The Power (of) Lunch and the Role of Incentives for Fostering Productive Interactions

Jason Sandvik*
University of Utah

Richard Saouma†
Michigan State University

Nathan Seegert‡
University of Utah

Christopher Stanton§
Harvard Business School and NBER

August 6, 2018

VERY PRELIMINARY and INCOMPLETE. PLEASE DO NOT CIRCULATE.

Abstract

Despite the perceived productivity benefits associated with workplace interactions, little experimental evidence exists on the efficacy of organizational practices to induce these interactions or whether they translate into output. This paper provides firm-based field experimental evidence on incentives and non-monetary HRM processes meant to encourage effort and knowledge-sharing through employee interactions. We randomly paired over 650 sales agents, assigned them to four different treatments alongside an off-site hold-out group, and measured their individual productivity. The treatments included a tournament against two other pairs for a prize valued at approximately $100; structured meetings between partners that included a free lunch; both the tournament and structured meetings; and an internal control group limited only to a partner pairing. Productivity gains are larger and more persistent in treatments that incorporate structured meetings—underscoring the importance of HRM processes. The productivity gains for agents that received pair incentives and structured meetings are submodular, suggesting that these practices failed to complement one another when implemented together.

JEL: D8, J3, L2, M5.
Keywords: management practices; organizational design, randomized controlled trial (RCT).
Acknowledgements: We thank Karen Bernhardt-Walther, Nick Bloom, Jen Brown, Ed Lazear, Luis Rayo, John Roberts, Raffaella Sadun, and Kathryn Shaw for helpful comments. **RCT registry number:** AEARCTR-0002332.

*jason.sandvik@business.utah.edu
†rs2@msu.edu
‡nathan.seegert@business.utah.edu
§cstanton@hbs.edu

1 Introduction

Managers increasingly view fostering peer interactions as a strategic priority. This is manifest in the rising use of open office-layouts, collaboration tools, and interdependent work teams (Myers, 2015; Ibarra and Hansen, 2011; Lazear and Shaw, 2007). An alternative means of inducing peer interactions is through group incentives (Friebel et al., 2017; Bandiera et al., 2013). While much careful work has gone into studying peer effects (Mas and Moretti, 2009), identification issues typically mean that any variation used for estimation purposes is orthogonal to intentional management practices that might harness peer gains. Thus, little is known about the relative efficacy of human resource management (HRM) processes[1] and group incentives to induce productive peer interactions, how these levers interact (e.g., are they complements or substitutes), and the mechanisms that drive them (e.g., increased effort or learning). We examine these issues using a pre-registered field experiment in a large sales organization.

We find that HRM processes that direct workers to have structured peer meetings are particularly effective at increasing individual sales productivity. The channel appears to be knowledge transfers from the most able employees, and group incentives are not necessary to reinforce the gains from these meetings. The fact that productivity increased from management-induced peer interactions suggests a link between management processes and agglomeration benefits of organizing production inside firms.[2] The fact that group incentives do not appear complementarity to these processes suggests incentive crowd-out or distortions (e.g. Bénabou and Tirole (2006); Ederer and Manso (2013)).[3] A potential reason these practices may be substitutes is that they induce different margins, which is consistent with

[1]While incentives are typically classified as a type of HRM practice (e.g. Bloom and van Reenen (2011)), we explicitly differentiate between *incentives* and *HRM processes*, where the latter is any HRM practice that does not involve monetary incentives.

[2]For discussion, see Glaeser and Gottlieb (2009), Glaeser, Kallal, Scheinkman, and Shleifer (1992), Agrawal, Kapur, and McHale (2008), and Catalini (2017).

[3]There exists a large literature in Economics and Psychology investigating the crowd-out effects of monetary incentives; see Titmuss (1970), Frey and Oberholzer-Gee (1997), Ariely et al. (2009a), Ariely et al. (2009b), and Gneezy et al. (2011) and Deci et al. (1999) for a review of the Economics and Psychology literatures, respectively.

our findings that structured meetings induced knowledge spillovers while incentives induced additional sales-effort.[4]

Our setting is an inbound sales firm, which provides a particularly well-suited laboratory to identify the effects of peer interaction treatments on worker productivity. First, the sales agents within the study firm perform a homogeneous and autonomous task—they respond to sales leads through incoming customer calls.[5] Second, the fact that workers do not rely on one another to complete their work ensures that our findings are free of production externalities. In contrast, the extant literature on peer spillovers highlights tasks with effort externalities (Mas and Moretti, 2009), the need to coordinate production (Friebel, Heinz, Krüger, and Zubanov, 2017), and social spillovers that come from selecting teams of co-workers (Bandiera, Barankay, and Rasul, 2005). Third, in our setting any observed productivity gains reflect changes to the agents' individual production functions (e.g., human capital development), their incentives (including perceived incentives—e.g., peer encouragement), and their resultant sales efforts.

The experiment included 653 treated agents who work in 52 groups reporting to separate front-line sales managers. We randomly assigned agents from two of the offices into one of four treatment cells at the manager level, whereas agents in the third office (located over 600 miles away from the first two) made up an external control group (labeled the *External Control*). We randomly paired treatment-eligible agents together, with half keeping their partner for all four weeks and the other half rotating partners each week. Each day, every agent-pair had their joint-performance measure published publicly (this performance measure captured sales gains relative to the pre-experimental period). The experiment lasted four weeks, and our analysis incorporates twelve weeks of data, extending from four weeks before

[4]For further discussion of identification of complementary practices and the need for an experiment to provide causal evidence on complementarity HRM processes, see (Athey and Stern (1998); Brynjolfsson and Milgrom (2013); Milgrom and Roberts (1990); Holmstrom and Milgrom (1994); Ichniowski, Shaw, and Prennushi (1997); Atkin, Chaudhry, Chaudry, Khandelwal, and Verhoogen (2017)).

[5]Call centers, like our study firm, that make inside sales remotely represent the fastest growing segment of the 14.5 million sales jobs in the United States. See the BLS website for sales employment: https://www.bls.gov/oes/current/oes410000.htm. See Krogue (2013) in Forbes on the shift to inside sales "What is Inside sales? The Definition of Inside Sales."

the experiment began to four weeks after the experiment concluded.[6]

In the first treatment, labeled the *Internal Control*, agents learned the identity of their partner and received no other additional instructions. The second treatment, labeled *Structured-Meetings*, was designed as a process intervention where agent-pairs would meet weekly to discuss their sales techniques with one another using (structured) worksheets, culminating with a free lunch towards the end of each week. The third treatment, labeled *Pair-Incentives*, had agent-pairs participate in a weekly tournament with prizes awarded based on joint gains in performance. The fourth treatment, labeled *Combined*, included all the elements of both the *Structured-Meetings* and the *Pair-Incentives* treatments. We measure productivity during the treatment and post-treatment period via the firm's focal performance metric, individual agents' revenue-per-call (RPC). We tracked the productivity of agents in the *External Control* group, who were unaware of the experiment, to control for Hawthorne effects, seasonal demand variation, and the possibility of rank-incentive effects in the *Internal Control* group.

The primary findings are:

1. The experimental intervention uncovered practices that increased worker productivity. Agents in the three active treatment groups experienced higher sales performance during the treatment period relative to *both* the control groups and relative to their own pre-treatment performance. The increase is present in every sales measure tracked by the firm, including revenue-per-call (RPC), revenue-per-hour (RPH), and total revenue per week.

2. The managerial practices in the *Structured-Meetings* treatment were particularly effective. The treatment yielded larger productivity gains (a 24% increase in revenue-per-call) during the treatment period relative to the *Pair-Incentives* treatment (a 13% increase). The per-agent cost of implementing the *Structured-Meetings* treatment was also approximately 50% lower than that of the *Pair-Incentives* treatment, resulting in

[6]We consider only the four weeks of data after the experiment concluded to ensure our results are not contaminated by a re-organization that shifted workers to different teams and different accounts. This re-organization was announced six weeks after the experiment ended.

higher margins.[7] The return-on-investment (ROI) from the *Structured-Meetings* treatment was approximately 500% across the four weeks spanning treatment period. All other active treatments had a positive ROI during the same 4-week period.

3. At this firm, monetary incentives and managerial practices are substitutes. Agents in the *Combined* treatment had smaller gains than would be implied by the sum of the *Structured-Meetings* and *Pair-Incentives* estimates. Providing incentives for peer interactions was not a necessary pre-condition for realizing the benefits of peer spillovers.

4. Treatments involving *Structured-Meetings* induced knowledge transfers between peers, while the incentive treatment did not.

   (a) The *Structured-Meetings* and *Combined* treatments yielded *persistent* performance increases that continued into the post-treatment period following the experiment.

   (b) Agents in the *Pair-Incentives* treatment had post-treatment productivity which was statistically indistinguishable from either control group.

   (c) Heterogeneous effects by partner ability are experimentally identified and allow us to distinguish knowledge transfers from self-improvement and sentiment increases. Agents in the *Structured-Meetings* and *Combined* treatments performed better during and after the study when they were paired with stars—those with pre-treatment productivity above the median. Star agents only improved in the *Structured-Meetings* and *Combined* treatments when paired with other stars.

   (d) Survey responses from participants and anecdotes from the firm indicate that the *Structured-Meetings* treatment induced agents to share best practices, while the *Pair-Incentives* treatment did not.

[7] While the experiment had two levels of incentive compensation ($0 and a prize worth about $50 per agent), extrapolation based on a linearity assumption suggests it would have taken a tournament prize with a monetary value of about $100 per agent to replicate the *Structured-Meetings* results. This extrapolation, however, is not experimentally identified, nor does it distinguish between agent effort and learning.

We find no evidence that these results are spurious. In particular, there were no Hawthorne effects; agents in the *Internal Control* group performed similarly to those in the *External Control* group, who were unaware of the experiment. In addition, none of the observed productivity differences across treatments can be attributed to differential turnover, employee attrition, changes in scheduling, or changes to time spent selling.

The results imply that the gains across treatments were achieved via different agent responses. We provide a parsimonious framework where agents allocate effort between sales and learning. Learning contributes to human capital, which makes selling effort more productive. The *Structured-Meetings* treatment lowers the cost of learning (i.e. from one's partner) and induces agents to exert positive learning-effort, whereas the *Pair-Incentives* treatment temporarily increases the agents' return on all forms of effort, suggesting that agents may respond with more learning-effort, more sales-effort, or both. Whereas human capital is undoubtedly complementary to sales-effort, our model highlights the fact that treatments designed to spur learning and selling need not be complementary, especially when the cumulative effects of sales-effort and learning-effort are concave (e.g., when agents face decreasing marginal returns to effort).

Education is the other obvious context for work on peer spillovers, and the insights from this literature suggest experimental variation is needed to understand tradeoffs from different policies. Garlick (2014) focuses on tracking policies versus random assignment to improve educational outcomes, and Carrell et al. (2013) evaluate policies to pair heterogeneous students together. In Carrell et al. (2013), like here, results from different practices and optimal policies could not be gleaned from observational data and thus required a policy change.[8]

From production environments, much of the work on interactions focuses on business-to-business relationships across firm and their associated benefits (Cai and Szeidl, 2016; Hasan

[8]For examples of other RCTs that uncover evidence invisible to observational data, see Nagin et al. (2002); Bertrand and Mullainathan (2004); Fehr and Goette (2007). Levitt and List (2009) claim that stronger inference can be achieved from field experiments than from uncontrolled data and data gathered in lab experiments. By directly and randomly constructing control groups, we create a sharp counterfactual against which to estimate treatment effects (Harrison and List, 2004) and distinguish underlying mechanisms (List and Rasul, 2011).

and Koning, 2017; Catalini, 2017) or on compensation schemes for groups of workers within the firm[9] Furthermore, little prior evidence from production environments exists on how to bundle practices to capture spillovers from peers.[10] This is one of the first experimental studies to consider the mix of different incentive and HRM processes. Our findings align with the growing body of evidence on the importance of management (Bloom et al., 2016, 2013; Bloom and van Reenen, 2011).

Given the magnitude of our empirical findings, why were these practices not attempted earlier? First, the outcomes were not obvious to the firm's management (nor to the authors). In particular, many leaders on the sales teams believed that incentive-based interventions would be sufficient to induce knowledge transfers, while operations and human resource managers wanted to test the efficacy of non-monetary practices. Second, given the initial uncertainty, earlier experimentation was under-powered due to coordination difficulties across division heads. Third, the large gains here mirror organizational interventions in other contexts. For example, akin to the worksheets used in our *Structured-Meetings* treatment, Lingard et al. (2008) required hospital teams to verbally discuss preoperative checklists, which led to improved patient outcomes, in part by surfacing knowledge gaps across the team and creating a safe-space in which the workers could teach one another.[11] Similarly, Jackson and Schneider (2015) find large productivity gains through the introduction of checklists in auto repair shops, but the gains in that context appear to come from better monitoring rather than knowledge enhancement. In these other settings, it was not obvious to management that small process interventions would have large effects on productivity.

Field experiments are a useful tool to implement and test new institutions before at-

[9]For instance, Bandiera et al. (2013) study how different incentives schemes impact team productivity across agricultural workers. Another careful study is Friebel et al. (2017), where team bonuses are shown to increase the team output in a chain of retail bakeries. Finally, Chan et al. (2014) examine how team incentives influence peer production in a department store setting.

[10] The closest work from a production environment is on the persistent effects of bosses on their subordinates (Lazear et al., 2015), though their focus is on front-line managers as opposed to our emphasis on co-workers.

[11]Psychological safety—or the lack thereof—is often cited as a determinant of team performance, see (Edmondson and Lei, 2014). Process interventions like ours may also prompt individuals to notice–and address–gaps (Hanna et al., 2014).

tempting firm-wide adoption (Carpenter et al., 2005). Based on the favorable outcomes of our experiment, the firm's management has replaced their traditional on-boarding process with a mentoring program to train all newly hired sales agents and to nudge their seasoned agents (mentors) to become more conscious of their own work.[12] The firm learned from this experience and now has staff members who are tasked with facilitating experimentation.

The paper proceeds as follows. Section 2, describes the experimental setting and provides details regarding the study firm, the sales agents, and the organizational culture. Section 3 introduces the experimental design and describes the different treatment groups. In Section 4 we derive a parsimonious model to develop hypotheses of how agents will respond within each treatment. Section 5 reports our findings, and Section 6 concludes.

