social comparisons, because salespeople may now evaluate themselves against their own quota levels, particularly when salespeople have heterogeneous abilities and may have personalized quota levels (e.g., quotas are a percentage of one's previous sales). We leave this question to future research.

## Acknowledgments

The paper has benefited significantly from valuable comments and suggestions by the senior editor, the associate editor, and the two anonymous referees. The first author acknowledges financial support from Marketing Science Institute (MSI). The suggestions by Tat Chan and Lamar Pierce, and their support to achieve the MSI financial support, are also deeply appreciated. The authors are grateful for all those involved in running the experiment.

## Appendix A: Derivation of Conditions for the Comparative Statics Analyses in Section 3

### IB Incentive

Under the IB incentive, the equilibrium effort of salespeople $H$ and $L$, who have social loss aversion, are $e_H^* = \alpha_H\{4V\gamma_H + m[8V^2 + \alpha_L^2(\gamma_H - \gamma_L)]\}/[2(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)]$ and $e_L^* = \alpha_L\{4V\gamma_L + m[8V^2 + \alpha_H^2(\gamma_H - \gamma_L)]\}/\ [2(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)]$, respectively. Taking the derivative of salespeople's equilibrium effort with respect to their coworker's ability, we have $\partial e_H^*/\partial\alpha_L = \alpha_H\alpha_L\gamma_H\{4V\gamma_L + m[8V^2 + \alpha_H^2(\gamma_H - \gamma_L)]\}/(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)^2$ and $\partial e_L^*/\partial\alpha_H = -\alpha_H\alpha_L\gamma_L\ \{4V\gamma_H + m\cdot[8V^2\ + \alpha_L^2(\gamma_H - \gamma_L)]\}/(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)^2$.

To obtain $\partial e_H^*/\partial\alpha_L > 0$, we need

$$\alpha_H\alpha_L\gamma_H\{4V\gamma_L + m[8V^2 + \alpha_H^2(\gamma_H - \gamma_L)]\} > 0 \tag{A.1}$$

since $(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L)^2 > 0$. Similarly, we require

$$-\alpha_H\alpha_L\gamma_L\{4V\gamma_H + m[8V^2 + \alpha_L^2(\gamma_H - \gamma_L)]\} < 0 \tag{A.2}$$

to have $\partial e_L^*/\partial\alpha_H < 0$. Furthermore, to maintain positive equilibrium effort (i.e., $e_H^* > 0$ and $e_L^* > 0$), we need

$$2(8V^2 + \alpha_H^2\gamma_H - \alpha_L^2\gamma_L) > 0 \tag{A.3}$$

given Inequalities (A.1) and (A.2).

From Inequalities (A.1) and (A.2), we obtain $\gamma_H > [(\alpha_H^2 m - 4V)\gamma_L - 8mV^2]/(\alpha_H^2 m)$ and $\gamma_H > m(\alpha_L^2\gamma_L - 8V^2)/(\alpha_L^2 m + 4V)$, respectively. Solving Inequality (A.3), we can get $\gamma_H > (\alpha_L^2\gamma_L - 8V^2)/\alpha_H^2$.

To sum up, we can show that when $\gamma_H > max\{(\alpha_L^2\gamma_L - 8V^2)/\alpha_H^2,\ [(\alpha_H^2 m - 4V)\gamma_L - 8mV^2]/(\alpha_H^2 m),\ m(\alpha_L^2\gamma_L - 8V^2)/(\alpha_L^2 m + 4V)\}$, $\partial e_H^*/\partial\alpha_L > 0$ and $\partial e_L^*/\partial\alpha_H < 0$ in the IB booths. In other words, when social loss aversion is sufficiently strong, the higher-ability salespeople will increase (decrease) effort when coworker's ability increases (decreases), while the lower-ability salespeople will decrease (increase) their effort when the coworker's ability increases (decreases).

### RS Incentive

Under the RS incentive, when salespeople have guilt aversion, the equilibrium effort of Salespeople $H$ and $L$ are $e_H^* = \alpha_H\{8V\psi_H + m[8V^2 - \alpha_L^2(\psi_L - \psi_H)]\}/[4(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L)]$ and $e_L^* = \alpha_L\{8V\psi_L + m[8V^2 - \alpha_H^2(\psi_L - \psi_H)]\}/[4(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L)]$ respectively. Taking the derivative of equilibrium effort with respect to their coworker's ability, we have $\partial e_H^*/\partial\alpha_L = \alpha_H\alpha_L\psi_H\{8V\psi_L + m[8V^2 - \alpha_H^2(\psi_L - \psi_H)]\}/[2(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L)^2]$ and $\partial e_L^*/\partial\alpha_H = -\alpha_H\alpha_L\psi_H\{8V\psi_H + m[8V^2 - \alpha_L^2(\psi_L - \psi_H)]\}/[2(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L)^2]$.