2 Experimental Setting

2.1 The Study Firm

The experiment took place in an inbound-sales call-center from July to August of 2017. The firm employs over 730 sales agents across three geographically separate offices,[13] contracting with television, phone, and internet providers to market and sell goods and services. Sales agents field inbound calls from potential customers and are tasked with determining customer needs and explaining the benefits of premium service packages (up-selling) when appropriate. Third-party sales contracting is a common practice in the United States, especially for nation-wide service providers.

The sales department of the firm consists of six large sales divisions and several smaller sales divisions. Sales divisions are headed by one or two division presidents and are uniquely characterized by the bundles of products, services, and brands offered for sale.[14] Divisions

[12]The firm's positive reaction to the experimental intervention is similar to that of the study firm in Bloom et al. (2015).

[13]The two offices involved in the experiment are within 50 miles of one another, whereas the third office is located over 600 miles away.

[14]For example, one division might only sell internet packages from provider A, while another might sell

are comprised of multiple groups consisting of approximately ten to fifteen sales agents led by a single sales manager. The 653 workers eligible for treatment included all sales agents from the six large sales divisions in the firm's two largest offices. Workers in the third location, 83, were not eligible for assignment to a treatment group, constituting a hold-out control group. Because there is minimal interaction between workers at different locations and because the hold-out location is geographically distant, workers at the third location were unaware of the experiment. Consistent with this lack of awareness, we find no break in productivity for these agents around the experiment.

Sales agents spend about 80% of their workday on the phone or waiting to field a call. When a call arrives, it is assigned to the next available agent in the queue. Only a small percentage, 3%, of calls are outbound, and most of these arise from agents following up with existing customers or returning a dropped call. When not receiving phone calls, agents participate in group- and division-wide meetings as well as in one-on-one discussions with managers. Most agents, 87%, work full-time schedules.[15] The sales floor is predominately male, 68%, and is relatively young, with 50% of sales agents under the age of 24. In a given week, the average agent takes 62 calls, approximately two calls for every hour available to answer the phone.

2.2 Sales Agent Training and Development Practices

When first hired, agents are enrolled in a formal, on-site sales training regimen for two-weeks. Throughout training, agents are paid to absorb information delivered largely through lectures and by listening in on other agents' phone calls. The trainees then spend up to four weeks in "integration," taking calls under the supervision of a temporary "integration manager." This manager familiarizes agents with the process of selling and with the specific products and services available for sale. Once trainees reach a threshold level of revenues under the

internet packages from provider B *and* satellite television packages.

[15]The threshold for full-time employment at the firm is 32 hours per week. The maximum number of weekly hours observed in our data is 46.

integration-manager, they join a permanent group on the sales floor where they continue to sell the same products and services on which they were trained. Sales agents that do not reach the threshold levels of performance within a designated number of weeks after training are usually let go.

When asked about what separates average sales agents from top performers, one of the division presidents explained, "top performers... have a solid understanding of product and [services] knowledge. More importantly they are the most consistent with following their sales process." The division president remarked that, while "soft skills" are important for building rapport with customers, the most important determinant of sales success is a comprehensive knowledge of the sales protocols and sales management systems. It is, thus, vital that agents efficiently acquire this combination of (largely) firm-specific knowledge, be it from management or from their coworkers.

2.3 Sales Agent and Manager Compensation Structure

Sales agents are compensated in three ways: (1) they receive an hourly wage for the time they are at work (and not on their scheduled breaks). Sales agents start at an hourly wage of approximately 150% of minimum wage and receive small hourly raises for every three months of tenure, with their hourly rate capped at approximately 200% of minimum wage; (2) they receive a weekly commission[16] based on multiple performance measures: the revenue they generate from sales,[17] their selling efficiency relative to their peers, and the audited quality of their customer service; and (3) they are able to receive bonuses from temporary promotional sales activities.

While the firm's leadership hierarchy is structured to facilitate ongoing agent development, sales managers' incentives do not necessarily reward them for allocating energy to

[16]The average (median) sales agent earns $217.78 ($185.45) per week in commissions.

[17]Partners pay the firm for every sale in accordance with pre-negotiated schedules—some of which vary with the total number of products or services sold by the firm. To insulate agents from the uncertainty surrounding aggregate sales and periodic contractual negotiations, the firm posts relatively fixed "transfer prices" that form the base revenue upon which agents are paid commissions. All use of the term "revenues" in this paper refers to sales priced in accordance with the internal transfer price schedule.

train under-performing agents. Managers are compensated based on their group's aggregate revenues, and—consistent with Bandiera et al. (2007)—our interviews suggest that these contracts induce managers to focus their efforts on wringing out performance from high-ability workers rather than developing less-able workers. Executives at the firm had previously observed such behavior during promotional periods (e.g. short-run contests between groups reporting to different managers), yet we were told the managerial incentive system in place is too coarse to target managerial attention into developing low-skilled workers.

2.4 Pre-Experiment Collaboration Among Agents

While management readily acknowledged the possible value of interactions amongst sales agents, two features of their imposed compensation system discouraged a culture of peer-learning or knowledge-sharing. First, peer-learning requires inter-agent communication, yet any time away from the phone results in fewer revenue generating opportunities for an agent which directly decreases their take-home pay (total commissions). Second, agents' commission rates—i.e., the fraction of their earned revenue paid out as commission—is a weakly decreasing function of their co-workers' success.[18] Accordingly, we did not observe significant or regular channels for knowledge-sharing between agents prior to the experiment.

2.5 Collaborating with the Firm

Call-centers—notably, inbound-sales call-centers—tend to have high rates of agent turnover.[19] The study firm's management attributed much of the difficultly in retaining sales agents to the variation in performance across workers. Specifically, managers explained that many of their applicants are drawn to the job because of the relatively large earning potential

[18] Commission rates are bucketed into four to five coarse categories that depend on relative performance on revenue-per-call and revenue-per-hour. For the same level of sales revenue, moving from the bottom bucket to the top bucket changes take-home commissions by about ten percent. Employees are fully aware of the incentive structure, but pre-experiment interviews revealed that employees would be willing to collaborate with others if encouraged to do so. This is likely because the probability of changing categories after helping one other person is small. In particular, agents described these categories as if they are relatively fixed.

[19] Industry average turnover is approximately 25% annually (Holman et al., 2007).

(if an agent is productive). We were told that some of their best-performing agents were also the slowest starters, and the firm felt that very capable agents often leave because they were unable to come up to speed "in time." Accordingly, management agreed to experiment with processes and incentives that would contribute to agent's (task-specific) human capital development.

Our collaboration began in June 2016, where we conducted interviews with sales managers and agents. During this time, we built trust with the organization and its leaders by sharing data analysis around significant organizational changes that occurred later that year. The experiment here post-dates those changes, occurring at an otherwise stable time for the firm.[20]

3 Experimental Design

The design was pre-registered before beginning the experiment.[21] The study was conducted in two neighboring call-center offices. We began by randomly assigning the 52 sales agent groups (based on the identity of a sales manager) into one of four treatment groups. Once groups were assigned to a treatment, agents within each treatment were randomly paired so long as they were in the same division and location and had overlapping schedules. The treatments are depicted in the four cells of Figure A (details follow).

	No Meetings Prompted	Meetings and Lunch
No Pair-Incentives	*Internal Control* Group	*Structured-Meetings* Treatment
Tournament Against Other Pairs	*Pair-Incentives* Treatment	*Combined* Treatment

Figure A: Treatment Assignment Matrix for Agents in Active-Treatment Eligible Locations

[20]The *Pair-Incentives* treatment resulted in a larger sales response with respect to incentives than found in our prior work with the firm in Sandvik et al. (2018). We discuss the relationship between these estimates in Section 6.4.

[21]The RCT registry number is AEARCTR-0002332. The IRB at the University of Utah approved the design with approval number IRB 00098156.

3.1 Structured-Meetings Treatment

Pairs in the *Structured-Meetings* treatment were encouraged to complete the following tasks: (1) fill out an individual self-reflective worksheet to prompt discussion prior to meeting with their partner;[22] (2) converse with their partner and record their partner's self-reflective responses on the backside of their own worksheet; and (3) turn in these worksheets to management by Wednesday each week. Completion of these tasks was optional but motivated by the receipt of a free catered lunch later in the week (both agents had to be present to receive lunch). During this lunch, agents were provided with high-end, local sandwiches[23] and were asked to discuss several talking-points related to the worksheets while they ate with their partner.[24]

It is important to note that while the *Structured-Meetings* treatment was largely self-guided: (1) it was a process, and (2) the process was directed. The distinction is important because relative to the *Peer-Incentives* treatment discussed below, the *Structured-Meetings* treatment required non-trivial, managerial attention; e.g. creating the worksheets, monitoring worksheet completion, answering questions related to the process, coordinating and administrating lunch, etc.

3.2 Pair-Incentives Treatment

Agents in the *Pair-Incentives* treatment were told that they were competing each week against two other pairs of agents for rewards valued at approximately \$50 per person.[25] The company indicated that tangible prizes rather than monetary payments, historically, were found to better motivate agents in individual contests. Management attributed the

[22] Example question: "Think about the least successful sales call you've had in the last week. How could you have done better?" These worksheets were (*ex-ante*) viewed as a necessary step to make the ensuing conversations more salient; points of emphasis on the worksheets were sourced from the firm's leadership to maximize the expected gains from focusing agents' attention. Documentation of this worksheet can be found in Appendix B.

[23] Price for the sandwiches was about \$7.

[24] Documentation of these talking-points can also be found in Appendix B.

[25] Competitors were chosen randomly at the end of the week so as to minimize the identity effects highlighted in Akerlof and Kranton (2000).

effectiveness of prizes to their immediacy, making them salient, rather than a monetary transfer that appears in a paycheck two weeks after the reward date. As a result, prizes for the *Pair-Incentives* treatment included rounds of golf, on-site massages, or tickets to extra-curricular activities. In the surveys, agents reported an average valuations for the prizes of \$40.20.[26] A \$40 incentive equates to an 18% increase in weekly commission pay for the average agent and a 22% increase for the median agent.

The tournament was designed to give each pair similar *ex-ante* odds of winning. In addition to randomly drawing new competitors each week, the tournament was scored according to each pair's change in productivity *relative to the pre-treatment baseline*. Further details are provided in Sections 3.5 and 3.7. The expected prize values were equal to over 8% of the median agent's take-home compensation. Far weaker group incentives, albeit in a setting with complementarity among workers, have been found to generate meaningful productivity increases (Friebel et al., 2017). Accordingly, we expected agents in the *Pair-Incentives* treatment to, at least, outperform agents in the control groups over the course of the study. While agents in the *Pair-Incentives* treatment were not explicitly encouraged to communicate—i.e., transfer knowledge—with their partners, they were free (and able) to do so.

3.3 Combined Treatment

Pairs in the *Combined* treatment were entered into identical, albeit separate, tournaments as those in the *Pair-Incentives* treatment[27] *and* they were also prompted to engage in the exact same type of structured conversations (the sharing of weekly self-reflective worksheet responses culminating with a catered lunch) as the agents in the *Structured-Meetings* treatment.

[26] See Table A.1.

[27] Pairs from different treatments never competed against one another in the tournaments.

3.4 Internal Control Group

Agents in the *Internal Control* group learned they were paired with a partner and had their joint change in productivity score published publicly, but they did not receive: worksheets, guided directions to meet, nor any other monetary incentives. Note, however, that this group is an *Internal Control* where the baseline comparison is agents who have knowledge of a pairing but no formal incentives to act on that pairing. When designing the experiment, we expected rank incentives to be minimal, but the design does, in principle, allow us to test for the effect of rank incentives.[28] The *External Control* group, or hold-out group, allows a comparison of each of these cells relative to agents with no pairings, knowledge of the experiment, nor prompts around collaboration. In practice, the *Internal Control* group tracked the *External Control* group throughout the experimental period.

3.5 Allocation of Agents to Treatments

The experiment is a clustered design, with agents allocated into one of the four cells in Figure A based on the identity of their sales manager. Agents were then randomly paired together so long as they worked in the same location, in the same division, had overlapping schedules, and were in groups that were designated to the same treatment. Pairings were assigned and publicly announced at the beginning of each week.

All three active interventions and the *Internal Control* group shared four common traits. First, all agents were randomly paired with a single partner within the same location, division, and treatment. Second, pairs were randomly designated as stable-pairs or rotating-pairs such that roughly half of all agents had a single partner throughout the entire four-week intervention, and the other half were randomly paired with a new partner each week (repeat

[28] While other studies have found that the introduction of public rank data (sometimes called rank incentives) may cause deviations from prior productivity (e.g. Bandiera et al. (2013)), rank incentives for individual agents were already present at this firm because commission rates partially depend on relative–albeit private–comparisons of agents. According to the contingency results in Blader et al. (2016), rank displays were in-line with prior practices and therefore may have had minimal effects relative to baseline.

assignments permitted).[29] Third, all pairs had their joint performance scores published daily on TV monitors and on their internal messaging platform.[30] The joint performance score normalized the *change* in the pair's average revenue-per-call (RPC) relative to their RPC in the four weeks immediately preceding the experiment, allowing for inter-divisional comparisons.[31] Finally, all agents in these four cells were notified that their, and their partners' individual productivity was being shared with a team of university researchers.

3.6 Pre-Experiment Homogeneity Across Treatments

Table 1 describes the demographic makeup and pre-treatment performance of the participating sales agents. These summary statistics highlight the merits of randomly allocating agents to treatments, as all treatment group characteristics and performance averages in the pre-experimental period are *not* statistically different from one another in these four cells. P-values of randomization tests of mean differences in the *Internal Control* and active treatment columns are reported. These tests are computed as the joint-hypothesis test of equality of treatment groups from a regression of the variable of interest on treatment assignment dummies after clustering standard errors based on manager identity (the level of assignment). The agents in the three active treatment groups appear to be quite homogeneous with each other and to those agents in the *Internal Control* group. Agents in the *External Control* group are, on average, less productive. However, we test for and find common pre-period trends in productivity between the agents assigned to different treatment arms and agents

[29] As new agents entered the sample (newly trained agents or agents moving in from other divisions), they too were randomly paired with a coworker and the pair was randomly designated to remain unchanged or to be repaired in subsequent weeks. Some pairs were dissolved as one or both agents left the sample (e.g., termination of employment, moved to a different division, took a leave of absence, etc.); the partners of these departing agents were paired with a new, randomly chosen, partner.

[30] A benefit of publicizing scores on TV monitors is that it kept agents aware of the tournament. However, with dynamic feedback about performance, it is possible some incentivized agents might quit or give up (Brown, 2011). As a result, we randomly assigned three pairs to each tournament at the end of each week, meaning all but the bottom two pairs had a chance to win. In addition, agents' individual commissions on each sale limited the attraction of completely minimizing effort. In later analyses we find minimal evidence that agents changed their effort based on past tournament performance.

[31] Management advised us to avoid displaying negative scores, hence scores were normalized around 100, where 100 reflected the pre-intervention productivity. Employees who joined the sales floor during the experiment were tagged with the median pre-treatment RPC.

in the *External Control* group, giving rise to a difference-in-differences estimator.

3.7 Implementing the Experiment

To communicate both the disclosures and the intervention guidelines, we did the following: (1) we solicited the help of senior executives who shared the details with division and sales managers; (2) we had a designer create and print posters which we placed in the firm's common areas and on their internal TV monitors in the weeks leading up to and during the experiment; (3) we set up an e-mail and phone hot-line to answer questions; (4) we set up a website that explained all aspects of the initiative including daily scores and frequently emailed questions; and (5) a subset of the authors were physically on-site at least three days a week at both locations to answer questions, distribute worksheets, and administer the catered lunches. Table A.2 in Appendix A reports agents' answers to several survey questions that suggest participants (1) knew about the treatments, (2) completed the worksheets meant to facilitate knowledge transfer, (3) interacted with their partners in a meaningful way, and (4) valued the rewards.

4 Theory Development

4.1 Model Setup

We provide a parsimonious agency model to isolate predictions and to illustrate different margins of adjustment. Sales agents are assumed to have task-specific human capital (e.g., selling skills) $\theta \geq 0$ (hereafter referred to simply as *human capital*). The agent combines sales effort, e, with human capital to convert calls into revenue, $Y = f(e\theta) + \varepsilon$ where $f(\cdot)$ is a monotonically increasing, concave function, ε is idiosyncratic noise that prevents the firm from contracting directly on effort, and $e\theta$ is the "aggregate input" put forth by the agent.

To set ideas, we define f as follows:

$$f(x) = \begin{cases} 0 & 0 \leq x < 1 \\ n \quad 2^{n-1} \leq x < 2^n \quad n \in \mathcal{N} & x \geq 1 \end{cases}$$

The step-function $f(x)$ is concave and everywhere increasing, with constant step-sizes although the length of each step is itself increasing with x. We use the step-function f because it reflects several features found in sales, and more generally, in professional environments: (1) the outcomes / output (e.g. services being sold) are discrete, (2) a modicum of sales-effort and human capital are required to attain even the smallest output or sale (since $f(x) > 0$ if and only if $e\theta \geq 1$), and most importantly, (3) reaching successively higher levels of output (up-selling) requires increasing levels of marginal effort.

The agent can increase his human capital in the present and (un-modeled) future periods should he exert learning effort, r. In particular, his effective human capital in period t (e.g. the treatment period) is given by his existing human capital plus his period t learning-effort: $\theta_t = \theta_{t-1} + r$. That is, we assume that learning-effort deterministically yields additional human capital–provided the agent is paired with another agent who has similar or relatively more human capital. In particular, for peer-learning to be feasible we assume that agent i's partner, say agent j, must have human capital θ_j where $\theta_{j,t-1} - \epsilon \geq \theta_{i,t-1}$ for some fixed $\epsilon > 0$. Accordingly, expected revenue increases when an (eligible) agent exerts learning-effort and increases his or her human capital, when an agent increases sales effort, or both–albeit the discrete benefits to additional aggregate input are marginally decreasing due to the concavity of f. To keep the exposition as simple as possible, we assume that the agent faces a binary learning decision; i.e. $r = 0$ and the agent incurs no learning costs, or $r = \delta$ and the agent incurs a cost of $w > 0$. We take the pre-existing (baseline) human capital of agents as given to concentrate on comparative statics. We assume that the agent's sales-effort, e, is continuous with cost $c(e) = \frac{e^2}{2}$, and that there are no costs associated with the dissemination

of knowledge. That is, agents can "teach" their partners without incurring any costs unless they themselves learn from their partner as well, in which case the cost w is incurred by both agents.

In exchange for generating a revenue of Y_i, agent i receives a commission rate, $B > 0$, in addition to a wage $\alpha > 0$. For a given learning effort, r_i, the agent selects a "revenue target," i.e. he selects among the minimum aggregate input levels needed to attain one of the discrete revenue levels, and we let $n_i^*(\theta_{i,t-1})$ denote the particular revenue level that maximizes the agent's payoff: $B(n) - \frac{1}{2}\left(\frac{2^{(n-1)}}{(\theta_{i,t-1}+r_i)}\right)^2$, implying that the agent's optimal sales-effort is given by $e_i^*(\theta_{i,t-1}) = \frac{2^{(n_i^*(\theta_{i,t-1})-1)}}{(\theta_{i,t-1}+r_i)}$. The following lemma characterizes the agent's choice of target, $n_i^*(\theta_{i,t-1})$.

Lemma 4.1. *Agent i will choose target $n+1$ over n if and only if:*

$$\sqrt{B}(\theta_{i,t-1} + r_i)^2 \geq 2^{\frac{2n-3}{2}}\sqrt{3}. \tag{1}$$

Continued application of the lemma shows that $n_i^*(\theta_{i,t-1})$ is the largest natural number, n, such that (1) holds.

Mapping treatments to the model, we assume that agents experience team-tournament winnings and any associated peer encouragement in the *Pair-Incentives* treatment through an increased commission rate, B.[32] Turning to the other active treatment, the *Structured-Meetings* treatment facilitates interactions and hence lowers to cost of peer-learning, which we model as a reduction in the cost of learning, w.

Proposition 1. *If structured-meetings increase revenues, then the agent must have exerted positive learning-effort, $r_i^* = \delta$, whereas an increase in revenues in the* Pair-Incentives *treatment could be the result of positive learning-effort and less sales-effort, no learning-effort and additional sales-effort, or more learning- and more sales-effort.*

[32]This follows Bandiera et al. (2013), who make the same reduced form assumption to capture group tournament incentives.

The fact that these treatments may operate through different mechanisms is important for understanding potential persistent gains, heterogeneous effects, and whether these treatments are complements or substitutes. The agent only increases effort if doing so allows them to reach a higher revenue target. Given the concavity of the revenue function, f, the marginal effort necessary to reach each additional hurdle is itself increasing, therefore while learning-effort and sales-effort are clearly complementary, the *effects* of the two active treatments–each encouraging different combinations of effort–need not be complements.

Proposition 2. *Pair-incentives and structured meetings can be substitutes or complements such that the productivity of a group receiving both may fall short of–or exceed–the sum of productivity gains across the two individual treatments respectively. The treatments are more likely to complement one another with agents have low-levels of (pre-existing) human capital.*

We highlight the tension between complementary efforts and concave output with an example illustrated in Figure B. Consider an agent with $\theta_i = 1$, where $B = 5$,[33] $w = 3$, and $\delta = \frac{1}{5}$. With these values, an agent who is eligible to learn from their partner (i.e. their partner has $\theta_j > \theta_i$) would not exert learning effort, as he targets $n_i^* = 1$ without learning and receives a payoff of \$3, whereas he would choose a target of $n_i = 2$ if he had exerted learning effort, albeit his payoff would be only \$1.44. Suppose now that the *Pair-Incentives* treatment involves a B equivalent value of 12, and the structured-meetings effectively reduce the cost of learning to \$1. With these treatment values, all actively treated agents (including those in the *Combined* treatment) target the same revenue level, $n_i^* = 2$, and all agents exert positive learning-effort, except those in the *Pair-Incentives* treatment, who alternatively attain the higher revenues via higher sales-effort alone. In this example, the gains to combining treatments is nil, as the aggregate input needed to attain a revenue of $n = 3$ requires significantly more efforts than the complementarity between sales-effort and learning-effort afforded to the agent.

[33]The commissions are paid on the basis of internally set prices, therefore they may exceed 1 without "selling the firm to the agent."

Target Revenues:	$n=1$	$n=2$	$n=3$
Control Group			
Learning	$0.61	$1.44	-$10.22
No learning	$3.00	$2.00	-$17.00
Structured-Meetings Treatment			
Learning	$2.61	$3.44	-$8.22
No learning	$3.00	$2.00	-$17.00
Pair-Incentives Treatment			
Learning	$7.61	$15.44	$10.78
No learning	$10.00	$16.00	$4.00
Combined Treatment			
Learning	$9.61	$17.44	$12.78
No learning	$10.00	$16.00	$4.00

Figure B: Optimal choice of learning-effort and target revenue. The baseline setting has $B = 5$, $w = 3$, $\theta_i = 1$, and $\delta = 0.2$, whereas the *Pair-Incentives* treatment increases the agent's incentives to an equivalent value of $B' = 12$, the *Structured-Meetings* treatment results in a learning cost of $w' = 1$, and the *Combined* treatment imposes both modifications. The agent's preferred choice in each treatment is highlighted in green.

From (1), the treatments can only complement one another if: (1) learning takes place in only the *structured-Meetings* and *Combined* treatments, (2) the *Pair-Incentives* treatment is sufficiently large (i.e. $\sqrt{B'} > \sqrt{B}$), and the level of relative learning $\frac{\delta}{\theta+\delta}$ is sufficiently large. Holding fixed the level of learning that takes place with peer-learning, δ, agents with relatively higher levels of human capital are less likely to yield complementary productivity gains because the increase in commission needed in the *Combined* and *Pair-Incentive* treatments is larger for these agents relative to the change needed for agents with relatively lower levels of pre-existing human capital.

Our example and Proposition 2 establish two important aspects of the treatments under study. First, we cannot *ex-ante* predict whether treatments targeting complementary efforts are themselves complements or substitutes, which in turn motives our experimental design. Second, these treatments are more likely to be complements when agents have relatively low-levels of human capital. This second aspect motivates a deeper study of the mechanisms underlying our two treatments, in particular, which of these treatments–if any–are inducing agents to increase their human capital. We address these questions in the following section by considering persistent gains, heterogeneous effects, and survey evidence.

5 Results

The presentation of results initially focuses on the firm's main performance measure, revenue-per-call (RPC). RPC is a natural focal performance measure for the firm because it is roughly comparable across agents who face different demand conditions (number of calls), filtering noise outside of agents' control. The RPC measure also aligns revenue with per-call customer acquisition costs used internally for accounting purposes. The firm did not change their customer acquisition strategy at any point throughout the entire 12-week data collection window, therefore relative RPC increases among treated agents represents incremental profits net of treatment costs (discussed below). We later evaluate alternate measures including revenue-per-hour (RPH), revenue-per-agent-week (Revenue), and agent turnover.[34] All revenue measures provide qualitatively similar results, while turnover patterns are similar across agents eligible for the active treatments.

Average RPC increased significantly in the active treatment groups relative to both controls. Figure 1 displays RPC (normalized to the grand-mean in the pre-treatment period) by treatment group in the 4-week pre-experimental period and the 4-week treatment period.[35] Average RPC in the *Pair-Incentives* treatment increased to about $71 from a pre-experiment baseline of $61. The treatment period increase was larger for the *Structured-Meetings* and *Combined treatments*, totaling about $77 in revenue-per-call. These changes represent sizeable percentage increases. In both the *Internal* and *External Control* groups, the change in mean RPC during the treatment period relative to the pre-period was approximately zero.

Figure 2, which plots the mean RPC by treatment group over time, shows parallel productivity trends in the pre-experimental period (weeks -7 through 0). This enables a difference-in-differences estimator for formal comparison of treatments. Every group in the figure has productivity normalized in the week prior to the experimental period (which begins in Figure

[34] Revenue-per-hour aligns revenue with the labor cost of staffing agents, while total revenue accounts for changes in hours worked or hours off the phone.

[35] The normalization is for consistency between Figure 1 and Figure 2. Table 1 presents the non-normalized, average log RPC for each group in the pre-treatment period.

2 at week 1). The figure shows that neither the *Internal* nor *External Control* groups reacted to the onset of the active treatments. We later return to Figure 2 to analyze persistence in productivity increases across the different treatments.

5.1 Productivity During the Experiment

Table 2 presents analysis of percentage changes in RPC and formal statistical tests of differences between treatment effects.[36] The sample contains the four weeks of data prior to the experiment and the four weeks of data concurrent with treatments. The estimating equation is:

$$\begin{aligned}\log(\text{RPC})_{i,t} =& \beta_0 + \beta_1 \text{Structured-Meetings}_i \text{ x Treatment-Period}_t \\ &+ \beta_2 \text{Pair-Incentives}_i \text{ x Treatment-Period}_t \\ &+ \beta_3 \text{Combined}_i \text{ x Treatment-Period}_t + \lambda_t + \theta_g + \varepsilon_{i,t}\end{aligned} \tag{2}$$

where i represents an agent, t represents week, g represents sales manager group, λ_t and θ_g are week and sales manager fixed effects, respectively, and $\varepsilon_{i,t}$ is an idiosyncratic error term. The variables Structured-Meetings x Treatment-Period, Pair-Incentives x Treatment-Period, and Combined x Treatment-Period are set to 1 during the experimental period for those randomly assigned to the treatment and are zero for all agents otherwise. The level effects of the treatment are subsumed by the sales manager fixed effects because all agents reporting to a sales manager are assigned to the same treatment.[37] Week fixed effects remove common time shocks that affect all treatments. Specifications in Columns (1) to (3) are relative to the *Internal Control*; Column (4) omits the *Internal Control* group while adding the *External control* as the baseline; Column (5) includes the *Internal Control* and tests for changes in

[36] Table A.3 reports similar results using log revenue-per-hour as the dependent variable. Analysis of total revenue per agent-week is in Table A.5. Across different measures, results are similar.