Given that $2(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L)^2 > 0$, to obtain $\partial e_H^*/\partial\alpha_L < 0$ we need

$$\alpha_H\alpha_L\psi_H\{8V\psi_L + m[8V^2 - \alpha_H^2(\psi_L - \psi_H)]\} < 0. \tag{A.4}$$

Similarly, we require

$$-\alpha_H\alpha_L\psi_H\{8V\psi_H + m[8V^2 - \alpha_L^2(\psi_L - \psi_H)]\} > 0 \tag{A.5}$$

to have $\partial e_L^*/\partial\alpha_H > 0$. Furthermore, we need

$$4(8V^2 + \alpha_H^2\psi_H - \alpha_L^2\psi_L) > 0 \tag{A.6}$$

to maintain $e_H^* > 0$ and $e_L^* > 0$, given Inequalities (A.4) and (A.5).

From Inequalities (A.4) and (A.5), we obtain $\psi_L > m(\alpha_H^2\psi_H + 8V^2)/(\alpha_H^2 m - 8V)$ and $\psi_L > [8\psi_H V + m(\alpha_L^2\psi_H + 8V^2)]/(\alpha_L^2 m)$, respectively. Solving Inequality (A.6), we can get $\psi_L > (\alpha_H^2\psi_H + 8V^2)/\alpha_L^2$.

To sum up, when $\psi_L > max\{(\alpha_H^2\psi_H + 8V^2)/\alpha_L^2,\ [8\psi_H V + m(\alpha_L^2\psi_H + 8V^2)]/(\alpha_L^2 m),\ m(\alpha_H^2\psi_H + 8V^2)/(\alpha_H^2 m - 8V)\}$, $\partial e_H^*/\partial\alpha_L < 0$ and $\partial e_L^*/\partial\alpha_H > 0$ in the RS booths. That means, if salespeople have strong guilt aversion, both the higher- and lower-ability salespeople will increase their effort under the RS incentive when the difference in ability between two salespeople increases.

## Appendix B: Additional Tables of Empirical Analyses

**Table B.1.** Full Estimation Results of Model 1

| | | | I: Stage 1 data; Condition C as base | | II: Stage 1 data; Condition B as base | | III: Stage 1 data; Condition A as base | | IV: Stage 2 data; Condition C as base | |
|---|---|---|---|---|---|---|---|---|---|---|
| Salesperson | Condition | Staffing composition | Estimated $\beta$ for IB booths | Estimated $\beta$ for RS booths | Estimated $\beta$ for IB booths | Estimated $\beta$ for RS booths | Estimated $\beta$ for IB booths | Estimated $\beta$ for RS booths | Estimated $\beta$ for IB booths | Estimated $\beta$ for RS booths |
| Salesperson 1 | A | 1-4 | –343.03*** (124.45) | 388.36*** (144.57) | –178.22** (84.14) | 207.12* (109.39) | Reference | Reference | –318.02*** (116.55) | 322.74** (164.45) |
| | B | 1-3 | –164.81 (128.89) | 181.23 (137.29) | Reference | Reference | 178.22** (84.14) | –207.12* (109.39) | –179.46 (227.75) | 135.38 (140.51) |
| | C | 1-2 | Reference | Reference | 164.81 (128.89) | –181.23 (137.29) | 343.03*** (124.45) | –388.36*** (144.57) | Reference | Reference |
| Salesperson 2 | A | 2-3 | 299.71** (137.37) | –160.39 (153.99) | 179.17* (107.84) | –218.09 (169.83) | Reference | Reference | 330.21** (152.35) | –194.30 (165.61) |
| | B | 2-4 | 120.54 (132.57) | 57.71 (146.23) | Reference | Reference | –179.17* (107.84) | 218.09 (169.83) | 126.09 (190.18) | 79.34 (151.19) |
| | C | 2-1 | Reference | Reference | –120.54 (132.57) | –57.71 (146.23) | –299.71*** (137.37) | 160.39 (153.99) | Reference | Reference |
| Salesperson 3 | A | 3-2 | –259.75** (134.76) | 309.60** (159.54) | 203.13* (107.47) | –96.85 (172.59) | Reference | Reference | –259.85* (146.62) | 303.69** (150.28) |
| | B | 3-1 | –462.88*** (139.66) | 406.45*** (148.64) | Reference | Reference | –203.13* (107.47) | 96.85 (172.59) | –445.03*** (151.48) | 393.02** (189.22) |
| | C | 3-4 | Reference | Reference | 462.88*** (139.66) | –406.45 (148.64) | 259.75** (134.76) | –309.60** (159.54) | Reference | Reference |
| Salesperson 4 | A | 4-1 | –626.19*** (132.76) | 612.13*** (154.22) | –318.99** (158.24) | 339.69** (169.98) | Reference | Reference | –626.61*** (137.37) | 553.07*** (171.58) |
| | B | 4-2 | –307.20*** (137.49) | 272.44** (146.43) | Reference | Reference | 318.99** (158.24) | –339.69 (169.98) | –381.57*** (136.24) | 231.98** (109.22) |
| | C | 4-3 | Reference | Reference | 307.20*** (137.49) | –272.44** (146.43) | 626.19*** (132.76) | –612.13*** (154.22) | Reference | Reference |