[37] The standard errors are clustered at the sales manager level to allow for arbitrary correlation within a sales manager group due to the clustered assignment rule. Standard errors are clustered at the sales manager group level in all subsequent specifications and, for brevity, we do not continue to reiterate this in the text.

the *Internal Control* relative to the *External Control*.

Agents in the active treatment groups had large increases in productivity. Table 2 indicates that agent-pairs provided with the worksheet exercise and catered lunch in the *Structured-Meetings* treatment increased performance by about 24% relative to either control group. Agent pairs who competed in weekly tournaments for individual prizes against other pairs in the *Pair-Incentives* treatment increased performance by about 13 - 14%. The difference between the productivity gains in the *Structured-Meetings* and *Pair-Incentives* treatments is statistically significant, as indicated by the p-value from the Wald test of equality in the penultimate row. As expected in a randomized experiment, the addition of agent or partner demographic controls across these columns minimally changes the estimates. The results are invariant to the choice of control group.

An important result from Column (5) is that the *Internal Control* group (with a partner pairing and a joint productivity score shared publicly) performed similarly to the (off-site) *External Control* group that was unaware of the experiment. This suggests the productivity increase is not due to Hawthorne effects or the monitoring of employee productivity. The difference between the *Internal Control* and the three active treatments shows that merely displaying ranks was not sufficient to improve performance, as agents in the *Internal Control* were aware of the experiment but had no changes in productivity.

Estimates from the experimental period underscore the value of guiding agent interactions (e.g. in the *Structured-Meetings* treatment), as the improvement in productivity with the *Structured-Meetings* treatment was larger than the increase from the *Pair-Incentives* treatment that lacked management direction. The relative cost of these two treatments combined with the productivity gains is decision relevant for firm profits. The *Structured-Meetings* treatment encouraged agents to fill out the worksheet in exchange for a \$7 catered lunch (\$14 total). In comparison, the *Pair-Incentives* treatment relied on prizes costing approximately \$50 (valued by the agents at approximately \$40), which were awarded to a third of the participating agents. The cost per six agents was \$42 in the *Structured-Meetings* treatment

and \$100 in the *Pair-Incentives* treatment. Nonetheless, the *Structured-Meetings* treatment induced productivity gains that were 10 percentage points higher than the *Pair-Incentives* treatment while costing only 40% as much.[38]

The differences in productivity gains across the active treatments provide several insights. First, the increase in performance of the agents in the *Pairs-Incentives* treatment is consistent with previous findings on the effectiveness of group-based incentives to increase effort exertion and productivity (Friebel et al., 2017; Bandiera et al., 2013). Second, the *Structured-Meetings* intervention, meant to test the effectiveness of a management-directed HRM process (Bloom and Van Reenen, 2011), resulted in larger productivity increases than the *Pairs-Incentives* treatment alone. Third, the treatment effect estimates are not due to differential turnover (discussed in Section 6.2) and are similar when including agent fixed effects (Table A.6). Finally, the similarity between the *Structured-Meetings* and *Combined* agents' average productivity gains suggests these practices are substitutes. The next section details tests for complementarity and implications of the outcome.

5.2 Structured-Meetings and Pair-Incentives: Complements or Substitutes?

The model developed in the previous section yields ambiguous predictions regarding whether the *Pair-Incentives* and *Structured-Meetings* treatments would be complements or substitutes. When practices are complements, giving agents both treatments together results in supermodular gains: the *Combined* treatment effect will be greater than the sum of the *Structured-Meetings* and *Pair-Incentives* treatments. Substitute practices means the gains will be submodular: the *Combined* treatment effect will be less than the sum of the underly-

[38] A different back-of-the-envelope calculation asks what the *Pair-Incentives* would have to be to induce the same productivity increase as the *Structured-Meetings* group if the treatment effect was linear. Accordingly, if a \$50 incentive resulted in a 14% increase in RPC relative to the control group, then to achieve the same 25% increase realized by agents in the *Structured-Meetings* treatment, the incentives would have to be approximately \$86 $\approx$ (.24 $\times$ \$50)/.14 per agent. We postpone a full return-on-investment discussion to section 6.5.

ing component treatments. The last line in Table 2 presents the p-value from a Wald test of the null hypothesis that the *Combined* treatment effect is larger than the sum of the *Pair-Incentives* and *Structured-Meetings* treatments. The hypotheses that the *Pair-Incentives* and *Structured-Meetings* treatments are complements is rejected in the data, indicating the practices are substitutes.

The fact that incentives failed to reinforce the peer-learning associated with structured-meetings was not obvious[39], underscoring the benefits of experiments with managerial practices to consider the portfolio of incentives and processes. Because the *Combined* treatment did have somewhat higher performance than the *Structured-Meetings* treatment, the optimal mix of practices may include high-powered incentives. Whether pair-incentives should be added to the *Structured-Meetings* treatment requires an assessment of the marginal benefit of the incentives on top of the structured-meetings. Table 2 reveals that adding pair-incentives to the structured-meetings did not yield a statistically distinguishable difference from the structured-meetings alone. Evaluating the difference in productivity at the point estimates, while recognizing the sampling variance associated with these estimates, the *Combined* treatment increased productivity by about 1% per call, which would have resulted in about $40.00 per week in additional revenue for each agent. Given that the per-agent cost of the pair-incentives was about $17 per week, we cannot reject that the incentives in the *Combined* treatment were optimal.

A comparison of the productivity changes across the post-treatment and treatment periods sheds light on why the *Pair-Incentives* and *Structured Meetings* treatments are substitutes. Theory suggests that productivity gains resulting from peer-learning ought to persist beyond the treatment window, though such gains may potentially come at the cost of lower sales-effort during treatment. Unfortunately, treatment-period productivity changes com-

[39]Incentives may be necessary to have agents "buy into" a new, non-mandatory process. However, incentives may distort other input choices. In extreme cases, the "optimal incentive structure may require the elimination or muting of incentives... [to inspire] cooperation and coordination" (Holmstrom and Milgrom, 1994). Relative to this extreme case where incentives may detract from productivity, adding the relatively high-powered pair-incentives to the structured-meetings did not appear to reduce the effect of the *Structured-Meetings* treatment.

mingle sales- and learning-effort, therefore to differentiate between the two, we next turn to the post-treatment period where any (relative) productivity gains can only be the result of peer-learning.

5.3 Persistence of Productivity Gains After the Experiment

Figure 2 shows that treatments involving structured-meetings had much greater persistence in the post-treatment period than the *Pair-Incentives* treatment. The productivity increase for agents in the *Pair-Incentives* group moderately tapered over the experimental period, but once incentives were removed, average RPC for these agents immediately returned to the pre-experimental level. Agents in the *Structured-Meetings* and *Combined* treatments, on the other hand, realized productivity gains after the experiment ended. Consistent with knowledge transfer and learning, agents who had structured-meetings saw persistent gains in average RPC.

Turning to the formal statistical results, we examine the persistence of the observed productivity gains using data from the four weeks preceding the experiment and the four weeks immediately following the culmination of the experiment. We restrict attention to the four weeks after the experiment rather than a longer time series because a reorganization was initiated six weeks after the experiment ended. This allows us some buffer between the end of the experiment and the re-organization.[40] The estimating equation is:

$$\begin{aligned} \log(\text{RPC})_{i,t} \;=\; & \beta_0 + \beta_1 \text{Structured-Meetings}_i \text{ x Post-Period}_t \\ & + \beta_2 \text{Pair-Incentives}_i \text{ x Post-Period}_t + \beta_3 \text{Combined}_i \text{ x Post-Period}_t \\ & + \lambda_t + \theta_g + \varepsilon_{i,t} \end{aligned} \tag{3}$$

where each of the treatment indicators is equal to one in the post-treatment period for agents

[40] The reorganization involved reassigning agents to managers with some agents reassigned across divisions (i.e. selling different products and services). Part of the re-assignment appears to have been based on lagged productivity.

originally assigned to that treatment cell. The specification mirrors that of Table 2.

The results, reported in Table 3, are consistent with Figure 2. Specifically, agents in the *Structured-Meetings* treatment continue to have productivity gains that are over 20% higher than agents in either control group after the study. This increase is precisely estimated. In contrast, agents in the *Pair-Incentives* treatment group have changes in post-study productivity that are indistinguishable from either control group. Finally, the *Combined* treatment had positive gains of about 24% after the treatment ended. Although the *Combined* group appears to offer the most persistence, we cannot reject equality of persistence between the *Structured-Meetings* and the *Combined* treatments.

The estimates suggest that the *Pair-Incentives* treatment resulted in increased effort but no knowledge transfer. In the *Combined* group, once the firm provided agents with the structured-meetings framework, the marginal incentives *may* have reinforced learning. The difference between the *Pair-Incentives* treatment and the *Combined* treatment effects suggests that, without managerial guidance, effort was largely directed at exploiting extant human capital in pursuit of short-run gains, as opposed to being applied to peer-learning.

Finally, these post-treatment results require some clarification regarding whether the treatments are complements. Although the *Combined* treatment appears to do well in the post-treatment period, this does not suggest the practices are complements; the thought experiment to assess complementarity presumes keeping the practices in place.

5.4 Partner Quality and Knowledge Transfer

We interpret the persistent effects of treatments involving structured-meetings as the result of knowledge transfers. Supporting evidence for this view comes from an examination of whether agents with more knowledgeable partners do better both during the experiment and in the post-treatment period. For this analysis, we classify agents based on their pre-experimental productivity in the eight weeks preceding the experiment. Agents with pre-experimental productivity above the median for their division are classified as "stars." We

then estimate how being paired with a star partner affects productivity in the different treatments. We also assess whether there are differences based on an individual agent's own human capital, allowing us to recover the directional nature of who is gaining in a partner pairing.

The estimating equation is

$$\begin{aligned}\log(\text{RPC})_{i,t} \;=\; & \beta_0 + \beta_1 \text{Structured-Meetings}_i \text{ x } \text{T}_t + \beta_2 \text{Pair-Incentives}_i \text{ x } \text{T}_t \\ & + \beta_3 \text{Combined}_i \text{ x } \text{T}_t + \gamma_1 \text{Structured-Meetings}_i \text{ x } \text{T}_t \text{ x Star-Partner}_t \\ & + \gamma_2 \text{Pair-Incentives}_i \text{ x } \text{T}_t \text{ x Star-Partner}_t + \gamma_3 \text{Combined}_i \text{ x } \text{T}_t \text{ x Star-Partner}_t \\ & + \gamma_4 \text{T}_t \text{ x Star-Partner}_t + \gamma_5 \text{Ever-Star-Partner}_i + \lambda_t + \theta_g + \varepsilon_{i,t}\end{aligned} \tag{4}$$

where the variable T_t is interacted with treatments as a placeholder to indicate either the treatment period or the post-treatment period. The parameters of interest are γ_1, γ_2, and γ_3, comparing how star partners affect productivity in different treatments. The parameter γ_4 captures the baseline effect of having a star partner for agents in the *Internal Control* group. A further test of random assignment comes through γ_5, which is an indicator that the agent was ever paired with a star partner.

Table 4 shows that agents randomly paired with a star partner in the *Structured-Meetings* and *Combined* treatments were more productive in both the treatment and post-experimental period. Agents in the *Pair-Incentives* treatment did not respond differentially to star partners. In the first 3 columns, star partners are defined based on the concurrent identity of the partner. Column (1) is the baseline specification that includes the pre-experimental and treatment periods. While all active treatments (coefficients in the last 3 rows) show increases in productivity for agents with non-star partners, agents paired with a star partner increase productivity by an additional 15% in the *Structured-Meetings* treatment and by an additional 14% in the *Combined* treatment. There is no additional productivity increase that comes from being randomly paired with a star partner for agents in the *Pair-Incentives*

treatment. Recall from the prior section that agents can only learn from their partners if the latter has similar or higher levels of human capital. Accordingly, we conclude that the positive productivity gains that follow being paired with a star partner in treatments involving structured-meetings are indicative of knowledge transfer.

Not all agents can be paired with star partners, but some assignment rules may fare better than others at realizing gains. We ask whether the difference in being paired with a star partner depends on the identity of an agent. Columns (2) and (3) split the sample based on whether the agent is himself or herself a star. Comparing these columns, all treatments at baseline with a non-star partner are more effective for non-star agents (Column 2). These non-star agents have positive star partner interaction effects in treatments involving structured-meetings. Having a star partner does not raise productivity for non-star agents in the *Pairs-Incentives* treatment. When the agent in question is a star (Column 3), treatments only increase productivity when the partner is also a star. No treatment induces productivity when star agents are paired with non-star partners. The direction of this result suggests that *star agents provide knowledge to non-stars, while stars only learn from other stars.* Among star agents, however, the *Pair-Incentives* treatment never has statistically distinguishable differences from zero. This is possibly due to pair-incentives having less power for agents who individually earn more than median commissions.

The star interactive effects for treatments involving structured-meetings continue into the post-treatment period. Column (4) adds the post-treatment period and omits the experimental period. In the post-treatment analysis, the star partner interaction equals 1 for an agent who was ever paired with a star partner. Past pairing with a star partner is responsible for most of the detectable productivity increases from structured-meetings in the post-experimental period. Taken together, these results confirm knowledge transfer between more and less skilled agents when management practices dictate a process to minimize peer-interaction frictions.

5.5 Additional Evidence on the Mechanism

Ex-ante, it was not obvious that structured-meetings would result in more knowledge transfer than pair-incentives. Agents in the *Pair-Incentives* treatment could have replicated nearly identical exchanges as those that occurred in the *Structured-Meetings* treatment, and former had higher-powered incentives to do find a system to raise the pair's productivity. And, while the estimates suggest that agents in the *Pair-Incentives* did find ways to raise their sales, the estimates suggest that the channel was not knowledge-sharing.

Survey responses, participant anecdotes, and information from the worksheet content help to provide context for why treatments involving structured-meetings induced knowledge transfers. Responses on the concluding survey reported that agents in the *Structured-Meetings* treatment (1) became more familiar with their partners during the study and (2) felt the interactions with their partners positively benefited their sales ability.[41] Specifically, before being paired, only 69% of agents knew their partners, whereas 87% felt they were well-acquainted with their partner after the study. Over 80% felt that the interactions they had with their partners positively benefited their ability to perform their job. More details are available in Appendix Table A.2.