*Notes.* Panels I–III show the results of the three versions of Model 1 that are estimated on the Stage 1 data. The only difference among them is the base or reference category selected for estimation. Panel IV shows Model 1 results estimated on the Stage 2 data, with Condition C as the reference category.

*, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

## Endnotes

[1] As will be detailed later in Sections 3 and 4, we use the changes in a salesperson's sales revenue as a proxy for the changes in her effort (as the latter is unobservable to the researchers). While the analytical modeling literature uses effort as the primary variable of interest, empirical studies (e.g., Kishore et al. 2013, Chung and Narayandas 2017) typically use sales output to examine these theoretical predictions. Our approach holds as long as greater effort translates into higher expected sales.

[2] The department store management and salespeople we interviewed told us that understanding customers' needs and preferences, and making correct suggestions are the most critical factors when selling high-end women's clothing. Therefore, to improve sales effectiveness, a successful salesperson will try to understand the customer better by asking well-designed questions, implementing active listening, and paying particular attention to the changes in the customer's facial expressions and body language when making suggestions.

[3] In our field experiment, we operationalized the change in coworker's ability by replacing a salesperson's coworker with another sales agent who has a different level of sales ability.

[4] We report the details of the derivation of this condition in Appendix A.

[5] It is worth noting that although both the higher- and lower-ability salespeople decrease their effort due to an increasing sales ability gap, their effort levels are higher than the equilibrium effort level derived from the self-interested model. This is because social comparisons drive salespeople to "overexert" effort even when they are compensated on an individual-based contract.

[6] This condition implies that the lower-ability salesperson could have a greater level of guilt aversion than her higher-ability coworker. This is reasonable because the lower-ability salesperson in a shift is more likely to be the party who drags down her coworker's earnings. In a recent study, Chen and Lim (2017) find that in a group-based contest, both the higher- and lower-ability team members exert higher effort relative to the self-interested prediction, but the magnitude of over-exertion is much greater for the lower-ability member. This finding suggests that the lower-ability member is more averse to dragging the team down.

[7] One can also interpret our study to be one that examines the interaction of team heterogeneity and compensation contracts. To begin, it is possible that there is a "main effect" of heterogeneity: When there is no incentive contract (i.e., salespeople are paid by fixed salary only), salespeople may still compare their sales revenues with those of their coworker's, because they care about their status in the team (Clark and Oswald 1998). When there is an IB incentive, the

effect of social comparisons will be much stronger because sales revenues translate into monetary payoffs for salespeople, so relative status is now determined by who actually earns more, and not just by who brings in more revenue. Hence, given the IB incentive, greater heterogeneity induces lower effort; however, the RS incentive moderates this effect.

[8] Of the non-four-salesperson booths, two have five salespeople and one has three salespeople. For these booths, a salesperson works in a third 6-hour shift (12 p.m. to 6 p.m.) every day. In the following sections, we focus our discussion on the booths with four salespeople.

[9] We believe that the past sales performance can be an appropriate proxy for sales ability in our context: First, because there were no changes in compensation schemes or staffing assignment in the 180 pretest days, sales were not affected by these potential confounds. Second, we also asked the store management to evaluate each salesperson's sales ability by ranking salespeople in each booth. These rankings are almost identical to those based on the average sales in the 180 pretest days. Of the 29 booths, there were only four cases where the two sets of rankings are not completely consistent. We reran our empirical models using the rankings based on the managers' evaluations, and the key findings do not change. Finally, we can assume sales ability to be fixed in our analyses, because sales ability is unlikely to change significantly over our short experimental period.