There are many examples of product-specific knowledge that was transferred between agents from the worksheets agents filled out in the *Structured-Meetings* treatment. For example, one agent shared with her partner that she was able to help customers overcome their initial skepticism by discussing two unique add-on features of the product. Another agent wrote that her customer wanted a very specific product mix, but that the requested mix was not feasible. She used her product knowledge to explain the problem and, instead, sell the customer a high-quality bundle of products. The content of both of these worksheets conveyed job-specific knowledge that could immediately be used by the agents' partners to interact with customers more effectively (recall that agents solicited their partner's responses in the treatments involving structured-meetings).

[41] This survey was administered immediately after the final week of the experiment ended.

Why didn't the *Pair-Incentives* treatment lead to these same conversations between coworkers? A recently hired sales agent in the *Structured-Meetings* treatment expressed her excitement to us when she learned that she had been paired with a very skilled coworker. Specifically, she said, "I would never have had the courage to approach him for help or advice. But since we are paired together for lunch, I get to learn from one of the best sales agents in the company!" As this anecdote is representative of similar feedback we received, it suggests active management was needed to reduce agent hesitance or other coordination frictions for fruitful peer interactions to occur.

6 Discussion

Our primary results compare the merits of incentives and management processes. While we find evidence suggesting that each approach improves performance via differing mechanisms, several questions remain, particularly around whether the productivity increases can be explained by adjustment on other margins. Thies section considers alternative explanations and discusses to what extent our findings generalize.

6.1 Did Treatments Detract from Ability to Take Calls?

There is no empirical evidence to suggest that the treatments themselves left the agents with less time to accept calls. Anecdotally, the agents only allowed the treatments to impinge on their slack time at the firm (approximately 20% of their time "on the clock"). The data confirms that an increase in RPC did not mask a change in time spent selling. Appendix Table A.3 shows that the treatment effects are similar when looking at either log revenue-per-call or log revenue-per-hour. Appendix Table A.4 shows that there was no discernible difference in adherence, defined as the time available to take calls, between the *Internal Control* and other treatments. Appendix Table A.5 also shows that the active treatments

increased total revenue in levels.[42] Importantly, while we have no evidence to suggest that our treatments reduced agents' time on the phones, any disruption due to treatments was more than offset by increases in sales revenue.

6.2 Are Turnover Differences Responsible for the Results?

In an earlier study, Sandvik et al. (2018) found that small changes at the extrinsic margin (turnover) could quickly drive observed mean performance. To address the possibility of turnover effects, we add agent fixed effects to the baseline results of Table 2 in Table A.6. The similarity of estimates suggest that the productivity gains are due to within-worker changes rather than differential turnover of unproductive agents across the different treatments.

The propensity for agents of (all abilities) to turnover, voluntarily and involuntarily, did not change for those in the *Structured-Meetings* or *Combined* treatments relative to the *Pair-Incentives*, *Internal Control*, or *External Control* groups during the treatment period. Panel A of Table 5 shows no difference in turnover rates during the experimental period. Panel B of Table 5 does show increases in turnover for agents in the *Internal Control* group relative to those in the *External Control* group, but these increases in turnover reflect seasonal differences in staffing across locations. In particular, locations with active treatments relied more heavily on college students and therefore featured higher natural attrition during the post-treatment period in the Fall. There are, however, no differences in turnover between active treatments and the *Internal Control*, all of which were located besides two college towns.

6.3 Did Pair-Incentives Agents Forfeit or Become Discouraged?

There is a possible trade-off between performance and quits or forfeitures in high-powered tournaments (Brown, 2011). Manager- and agent-based testimony suggests that agents in

[42] The total revenue figure must be imputed for the *External Control* due to reporting differences across establishments.

both the *Pair-Incentives* and *Combined* tournaments did not appear to give up.[43] To test for the possibility of "quitting" in the data, we assess how agents' performance responds to past tournament outcomes. Winners of the tournament in the last week do slightly less well than non-winners in the prior week (see Table A.7). This might be due to mean reversion after a win or possibly to income effects. The excluded category is past tournament losers, so this evidence is inconsistent with agents "quitting" by means of reducing effort in response to tournament cues.

Another possibility is that some agents simply don't begin to compete in the contest. The tournament was designed to make every agent group ex-ante similar to other groups because pre-experiment productivity was filtered out; the tournament rankings were based on gains. It is possible that star performers find it more difficult to improve their own performance in the tournament, but our pre-experiment discussions with management suggested this tournament design would best encourage non-stars to compete. If a star found it difficult to improve on their own, a more productive path would be to help one's partner.

6.4 Comparing Pair-Incentives and Estimates of Other Incentive Changes

In Sandvik et al. (2018), we analyze sales changes after one division of the firm reduced agents' effective incentive pay by altering the mapping between sales and commission-eligible revenue. The reduction occurred in the Fall/Winter of the prior year and was not concurrent with this experiment.

Sandvik et al. (2018) primarily find adjustment on the turnover margin while sales effort responses to commission reductions were muted. The elasticities of effort with respect to incentive changes thus varies across the two studies. What potentially explains the differences? First, the *Pair-Incentives* were transitory and avoided the income effects typically associated

[43] While agents were provided with daily updates of their pair's relative ranking, membership in each three-team tournament bracket was randomly determined ex-post.

with permanent wage changes (Stafford, 2015). Second, Sandvik et al. (2018) focuses on the limited long-run response to the commission reduction, whereas the effort response to the *Pair-Incentives* treatment is short-term (and begins to fade out during the experimental period).[44] Hence, the parameters across the two papers capture different time frames for effort adjustment, with the long-run response including income effects that may reduce the substitution response to changes in incentives. Finally, the *Pair-Incentives* treatment may have had a social component that increased the salience of the group incentives relative to individual incentive changes in Sandvik et al. (2018).

6.5 The Firm's Return on Investment

The economic significance of our findings was apparent for the firm. The firm previously relied exclusively on marginal incentives aimed to influence agents' performance. Interviews with management suggest that, while these interventions generally led to increased productivity, all associated gains would evaporate as soon as the incentives were removed. As modeled in Section 4.1, agent performance is a function of both sales-effort and human capital. The lack of persistent gains suggests that prior incentive changes merely triggered sales-effort. The experiment was designed from the outset to test the efficacy of various incentives at motivating peer-learning as a means of growing agents' human capital.

The *Structured-Meetings* treatment attempts to reduce the cost of knowledge transmission, whereas the *Combined* and *Pair-Incentives* treatment each contained a component of marginal incentives—albeit such incentives were perfectly aligned with a partner agent so as to incentivize agents to train one another for a common goal. Using the estimated treatment effects, we find that the performance improvement of the treated sales agents resulted in approximately $1M USD of extra-marginal revenue.

At the intervention level, we find that the *Structured-Meetings* treatment resulted in a return on investment (ROI) of over 530%, where the investment base was the cost of having

[44]Examining the short-term responses in the Sandvik et al. (2018) data, we do observe a limited-term reduction in sales, though the design is insufficiently powered to rule out a sales increase.

catered meals delivered to the firms' campuses and the top-line value was the sales-margin multiplied by the extra revenue recorded exclusively during the 4-week treatment period (all ROI details can be found in Appendix A.3). We estimate the ROI from the *Pair-Incentives* treatment at approximately 61%, as the estimated profits therein outperformed the *Internal Control* group, but the associated incentives were more costly—at least double those of the *Structured-Meetings* treatment. Finally, the *Combined* treatment had an estimated ROI of nearly 130%. This intervention was the most costly, as both lunches and monetary incentives were provided, but these sales agents did outperform the control group by the largest margin, hence the relatively large ROI.

We stress that the reported ROI values do not impound the persistent, relative revenue gains identified in the *Structured-Meetings* and *Combined* treatments. While doing-so would obviously cause the reported ROI to increase, the post-treatment observation window is too short to reliably forecast and discount back the gains observed in the post-treatment period.

Taken together, the increased productivity resulting from our relatively low-cost interventions made the randomized control trial a success with management. Pursuant to sharing our results, management has prioritized supporting future randomized control trials for organizational design.

7 Conclusion

Harnessing co-worker interactions may dramatically improve workplace efficiency, but getting individuals to interact effectively—either sharing knowledge or encouraging one another—raises a host of economic questions. An extensive literature has examined group incentives for this purpose, but little evidence from the field exists on the role of non-monetary Human Resource Management (HRM) practices. This article compares monetary incentives and HRM processes meant to foster productive co-worker interactions. We ran a field experiment that randomly paired more than 650 call center sales agents together and then assigned the

pairs to different treatments. This design allows us to test whether monetary incentives (in the form of tournaments where pairs of workers compete against other pairs) resulted in similar or superior productivity gains compared to a managerial intervention that established a process for workers to meet and share knowledge over lunch.

The *Pair-Incentives* treatment with group tournaments yielded 13% increases in performance relative to a control group that was unaware of the experiment. The *Structured-Meetings* treatment yielded 24% increases in performance relative to the same outside control group. Importantly, agents in the *Structured-Meetings* treatment had persistent performance gains, while the performance gains from the *Pair-Incentives* treatment subsided at the end of the treatment period. The tournament appeared to increase effort but did not lead to knowledge transfers or other lasting improvements in agents' human capital. The management-lead process intervention, on the other hand, resulted in greater productivity during the study period and after. A number of additional results suggest that the management-lead process resulted in knowledge transfers, primarily from skilled agents to others.

These findings add to a small but growing set of studies showing that simple management interventions can dramatically raise productivity (Bloom and van Reenen, 2011; Bloom et al., 2015, 2013; Jackson and Schneider, 2015; Haynes et al., 2009). The experimental design also enables a test of whether workers who received both high-powered incentives and peer lunch meetings performed better or worse than the sum of the productivity changes for the individual treatments. While previous research emphasizes the importance of clustering complementary management practices (Ichniowski et al., 1997; Athey and Stern, 1998), few studies have induced experimental variation to estimate what types of practices, when combined, result in output gains that are greater than the sum of their respective parts. Agents in the *Combined* treatment had submodular gains, which suggests the *Pair-Incentives* and *Structured-Meetings* treatments are substitutes.

Although our randomized controlled trial took place in a sales organization, our findings speak more broadly to peer-learning in other organizations. Many settings provide perfor-

mance incentives and opportunities to interact with coworkers. A fruitful area of future research surrounds how to match workers for these purposes.

References

Agrawal, Ajay, Devesh Kapur, John McHale. 2008. How do spatial and social proximity influence knowledge flows? evidence from patent data. *Journal of urban economics* **64**(2) 258–269.

Akerlof, George A, Rachel E Kranton. 2000. Economics and identity. *The Quarterly Journal of Economics* **115**(3) 715–753.

Ariely, Dan, Anat Bracha, Stephan Meier. 2009a. Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review* **99**(1) 544–55.

Ariely, Dan, Uri Gneezy, George Loewenstein, Nina Mazar. 2009b. Large stakes and big mistakes. *The Review of Economic Studies* **76**(2) 451–469.

Athey, Susan, Scott Stern. 1998. An empirical framework for testing theories about complimentarity in organizational design (6600). doi:10.3386/w6600. URL http://www.nber.org/papers/w6600.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K Khandelwal, Eric Verhoogen. 2017. Organizational barriers to technology adoption: Evidence from soccer-ball producers in pakistan. *The Quarterly Journal of Economics* **132**(3) 1101–1164.

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

Bandiera, Oriana, Iwan Barankay, Imran Rasul. 2007. Incentives for managers and inequality among workers: Evidence from a firm-level experiment. *The Quarterly Journal of Economics* **122**(2) 729–773.

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

Bénabou, Roland, Jean Tirole. 2006. Incentives and prosocial behavior. *American Economic Review* **96**(5) 1652–1678.

Bertrand, Marianne, Sendhil Mullainathan. 2004. Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review* **94**(4) 991–1013.

Blader, Steven, Claudine Madras Gartenberg, Andrea Prat. 2016. The contingent effect of management practices .

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

Bloom, Nicholas, James Liang, John Roberts, Zhichun Jenny Ying. 2015. Does working from home work? evidence from a chinese experiment *. *The Quarterly Journal of Economics* **130**(1) 165–218.

Bloom, Nicholas, Raffaella Sadun, John Van Reenen. 2016. Management as a technology? Tech. rep., National Bureau of Economic Research.

Bloom, Nicholas, John van Reenen. 2011. Human resource management and productivity. 1st ed., chap. 19. Elsevier, 1697–1767. URL https://EconPapers.repec.org/RePEc:eee:labchp:5-19.

Bloom, Nicholas, John Van Reenen. 2011. Human resource management and productivity. *Handbook of labor economics* **4** 1697–1767.

Brown, Jennifer. 2011. Quitters never win: The (adverse) incentive effects of competing with superstars. *Journal of Political Economy* **119**(5) 982–1013.

Brynjolfsson, Erik, Paul Milgrom. 2013. Complementarity in organizations. *The handbook of organizational economics* 11–55.

Cai, Jing, Adam Szeidl. 2016. Interfirm relationships and business performance. *The Quarterly Journal of Economics* .

Carpenter, Jeffrey P, Glenn W Harrison, John A List. 2005. *Field experiments in economics*. Elsevier JAI.

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

Catalini, Christian. 2017. Microgeography and the direction of inventive activity. *Management Science* .

Chan, Tat Y, Jia Li, Lamar Pierce. 2014. Compensation and peer effects in competing sales teams. *Management Science* **60**(8) 1965–1984.

Deci, Edward L, Richard Koestner, Richard M Ryan. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* **125**(6) 627.

Ederer, Florian, Gustavo Manso. 2013. Is pay for performance detrimental to innovation? *Management Science* **59**(7) 1496–1513.

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

Fehr, Ernst, Lorenz Goette. 2007. Do workers work more if wages are high? evidence from a randomized field experiment. *American Economic Review* **97**(1) 298–317.

Frey, Bruno S, Felix Oberholzer-Gee. 1997. The cost of price incentives: An empirical analysis of motivation crowding-out. *The American Economic Review* **87**(4) 746–755.

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

Garlick, Robert. 2014. Academic peer effects with different group assignment rules: Residential tracking versus random assignment. Tech. rep., Mimeo, Duke.

Glaeser, Edward L, Joshua D Gottlieb. 2009. The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature* **47**(4) 983–1028.

Glaeser, Edward L, Hedi D Kallal, Jose A Scheinkman, Andrei Shleifer. 1992. Growth in cities. *Journal of political economy* **100**(6) 1126–1152.

Gneezy, Uri, Stephan Meier, Pedro Rey-Biel. 2011. When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives* **25**(4) 191–210.

Gosling, Samuel D, Peter J Rentfrow, William B Swann. 2003. A very brief measure of the big-five personality domains. *Journal of Research in personality* **37**(6) 504–528.

Hanna, Rema, Sendhil Mullainathan, Joshua Schwartzstein. 2014. Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics* **129**(3) 1311–1353.

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

Hasan, Sharique, Rembrand Koning. 2017. Conversational peers and idea generation: Evidence from a field experiment .

Haynes, Alex B, Thomas G Weiser, William R Berry, Stuart R Lipsitz, Abdel-Hadi S Breizat, E Patchen Dellinger, Teodoro Herbosa, Sudhir Joseph, Pascience L Kibatala, Marie Carmela M Lapitan, et al. 2009. A surgical safety checklist to reduce morbidity and mortality in a global population. *New England Journal of Medicine* **360**(5) 491–499.

Holman, David, Rosemary Batt, Ursula Holtgrewe. 2007. The global call center report: International perspectives on management and employment .

Holmstrom, Bengt, Paul Milgrom. 1994. The firm as an incentive system. *The American Economic Review* 972–991.

Ibarra, Herminia, Morten T Hansen. 2011. Are you a collaborative leader. *Harvard business review* **89**(7/8) 68–74.

Ichniowski, Casey, Kathryn Shaw, Giovanna Prennushi. 1997. The effects of human resource management practices on productivity: A study of steel finishing lines. *The American Economic Review* 291–313.

Jackson, C Kirabo, Henry S Schneider. 2015. Checklists and worker behavior: A field experiment. *American Economic Journal: Applied Economics* **7**(4) 136–68.

Lazear, Edward P, Kathryn L Shaw. 2007. Personnel economics: The economist's view of human resources. *Journal of economic perspectives* **21**(4) 91–114.

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

Levitt, Steven D, John A List. 2009. Field experiments in economics: The past, the present, and the future. *European Economic Review* **53**(1) 1–18.

Lingard, Lorelei, Glenn Regehr, Beverley Orser, Richard Reznick, G Ross Baker, Diane Doran, Sherry Espin, John Bohnen, Sarah Whyte. 2008. Evaluation of a preoperative checklist and team briefing among surgeons, nurses, and anesthesiologists to reduce failures in communication. *Archives of surgery* **143**(1) 12–17.

List, John A, Imran Rasul. 2011. Field experiments in labor economics. *Handbook of labor economics* **4** 103–228.

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

Milgrom, Paul, John Roberts. 1990. The economics of modern manufacturing: Technology, strategy, and organization. *The American Economic Review* **80**(6) 511–528.

Myers, Christopher. 2015. Is your company encouraging employees to share what they know. *Harvard Business Review* 1–9.

Nagin, Daniel S, James B Rebitzer, Seth Sanders, Lowell J Taylor. 2002. Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review* **92**(4) 850–873.

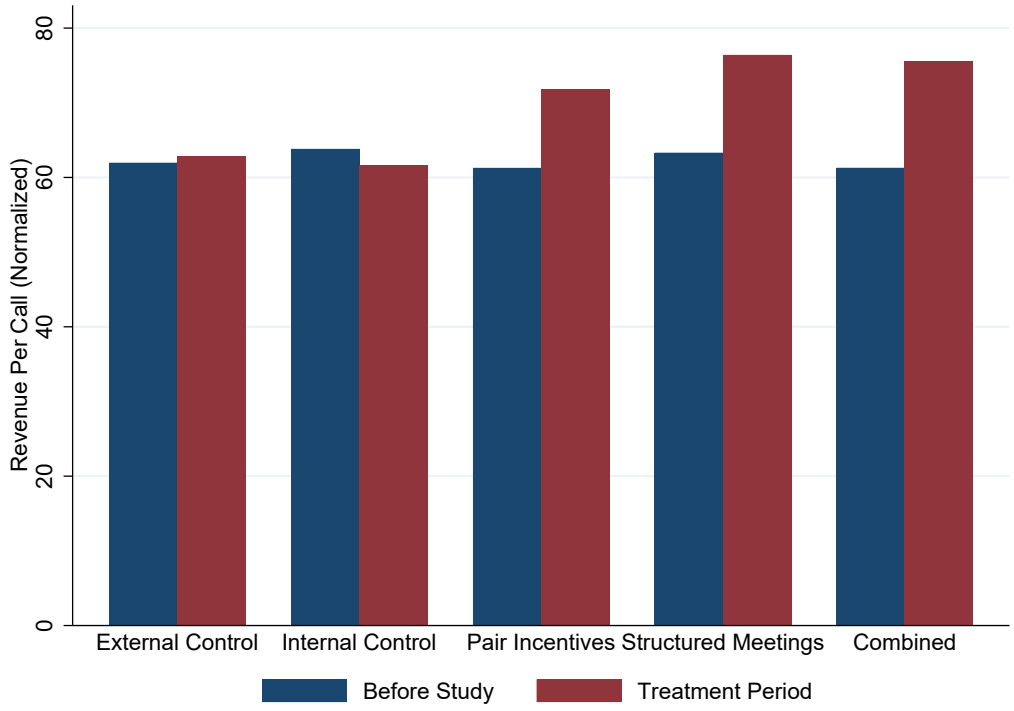
Sandvik, Jason, Nathan Seegert, Richard Saouma, Christopher Stanton. 2018. Analyzing the aftermath of a compensation reduction .

Stafford, Tess M. 2015. What do fishermen tell us that taxi drivers do not? an empirical investigation of labor supply. *Journal of Labor Economics* **33**(3) 683–710.

Titmuss, Richard Morris. 1970. *The Gift Relationship. From Human Blood to Social Policy.*. London: George Alien & Unwin Ltd.

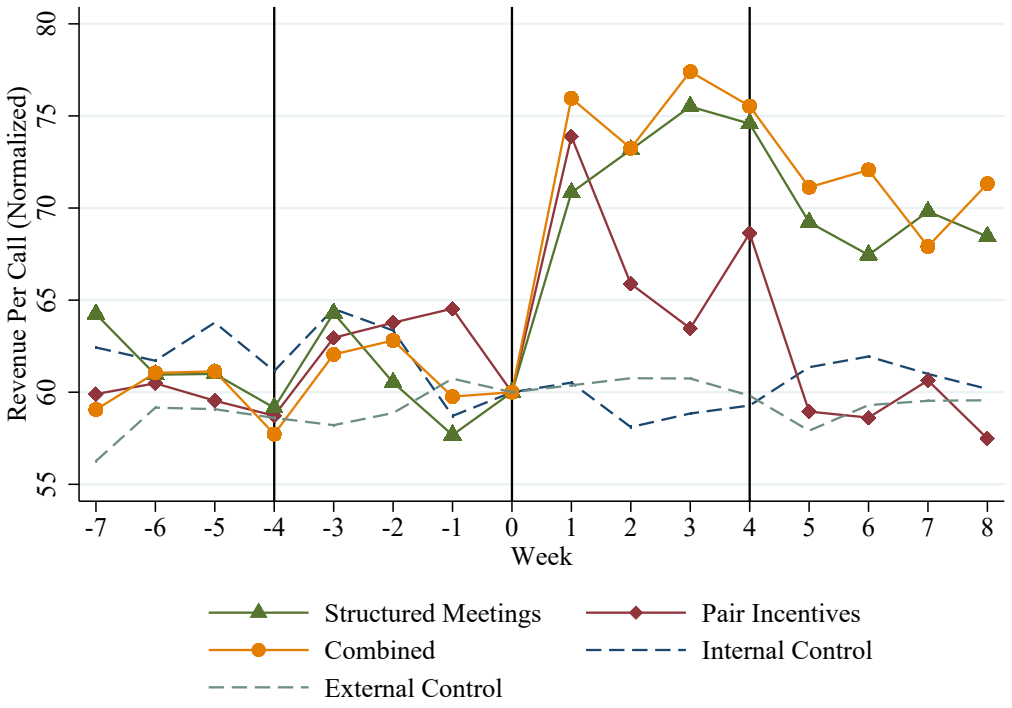
8 Tables and Figures

Figure 1: Mean RPC by Treatment During the Pre-Study and Experimental Period



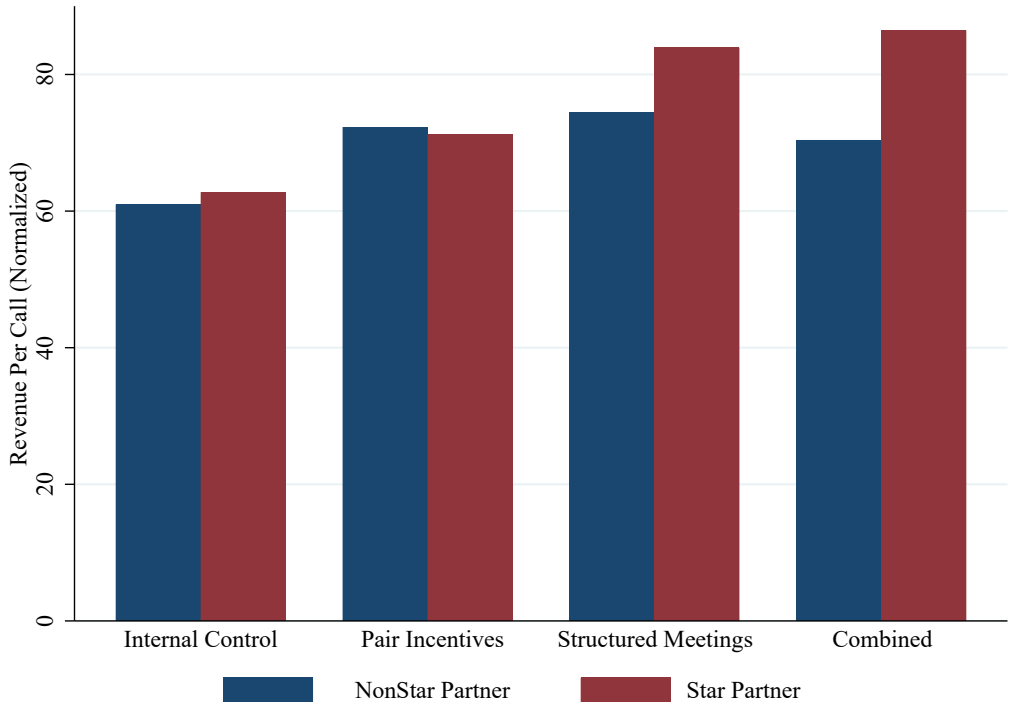
Average revenue-per-call (RPC) by treatment group. The bars for each treatment are normalized around the grand-mean of RPC in the pre-treatment period.

Figure 2: Treatment Effects by Week



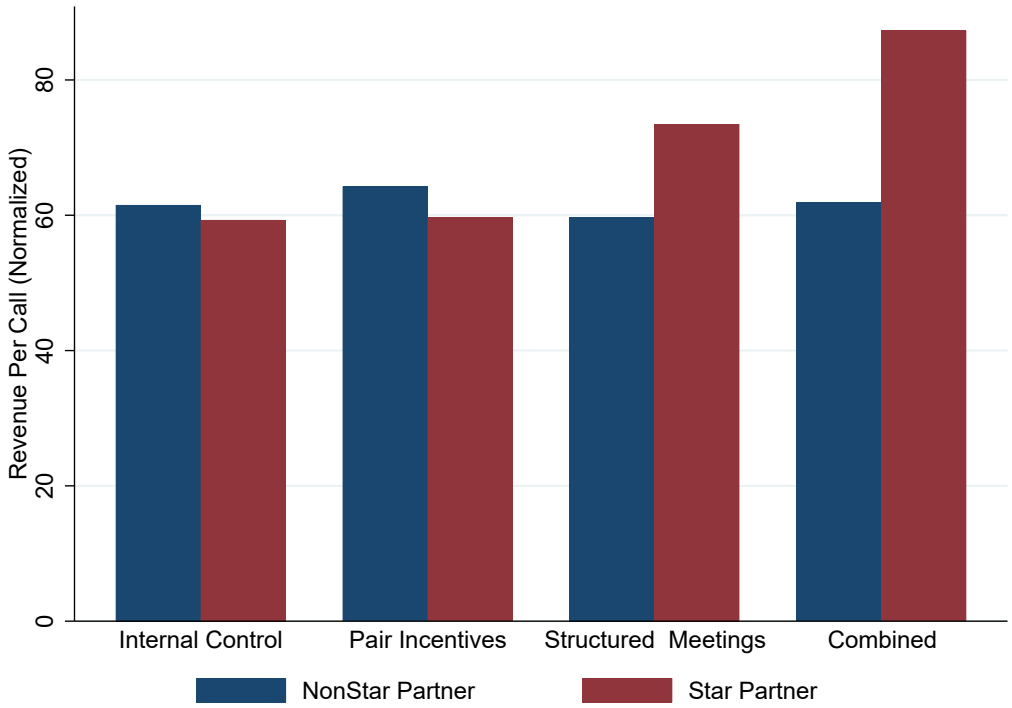
Average revenue-per-call (RPC) by week and treatment group. Each series is normalized to the grand-mean of RPC in week 0. The experimental intervention begins in week 1 and continues to week 4. After week 10, the firm began a post-summer re-organization that partially depended on lagged productivity, so we restrict analysis to the 8 weeks that follow the start of the experiment.

Figure 3: Mean Productivity During the Experimental Period by Star Partner Assignment



Average revenue-per-call (RPC) by treatment group during the treatment period based on whether the concurrent partner is a Star agent (defined as above median productivity in the pre-treatment period).

Figure 4: Mean Productivity During the Post-Experimental Period by Star Partner Assignment



Average revenue-per-call (RPC) by treatment group during the post-treatment period based on whether the agent was ever partnered with a Star (defined as above median productivity in the pre-treatment period).