[10] Statistical tests indicate that there are no differences in the individual average daily sales revenue during the 180 pretest days at the 5% level for Salespersons 1, 2, 3, and 4, respectively, across the IB and RS booths.

[11] In the rest of the article, we use the Chinese Renminbi (¥). During our experimental period, the exchange rate between the Chinese Renminbi and the U.S. dollar was approximately 7:1.

[12] During Stage 1, there were occasional instances where the staffing composition in a booth did not follow the experimental design. Specifically, one salesperson in an IB booth took a 3-day sick leave, one salesperson in another IB booth had a 1-day family emergency leave, and one salesperson in an RS booth took a 1-day sick leave. On the days these occurred, a three-shift system (9 a.m. to 3 p.m., 3 p.m. to 9 p.m., and 12 p.m. to 6 p.m.) was used to replace the original two-shift system. In the three-shift system, one salesperson was assigned to each shift. Once the absent salesperson returned, the staffing composition for that booth resumed immediately based on our experimental design. When conducting the empirical analysis, for these three booths we exclude those days where there was a salesperson absence.

[13] Statistical tests confirm that the percentage of days assigned to each of the three conditions do not differ significantly. Details are available from the authors on request.

[14] The store management permitted us to use only 10 booths for the fixed staffing assignment in Stage 2. For these 10 booths, the salespeople were informed that the store management was piloting a new staffing assignment for 3 weeks.

[15] This resulted in the following realizations: Salesperson 1 was paired with Salesperson 4 (and, consequently, Salesperson 2 was paired with Salesperson 3) in four booths, and Salesperson 1 was paired with Salesperson 3 (and, consequently, Salesperson 2 was paired with Salesperson 4) in two booths. We verify that the empirical results are not driven by any one particular assignment. Details are available from the authors on request.

[16] Although we label $\theta_j$ the "salesperson-level fixed effect," it captures the factors that affect sales revenue not only at the individual salesperson level, but also at the booth level such as brand image, quality of products, compensation plan, etc. In other words, the salesperson-level fixed effects subsume the booth level fixed effects. Specifying the fixed effects in the model enables us to attribute the changes in sales to the changes in effort rather than to other factors, which are absorbed in the fixed effects. Moreover, since we are unable to randomly allocate booths to the IB and RS incentives or assign salespeople to different booths, the fixed effects capture the variation due to the allocation of incentives and salespeople into booths as well.

[17] Other potential control variables include the rank of each salesperson's ability across all salespeople in the 29 booths and the clothing style of each brand. Our robustness checks show that including these variables does not change the results.

[18] We also rank salespeople in each booth based on their sales performance during Stage 1 (where the staffing assignment is randomized) and find that the rankings in 22 (out of 29) booths are consistent with those based on sales performance in the 180 pretest days. In each of the remaining seven booths, only two salespeople's ranks switched (either between Salespersons 1 and 2 or between Salespersons 3 and 4).

[19] As a robustness check, we also estimate Model 1 with the standard errors clustered at the day and salesperson level, respectively. All of the main conclusions hold.

[20] This is significant at the 1% level. In the rest of this article, unless otherwise specified, all of the estimates reported are statistically significant, at least at the 10% level.

[21] The estimation results (see also the fourth and ninth columns of Table B.1) suggest that under the IB incentive, booth-level daily sales in staffing Condition C (the most homogeneous staffing condition) are the highest; while under the RS incentive, staffing Condition A (the most heterogeneous staffing condition) is the best one for the booth. It is also worth noting that the optimal staffing assignment for the firm might not necessarily be in alignment with individual salesperson's best interest. For example, Condition C yields the highest daily sales at the booth level under the IB incentive, but Salesperson 2 earns the highest commissions under Condition A. We leave the question of how to better align the firm and salespeople's interest to future research.

[22] We thank an anonymous referee for this suggestion. This result also suggests that the compensation contracts can moderate the effect of team heterogeneity on sales performance.

[23] To further examine if there are potential learning effects, we examine the booth-level daily sales of both the IB and RS booths and find that sales are quite stable across days. Details are available from the authors on request.

[24] An alternative sales output function would be $y_i = \alpha_j e_j + \alpha_{-j} e_{-j}$. Using this specification, we obtain $\partial e_j^* / \partial \alpha_{-j} = 0$, which does not find support in the empirical results.