Table 1: **Agent Demographics and Productivity Statistics Prior to the Experiment**

Sales agent demographic information (age, tenure with the firm, and gender ratio) and weekly performance measures are listed below for the full sample of participating agents and broken out by treatment group. Except for attrition, these numbers are derived from a four-week average immediately before the experiment began. Other than age, tenure, percent female, and attrition, each productivity variable is a weekly measure. A *Phone Hour* is a measure of time an agent is logged into the phone system, and *Adherence* is calculated as the sum of an agent's time available to receive a call and his time on calls, all divided by the total time he is logged into the phone system. *Attrition* is measured as the number of agents that quit during the pre-experimental period, experimental period, and post-experimental period divided by the total number of agents in that treatment. P-values of randomization tests of mean differences in the *Internal Control* and active treatment columns are reported. These tests are computed as the joint-hypothesis test of equality of treatment groups from a regression of the variable of interest on treatment assignment dummies after clustering standard errors based on manager identity (the level of assignment).

	Full Sample	Structured Meetings	Pair Incentives	Combined	Internal Control	External Control	P-Value
Age (yrs.)							
Mean	26.08	25.76	26.61	26.43	25.14	27.19	0.62
Median	23.39	22.51	23.55	24.02	22.97	24.63	
Std Dev.	8.14	8.20	9.61	8.10	6.66	8.41	
Tenure (log days)							
Mean	5.25	5.14	5.38	5.59	5.18	4.67	0.61
Median	5.15	4.62	5.40	5.37	4.62	5.18	
Std Dev.	1.18	1.12	1.07	1.10	1.22	1.24	
Percent Female							
Mean	0.32	0.32	0.31	0.33	0.34	0.25	0.95
Revenue per Call (log)							
Mean	3.92	3.90	4.06	3.94	3.92	3.62	0.69
Median	3.97	4.04	4.09	3.99	3.99	3.69	
Std Dev.	0.49	0.52	0.37	0.47	0.55	0.33	
Revenue per Hour (log)							
Mean	4.51	4.48	4.69	4.56	4.51	4.11	0.54
Median	4.62	4.64	4.78	4.65	4.63	4.18	
Std Dev.	0.60	0.69	0.45	0.59	0.59	0.46	
Commission							
Mean	217.78	202.65	230.41	230.64	202.31		0.75
Median	185.45	168.42	192.28	209.73	169.73		
Std Dev.	155.61	159.99	156.09	157.73	147.70		
Total Calls							
Mean	61.53	57.56	64.16	65.81	58.89		0.33
Median	60.43	57.22	62.41	65.29	58.63		
Std Dev.	21.32	19.16	22.02	20.81	22.43		
Phone Hours							
Mean	32.61	32.52	33.76	33.17	31.22		0.32
Median	34.05	34.08	34.77	33.33	33.75		
Std Dev.	7.36	7.01	6.09	6.74	8.95		
Adherence							
Mean	0.80	0.80	0.84	0.79	0.77		0.19
Median	0.83	0.83	0.85	0.83	0.82		
Std Dev.	0.14	0.11	0.07	0.14	0.21		
Attrition							
Term Rate	0.22	0.22	0.18	0.24	0.24		0.19
N_A (Agents)	736	158	135	174	186	83	

Table 2: **Difference-in-differences Estimates of Performance Effects of the Treatments.**

This table reports regression estimates of the log of revenue-per-call using data from the 4 weeks before and 4 weeks during the experiment. The variables *Structured-Meetings*, *Pair-Incentives*, and *Combined* are shorthand for "Structured-Meetings x Treatment-Period," "Pair-Incentives x Treatment-Period," and "Combined x Treatment-Period" and are set to 1 in the experimental period for those randomly assigned to those treatments, and zero otherwise. Using the abbreviated variable names, the baseline estimating equation is:

$$\log(\text{RPC})_{i,t} = \beta_0 + \beta_1 \text{Structured-Meetings}_{i,t} + \beta_2 \text{Pair-Incentives}_{i,t} + \beta_3 \text{Combined}_{i,t} + \lambda_t + \theta_g + \varepsilon_{i,t}$$

where i represents an agent, t represents week, g represents sales manager group, λ_t and θ_g are week and sales manager fixed effects, and $\varepsilon_{i,t}$ is an idiosyncratic error term. Indicators for treatment assignment in the pre-period are absorbed by θ_g as randomization is at the sales manager group level. In Columns (1)–(3) the *Internal Control* (passive pairs) is the omitted category. Column (4) omits the *Internal Control* group and instead uses the *External Control* group (that was not aware of the experiment and had no partner pairing) as the excluded category. Column (5) includes both control groups, with an indicator for the *Internal Control* during the experimental period. Specifications with agent or partner demographics include age, gender, and tenure with the firm. The p-values in the bottom rows report results from Wald tests of two null hypotheses: i) equality of effects between *Pair-Incentives* and *Structured-Meetings* and ii) that the *Combined* group had supermodular gains. Standard errors are clustered at the sales manager level and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

Control Group:	Internal (Passive Pairs)			External (No Pairs)	Both
	(1)	(2)	(3)	(4)	(5)
Structured-Meetings	0.241***	0.241***	0.241***	0.247***	0.244***
	(0.045)	(0.044)	(0.044)	(0.073)	(0.071)
Pair-Incentives	0.131***	0.136***	0.136***	0.124*	0.134*
	(0.048)	(0.047)	(0.047)	(0.071)	(0.071)
Combined	0.255***	0.258***	0.256***	0.259***	0.259***
	(0.043)	(0.044)	(0.044)	(0.069)	(0.069)
Internal Control					0.004
					(0.065)
Demographics		✓	✓	✓	✓
Partner Demo.			✓		
Manager FE (θ_g)	✓	✓	✓	✓	✓
Week FE (λ_t)	✓	✓	✓	✓	✓
Adj. R-Square	0.470	0.482	0.483	0.459	0.476
Observations	3,418	3,418	3,418	2,856	3,821
Agents	653	653	653	580	736
Managers	52	52	52	45	58
P-Values:					
H_0: Meetings = Incent.	0.048	0.051	0.049	0.033	0.050
H_0: Meetings+Incent.≤Comb.	0.049	0.042	0.039	0.110	0.094

Table 3: **Do Some Treatments Result in Greater Persistence of Productivity Gains?**

This table reports regression estimates of the log of revenue-per-call using data from the 4 weeks before the experiment and 4 weeks after the end of the experiment. The variables *Structured-Meetings*, *Pair-Incentives*, and *Combined* are shorthand for "Structured-Meetings x Post-Period", "Pair-Incentives x Post-Period", and "Combined x Post-Period" are set to 1 in the post-experimental period for those randomly assigned to those treatments, and zero otherwise. The baseline estimating equation is the same as in Table 2 except the treatment period is omitted and the post-experimental period is included. Standard errors are clustered at the sales manager level and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

Control Group:	Internal (Passive Pairs)			External (No Pairs)	Both
	(1)	(2)	(3)	(4)	(5)
Structured-Meetings	0.208***	0.204***	0.204***	0.214***	0.214***
	(0.061)	(0.061)	(0.061)	(0.077)	(0.076)
Pair-Incentives	0.030	0.035	0.035	0.028	0.037
	(0.057)	(0.057)	(0.056)	(0.072)	(0.071)
Combined	0.242***	0.246***	0.244***	0.247***	0.247***
	(0.062)	(0.063)	(0.063)	(0.076)	(0.075)
Internal Control					0.005
					(0.070)
Demographics		✓	✓	✓	✓
Partner Demo.			✓		
Manager FE (θ_g)	✓	✓	✓	✓	✓
Week FE (λ_t)	✓	✓	✓	✓	✓
Adj. R-Square	0.418	0.429	0.430	0.440	0.448
Observations	3,252	3,252	3,252	3,295	3,627

Table 4: **How do Partner and Agent Baseline Productivity Influence Treatment Outcomes?**

This table examines the effect of star partners on treatment outcomes. An agent is defined as a star if they are an above median performer within their own division 0 to 8 weeks before the study. The *Internal Control* group is the excluded category. The table reports regressions of $log(RPC)$ that add interactions for treatment assignment and being paired with a star partner. Columns (1)—(3) include the pre-period and experimental period while Column (4) includes the pre-period and post-treatment period. During the treatment, star partner assignment is defined based on the concurrent partner. For the post-experimental period, *Star Partner* is defined based on whether the agent was ever paired with a star. The variable "Star Partner x Study/Post" is set to 1 if the agent is paired with a star partner. The variable *Ever Star Partner* is set to one during the entire sample if the agent was ever paired with a star, serving as a test of non-random partner assignment. Columns (2) and (3) condition on the productivity of the agent in question and examine heterogeneous effects of star partnership by agent productivity. Each regression includes time fixed effects and sales manager group fixed effects. Note that observation counts differ due to agents whose assigned partner was not available or was absent that week. These agents are included in prior tables, but based on lack of pair data, are not included in this subsample. Standard errors are clustered at the sales manager level and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

	Baseline	Non-Star Agents	Star Agents	Post-Period
	(1)	(2)	(3)	(4)
Meetings × Star Partner	0.150***	0.210**	0.147*	0.153**
	(0.054)	(0.087)	(0.078)	(0.075)
Incentives × Star Partner	-0.021	-0.031	0.069	-0.104
	(0.067)	(0.088)	(0.085)	(0.081)
Combined × Star Partner	0.136**	0.250***	0.206**	0.232***
	(0.063)	(0.071)	(0.077)	(0.076)
Star Partner × Study/Post	-0.031	-0.009	-0.071	0.012
	(0.049)	(0.062)	(0.068)	(0.062)
Ever Star Partner	-0.016	-0.015	-0.043	0.039
	(0.026)	(0.044)	(0.033)	(0.024)
Structured-Meetings	0.222***	0.241***	0.068	0.030
	(0.055)	(0.075)	(0.072)	(0.048)
Pair-Incentives	0.105*	0.290***	-0.012	0.073
	(0.061)	(0.064)	(0.082)	(0.062)
Combined	0.190***	0.372***	0.061	0.058
	(0.057)	(0.064)	(0.062)	(0.049)
Manager FE (θ_g)	✓	✓	✓	✓
Week FE (λ_t)	✓	✓	✓	✓
Adj. R-Square	0.488	0.491	0.450	0.441
Observations	3,287	1,465	1,818	3,024

Table 5: **Difference-in-differences Estimates of Worker Turnover.**

Panel A of this table reports regression estimates of employee turnover using data from the 4 weeks before and 4 weeks during the experiment. Panel B uses data from the 4 weeks before and 4 weeks after the experiment. The variables *Structured-Meetings*, *Pair-Incentives*, and *Combined* are set to 1 in the treatment/post period for those randomly assigned to those treatments, and zero otherwise. The baseline estimating equation is:

$$\text{Turnover}_{i,t} = \beta_0 + \beta_1 \text{Structured-Meetings}_{i,t} + \beta_2 \text{Pair-Incentives}_{i,t} + \beta_3 \text{Combined}_{i,t} + \lambda_t + \theta_g + \varepsilon_{i,t}$$

where i represents an agent, t represents week, g represents sales manager group, λ_t and θ_g are week and sales manager fixed effects, respectively, and $\varepsilon_{i,t}$ is an idiosyncratic error term. Indicators for treatment assignment in the pre-period are absorbed by θ_g as randomization is at the sales manager group level. In Columns (1)–(3) the *Internal Control* (passive pairs) is the omitted category. Column (4) omits the *Internal Control* group and instead uses the *External Control* group (that was not aware of the experiment and had no partner pairing) as the excluded category. Column (5) includes both control groups, with an indicator for the *Internal Control* during the experimental period. Specifications with agent or partner demographics include age, gender, and tenure with the firm. Standard errors are clustered at the sales manager level and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

Control Group:	Internal (Passive Pairs)			External (No Pairs)	Both
	(1)	(2)	(3)	(4)	(5)
Panel A: Pre-Period and Treatment-Period					
Structured-Meetings	-0.018	-0.018	-0.018	-0.025	-0.023
	(0.021)	(0.021)	(0.021)	(0.041)	(0.041)
Pair-Incentives	-0.011	-0.011	-0.011	-0.017	-0.016
	(0.021)	(0.021)	(0.021)	(0.041)	(0.040)
Combined	0.006	0.006	0.006	0.002	0.002
	(0.023)	(0.023)	(0.023)	(0.041)	(0.041)
Internal Control					-0.005
					(0.045)
Adj. R-Square	0.029	0.029	0.029	0.020	0.027
Observations	3,418	3,418	3,418	2,856	3,821
Panel B: Pre-Period and Post-Period					
Structured-Meetings	-0.014	-0.014	-0.014	0.028**	0.027**
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Pair-Incentives	-0.011	-0.012	-0.012	0.029***	0.029***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Combined	-0.010	-0.011	-0.011	0.030**	0.030**
	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)
Internal Control					0.041***
					(0.012)
Demographics		✓	✓	✓	✓
Partner Demo.			✓		
Manager FE (θ_g)	✓	✓	✓	✓	✓
Week FE (λ_t)	✓	✓	✓	✓	✓
Adj. R-Square	0.010	0.010	0.010	0.009	0.009
Observations	3,252	3,252	3,252	3,295	3,627

Table A.1: **Preliminary Survey Responses for Treatment-Eligible Agents**

These are selected questions from the preliminary survey that we administered one week prior to the start of the experiment. The survey was not administered to the *External Control* group. The question wording, as it is displayed in the table, has been adapted from its original form to eliminate institutionally unique jargon. Agents were provided with a link to the quiz and were asked to complete it while at work. Agents were not aware of which treatment they were going to be placed in at the time they took the survey. The question regarding the dollar value of the proposed incentives is the average valuation for the set of prizes offered in the *Pair-Incentives* tournament.