[25] To verify the robustness of these results, we estimate two additional specifications of Model 3: (a) We classify each shift as "homogeneous" or "heterogeneous" based on the difference in ability ranks between the two salespeople in a shift. Specifically, any shift with difference in ranks lower than 1.66 (the average difference in ranks when the staffing composition is randomized) is classified as "homogeneous" (such as the shift which consists of Salespersons 2 and 3, because the difference in ability ranks is now 1), and "heterogeneous" otherwise (such as the shift which consists of Salespersons 1 and 4). (b) A continuous version of Model 3 where the difference in ability is captured by a continuous variable (based on the 180 pretest days sales). Note that in these two models, the unit of analysis becomes the average daily sales of a shift, rather than that of a booth. The estimation results of these alternative models are fully consistent with those of Model 3. Details are available from the authors on request.

## References

Bantel KA, Jackson SE (1989) Top management and innovations in banking: Does the composition of the top team make a difference? *Strategic Management J.* 10(1):107–124.

Brickson S (2000) The impact of identity orientation: Individual and organizational outcomes in demographically diverse settings. *Acad. Management Rev.* 25(1):82–101.

Caldieraro F, Coughlan AT (2009) Optimal sales force diversification and group incentive payments. *Marketing Sci.* 28(6): 1009–1026.

Chan T, Li J, Pierce L (2014) Compensation peer effects in competing sales teams. *Management Sci.* 60(8):1965–1984.

Charness G, Dufwenberg M (2006) Promises and partnership. *Econometrica* 74(6):1579–1601.

Chen H, Lim N (2013) Should managers use team-based contests? *Management Sci.* 59(12):2823–2836.

Chen H, Lim N (2017) How does team composition affect effort in contests? A theoretical and experimental analysis. *J. Marketing Res.* 54(1):44–60.

Chen H, Ham SH, Lim N (2011) Designing multi-person tournaments with asymmetric contestants: An experimental study. *Management Sci.* 57(5):864–883.

Chung DJ, Narayandas D (2017) Incentives vs. reciprocity: Insights from a field experiment. *J. Marketing Res.* 54(4):511–524.

Clark AE, Oswald AJ (1998) Comparison-concave utility and following behaviour in social and economic settings. *J. Public Econom.* 70(1):133–155.

Cunningham GB, Chelladurai P (2004) Affective reactions to cross-functional teams: The impact of size, relative performance, and common in-group identity. *Group Dynam.* 8(2):83–97.

Ely RJ, Thomas DA (2001) Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Admin. Sci. Quart.* 46(2):229–273.

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Quart. J. Econom.* 114(3):817–868.

Fischbacher U, Gächter S (2010) Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *Amer. Econom. Rev.* 100(1):541–556.

Guzzo RA, Dickson MW (1996) Teams in organizations: Recent research on performance and effectiveness. *Annual Rev. Psych.* 47(1):307–338.

Hamilton B, Nickerson J, Owan H (2003) Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *J. Political Econom.* 111(3):465–497.

Hossain T, List JA (2012) The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Sci.* 58(12):2151–2167.

Jackson SE, May KE, Whitney K (1995) Understanding the dynamics of diversity in decision-making teams. Guzzo RA, Salas E, eds. *Team Effectiveness and Decision Making in Organizations* (Jossey-Bass, San Francisco), 204–261.

Kishore S, Rao RS, Narasimhan O, John G (2013) Bonuses vs. commissions: A field study. *J. Marketing Res.* 50(3):317–333.

Knez M, Simester D (2001) Firm-wide incentives and mutual monitoring at continental airlines. *J. Labor Econom.* 19(4):743–772.

Lazear EP (1998) *Personnel Economics for Managers* (John Wiley & Sons, New York).

Lim N (2010) Social loss aversion and optimal contest design. *J. Marketing Res.* 47(4):777–787.

Lim N, Chen H (2014) When do group incentives for salespeople work? *J. Marketing Res.* 51(3):320–334.

Lim N, Ahearne MJ, Ham SH (2009) Designing sales contests: Does the prize structure matter? *J. Marketing Res.* 46(3):356–371.

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

Stewart GL (2006) A meta-analytic review of relationships between team design features and team performance. *J. Management* 32(1):29–54.

Van der Vegt, GS, Janssen O (2003) Joint impact of interdependence and group diversity on innovation. *J. Management* 29(5):729–751.

Wageman R (1995) Interdependence and group effectiveness. *Admin. Sci. Quart.* 40(1):145–180.