	Full Sample	Internal Control	Pair-Incentives	Structured-Meetings	Combined
On a scale of 1-5, how connected do you feel to others within the firm?	3.7	3.7	3.4	3.7	3.8
How many work-related interactions do you initiate in an average work week?	5.8	5.0	5.3	7.1	6.1
On a scale of 1-5, how beneficial are these interactions to you personally?	3.9	3.9	3.9	4.0	4.0
How many work-related interactions do others initiate with you in an average work week?	6.2	5.8	5.3	6.8	6.9
On a scale of 1-5, how beneficial do you feel these interactions are for your coworkers?	4.2	4.2	4.2	4.2	4.3
What dollar value would you be willing to spend on the proposed incentives?	\$40.20				
Gosling Personality Measures					
Extroversion	5.1	5.1	5.2	5.1	5.1
Agreeableness	5.0	5.0	4.9	5.1	5.1
Conscientiousness	5.8	5.7	5.8	6.0	5.7
Emotional Stability	5.3	5.4	5.4	5.6	5.2
Openness to Experiences	5.8	5.6	5.8	6.0	5.8
N_A (Agents)	312	86	66	62	98

Table A.2: **Intervention Effectiveness/Salience Survey Responses for Treatment-Eligible Agents**

These are selected questions that were asked on the concluding survey. The question wording, as it is displayed in the table, has been adapted from its original form to eliminate institutionally unique jargon. Agents were provided with a link to the survey and were asked to complete it while at work. These questions are meant to assess the overall salience of the competition and the effectiveness of the *Structured-Meetings* and *Combined* treatments in encouraging and facilitating peer interactions.

	Full Sample	Internal Control	Pair-Incentives	Structured-Meetings	Combined
I was aware of the treatment that took place this past month.	82.5%	77.4%	78.3%	84.8%	92.0%
I was well-acquainted with my partner before the study.	68.4%	54.4%	73.1%	69.0%	76.1%
I was well-acquainted with my parter after the study.	84.3%	65.6%	83.6%	86.8%	96.8%
We turned in a completed worksheet each week.				82.6%	88.2%
I spent [ ] minutes with my partner on the worksheet.				6.3	7.3
These interactions with my partner were beneficial.				78.6%	76.0%
N_A (Agents)	378	115	83	105	75

Table A.3: **Does Revenue Per Hour Change with Treatments?**

This table replicates the specification in Table 2 but changes the dependent variable to the log of revenue-per-hour, log(RPH). All specifications include week and manager fixed effects. Standard errors are clustered at the sales manager level and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

Control Group:	Internal (Passive Pairs)			External (No Pairs)	Both
	(1)	(2)	(3)	(4)	(5)
Structured-Meetings	0.262***	0.262***	0.263***	0.232**	0.220**
	(0.078)	(0.076)	(0.076)	(0.093)	(0.090)
Pair-Incentives	0.152**	0.158**	0.159**	0.108	0.115
	(0.060)	(0.060)	(0.060)	(0.074)	(0.074)
Combined	0.213***	0.218***	0.215***	0.172**	0.174**
	(0.061)	(0.061)	(0.061)	(0.075)	(0.074)
Internal Control					-0.043
					(0.073)
Demographics		✓	✓	✓	✓
Partner Demo.			✓		
Manager FE (θ_g)	✓	✓	✓	✓	✓
Week FE (λ_t)	✓	✓	✓	✓	✓
Adj. R-Square	0.470	0.482	0.483	0.459	0.476
Observations	3,418	3,418	3,418	2,856	3,821

Table A.4: **Does Ability to Answer Calls Change with Treatments?**

Here we compare the change in scheduled time to take calls, defined as adherence, across the three different treatment groups using data from the 4 weeks before and 4 weeks during the study. The variables *Structured-Meetings*, *Pair-Incentives*, and *Combined* are set to 1 in the study weeks for those randomly assigned to those treatments, and zero otherwise. Each specification contains week and sales manager fixed effects. The *Internal Control* group is the omitted category, as we do not have data on adherence for the *External Control* group. Standard errors are clustered at the sales manager level and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

	(1)	(2)	(3)
Structured-Meetings	-0.041	-0.040	-0.041
	(0.031)	(0.031)	(0.031)
Pair-Incentives	-0.036	-0.036	-0.036
	(0.030)	(0.030)	(0.030)
Combined	-0.019	-0.019	-0.020
	(0.047)	(0.046)	(0.047)
Demographics		✓	✓
Partner Demo.			✓
Manager FE (θ_g)	✓	✓	✓
Week FE (λ_t)	✓	✓	✓
Adj. R-Square	0.114	0.114	0.116
Observations	2,550	2,550	2,550

Table A.5: **Before and After Analysis of Total Revenue Changes by Treatment**

This table compares total revenue levels using data from the 4 weeks before and 4 weeks during the study. The specification follows that in Table 2. Note, however, that total revenue must be imputed for the *External Control* group as hours and total calls are not recorded for this group. The imputation uses estimates of changes in total calls by division and applies them to the *External Control* group. Standard errors are clustered at the sales manager level and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

Control Group:	Internal (Passive Pairs)			External (No Pairs)	Both
	(1)	(2)	(3)	(4)	(5)
Structured-Meetings	562.28*	568.68*	575.21*	562.26***	528.10***
	(296.54)	(299.77)	(301.92)	(178.23)	(184.33)
Pair-Incentives	531.18*	544.90*	544.71*	464.94**	496.55***
	(294.29)	(298.45)	(297.27)	(174.03)	(172.52)
Combined	716.80**	733.55**	708.98**	696.17***	704.25***
	(317.00)	(321.22)	(326.25)	(225.41)	(221.04)
Internal Control					-27.78
					(239.54)
Demographics		✓	✓	✓	✓
Partner Demo.			✓		
Manager FE (θ_g)	✓	✓	✓	✓	✓
Week FE (λ_t)	✓	✓	✓	✓	✓
Adj. R-Square	0.286	0.292	0.297	0.458	0.430
Observations	3,186	3,186	3,186	2,709	3,589

Table A.6: **Main Results are Robust to the Inclusion of Agent Fixed Effects**

This table replicates the specification in Table 2 while including sales agent fixed effects in addition to week fixed effects. Standard errors are clustered at the sales manager level and are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are indicated by *, **, and ***, respectively.

Control Group:	Internal (Passive Pairs)			External (No Pairs)	Both
	(1)	(2)	(3)	(4)	(5)
Structured-Meetings	0.241***	0.241***	0.241***	0.253***	0.250***
	(0.038)	(0.037)	(0.037)	(0.053)	(0.052)
Pair-Incentives	0.131***	0.136***	0.136***	0.134***	0.145***
	(0.038)	(0.037)	(0.037)	(0.052)	(0.052)
Combined	0.255***	0.258***	0.256***	0.268***	0.269***
	(0.033)	(0.034)	(0.034)	(0.049)	(0.049)
Internal Control					0.010
					(0.048)
Demographics		✓	✓	✓	✓
Partner Demo.			✓		
Manager FE (θ_g)	✓	✓	✓	✓	✓
Week FE (λ_t)	✓	✓	✓	✓	✓
Adj. R-Square	0.470	0.482	0.483	0.465	0.485
Observations	3,418	3,418	3,418	2,856	3,821

Table A.7: **Do Agents Give-up After Winning or Losing the Tournament**

This table reports regressions of log RPC on an indicator that the agent received a tournament prize in the prior week. The estimate is relative to a baseline of agents who did not win in the previous week. The sample contains only those agents in either the *Pair Incentives* or *Combined* treatments during the experimental period.

	Pair-Incentives	Combined	Both
	(1)	(2)	(3)
Won Last Week	-0.035**	-0.027*	-0.033***
	(0.014)	(0.014)	(0.010)
Demographic Controls	✓	✓	✓
Manager FE (θ_g)	✓	✓	✓
Week FE (λ_t)	✓	✓	✓
R-Square	0.114	0.159	0.136
Observations	673	770	1,446

A Appendix

A.1 Survey Responses and Robustness Tables

Several survey results are compiled in Table A.1. All surveys were administers through Qualtrics and distributed via email and links on the experiment website. Over 300 agents completed the preliminary survey, answering questions about their social and work-related conversations with coworkers and providing self-reflective personality scores based on the Gosling personality measures of extroversion, agreeableness, conscientiousness, emotional stability, and openness to experiences (Gosling et al., 2003).

Post-experiment survey results are in Table A.2. These questions allow us to get an approximate measure of the effectiveness and salience of the experiment as a whole and of the *Structured-Meetings* treatment specifically. Several questions were asked on all three surveys (preliminary, intermediate, and post).

A.2 Institutional Detail

A *Phone Hour* is a measure of time an agent is logged into the phone system and is, therefore, able to receive inbound calls. This measure relates very closely to an important metric within the firm: *Adherence*, which is the sum of an agent's time available to receive a call and his time on calls, all divided by the total time he is logged into the phone system (i.e., not in meetings, at lunch, or on scheduled breaks). Agents can be logged into the phone system and mark themselves as "unavailable" (e.g., when they use the restroom or get a drink), and this time goes against their adherence. Agents are required to maintain a level of adherence above 80% or they run a serious risk of being terminated, and this is reflected in the mean (median) adherence being 79% (84%).[45]

A.3 The Study Firm's Return on Investment

Here we detail the procedure used to estimate the return on investment of each intervention and the total extra-ordinary sales revenue generated by the experiment. The *Internal Control* group

[45] In fact, in our full data sample, agents achieve weekly adherence above 80% over 75% of the time.

revenue-per-call (RPC) during the four week experiment was, on average, \$64.01. If we multiple these by our resultant treatment effects (24.1%, 13.1%, and 25.5% for the *Structured-Meetings*, *Pair-Incentives*, and *Combined* treatment, respectively), we get the additional revenue-per-call generated by each intervention.

- *Structured-Meetings*: \$64.20 $\times$ 24.1% $\approx$ \$15.50 extra per call.
- *Pair-Incentives*: \$64.20 $\times$ 13.1% $\approx$ \$8.40 extra per call.
- *Combined*: \$64.20 $\times$ 25.5% $\approx$ \$16.40 extra per call.

We then multiply these numbers by the average number of calls per agent per week within each intervention during the four weeks of the study—58, 64, and 66 calls per week for the *Structured-Meetings*, *Pair-Incentives*, and *Combined* treatment, respectively.

- *Structured-Meetings*: \$15.50 $\times$ 58 = \$899 extra per agent per week.
- *Pair-Incentives*: \$8.40 $\times$ 64 = \$538 extra per agent per week.
- *Combined*: \$16.40 $\times$ 66 = \$1,082 extra per agent per week.

Next we multiply these amounts by the number of agent-weeks in each intervention to get to total amount of extra revenue generated by the four-week intervention.

- *Structured-Meetings*: \$899 $\times$ 379 = \$340,721 extra revenue earned across all four weeks.
- *Pair-Incentives*: \$538 $\times$ 396 = \$213,048 extra revenue earned across all four weeks.
- *Combined*: \$1,082 $\times$ 353 = \$381,946 extra revenue earned across all four weeks.

Now we consider the costs of implementing each intervention.

- *Structured-Meetings*: \$7 was spent on all agents for lunch each week: \$7 $\times$ 379 = \$2,700 in treatment costs (rounded up).
- *Pair-Incentives*: 1/3 of the agents won a prized valued at \$50: \$50 $\times$ 1/3 $\times$ 396 $\approx$ \$6,600 in treatment costs (rounded up)
- *Combined*: Consider both of these two cost structures for the 353 agent-weeks in this intervention: (\$7 $\times$ 353) + (\$50 $\times$ 1/3 $\times$ 353) $\approx$ \$8,400 in treatment costs (rounded up)

Finally, we calculate the return on investment of a individual intervention as

$$\text{ROI} = \frac{(\text{Extra Revenue} \times \text{Profit Fraction}) - \text{Treatment Cost}}{\text{Treatment Cost}} \tag{A.1}$$

where *Extra Revenue* equals the extra revenue earned from the given intervention across all four weeks of the experiment, *Profit Fraction* equals 5%,[46] and *Treatment Cost* equals the treatment cost calculated above. Performing computations leads to ROIs of 531%, 61%, and 127% for the *Structured-Meetings*, *Pair-Incentives*, and *Combined* treatments, respectively. Summing up the three extra revenue earned numbers results in a total extra revenue earned of ($340,721 + $213,048 + $381,946) = $935,715.

[46] This is a conservative estimate that is motivated by conversations had with firm executives.

B Documentation

The following are materials that were provided to participating sales agents and their supervisors in an effort to streamline the communication, increase the salience of the competition, and gather self-reported data. The first two sheets show the front and back sides of the collaboration worksheets completed by agents in the *Structured-Meetings* and *Combined* treatments. The next sheet contains the lunchtime talking points that we encouraged partners to discuss as they ate their free lunch (those in the *Structured-Meetings* and *Combined* treatments only).

[redacted] Sales Representative Collaboration Worksheet
PLEASE PRINT LEGIBLY

Your Full Name: _____ [redacted]: _____

Think about the **most successful** sales call **you** had in the last week. What did you do that made it successful?

Think about the **least successful** sales call **you've** had in the last week. How could you have done better?

Describe the most difficult deal-breaker that **you've** come across **in the last week**; for example: *upgrading callers to a specific new bundle*.

Please write down two goals for **you** to work on for the **rest of this week**; for example: *be braver in suggesting products*.

Goal 1:

Goal 2:

If you did the same exercise **last week**: were you successful in executing your goals? If no, why not?

Goal 1:

Goal 2:

Your Partner's Full name and [redacted]ID: _____

Please TALK to your partner about the questions below and write down their responses.

Ask **your partner** about their **most successful** sales call **last week,** what did they do right?

Ask **your partner** about their **least successful** sales call from the **last week**. What advice did **you** offer your partner?

Was **your partner** successful in accomplishing their goals **last week**? If no, why not?

Goal 1:

Goal 2:

What are **your partner's** two goals for **rest of this week**?

Goal 1:

Goal 2:

Please write down one thing **your partner recommended you** to try:

What day would you and our partner like to pick-up lunch from 12-2? (must hand in with 24 hour notice!): ☐ Thursday ☐ Friday

RXd by Adviser: date/time:

Lunch Talking Points

[redacted] Tournament 3 & 4

You do not need to turn this sheet in, put please read through it: is designed to help you make the most of the time with your partner.

1) *Bon Appetit!*

2) Have either of you had an awesome sale since you last met? What made it great?

3) Have either of you had a call go completely sideways? What happened? Does your partner have any advice?

4) Your partner gave you some advice on how to handle difficult stations earlier this week. Did it help?

5) You and your partner each made goals earlier this week, what progress have you each made on those goals?

6) If you have suggestions on how this lunch program could be more productive, please let your adviser know—we greatly appreciate your feedback.

Thank you!