

The Negative Consequences of Loss-Framed Performance Incentives[†]

By LAMAR PIERCE, ALEX REES-JONES, AND CHARLOTTE BLANK*

Behavioral economists have proposed that incentive contracts result in higher productivity when bonuses are “loss framed”—prepaid then clawed back if targets are unmet. We test this claim by randomizing the pre- or postpayment of sales bonuses at 294 car dealerships. Although somewhat statistically imprecise, our analysis provides strong indications that the random assignment of loss framing had quantitatively important negative effects. We document that the negative effects of loss framing can arise due to an increase in incentives for “gaming” behaviors. Based on these claims, we reassess the common wisdom regarding the desirability of loss framing. (JEL C93, D91, J24, J33, L62, L81)

Worker incentives commonly feature bonuses for exceeding a production target. At least since the work of Hossain and List (2012), behavioral economists have often advised that the benefits of such incentives can be raised by prepaying the bonus with a threat of clawback if the target is not met. If workers are loss averse and if prepayment leads workers to perceive failure to retain the bonus as a loss, then this intervention could induce an endowment effect (Kahneman, Knetsch, and Thaler 1991) that makes hitting the target seem more important. Greater effort and productivity could then follow.

In this paper, we report the results of a field experiment where a car manufacturer applied this advice to its dealer sales incentive program. In this field experiment, random assignment of the prepayment of bonuses did not result in the predicted increase in sales. Instead, we find a marginally significant and quantitatively meaningful *negative* effect. This finding is surprising when viewed through the lens of

*Pierce: Olin Business School, Washington University in St. Louis (email: pierce@wustl.edu). Rees-Jones: The Wharton School, University of Pennsylvania and NBER (email: alre@wharton.upenn.edu). Blank: Jaguar Land Rover NA (email: cblank1@jaguarlandrover.com). Erzo F.P. Luttmer was coeditor for this article. The authors are grateful to Maritz, LLC for making possible the natural field experiment in this paper and to Meghan Busse and Florian Zettelmeyer for help with design of the experiment. Pierce and Rees-Jones accepted no compensation from any party for this research. Blank served as Chief Behavioral Officer of Maritz, LLC while conducting this research. For especially helpful comments, we thank Greg Besharov, Stefano DellaVigna, Jon de Quidt, Alex Imas, Ian Larkin, and Ted O'Donoghue. We also thank seminar participants at Cornell University, George Washington University, Goethe University Frankfurt, INSEAD, the University of Mannheim, Purdue University, the University of Chicago, the University of Maryland, the University of South Florida, and Yale University, as well as conference participants at the AEA Annual Meetings, the Society for Judgment and Decision Making Annual Meetings, the Stanford Institute for Theoretical Economics Psychology and Economics Conference, and the Institutions and Innovation Conference. This study was ruled exempt from IRB review by the Northwestern University IRB (Study ID: STU00205062).

[†]Go to <https://doi.org/10.1257/pol.20220512> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

the conventional wisdom on the effects of loss aversion and when considered in the context of a purely rational model.¹ Our analysis suggests that this negative effect is at least partly influenced by the interaction between loss framing and the perverse incentives that arise in multitasking environments. Motivated by these empirical results, we theoretically explore the positive and negative consequences of loss framing and establish that an important negative consequence is the exacerbation of broadly defined “gaming” responses to incentives.

This project began when the car manufacturer became aware of the behavioral economics literature on loss-framed incentives. This literature, and its prescriptions, appeared directly relevant: the manufacturer provides car dealers with separate monthly targets for the sales of two groups of cars (which we will refer to as *model groups*) and pays a large bonus for exceeding each target. Based on the line of logic presented above, the manufacturer believed they might induce greater dealer effort by changing their existing postpayment scheme to a prepaid one. Toward that end, they recruited the authors of this paper to help implement the new bonus scheme.

Due to concerns about both the efficacy and desirability of loss framing, the authors recommended that this policy be tested in a randomized controlled trial (RCT) prior to implementation. While many “nudge” interventions are extremely low cost and may therefore be productively adopted even if their impact is low (Benartzi et al. 2017), this intervention differs: prepayment of incentives is in effect an interest-free loan, the cost of which is substantial given the scale of the bonus program. To directly test if the benefits of prepayment justify these costs, 294 car dealerships—with monthly sales of approximately 15,000 vehicles and monthly revenue of \$600 million—had \$66 million in bonuses randomized to pre- or postpayment conditions. Except for the difference in timing, financial incentives were identical across treatment groups.

In this RCT, the loss-framed prepayment scheme resulted in lower total sales. Our initial analyses are based on a difference-in-differences design in which we compare the sales path of pre- and postpaid dealers in the four months before and after randomization. During the treatment window, dealers assigned to prepayment sell 2.31 fewer cars, on average ($SE = 1.34$; $p = 0.086$), or 3.8 percent less in a Poisson specification ($SE = 2.2\%$; $p = 0.081$). These estimates are marginally significant but economically large. To illustrate, a 2.31 car sales decline per dealer per month would imply that the treated dealers lost \$45 million in revenue during the four-month treatment window. Forecasting further, this estimate would imply that adopting this policy network wide for a year would reduce revenue by over \$1 billion.

Our estimates suggest that the negative total sales effect of prepayment was driven by a decline in sales in one of the two model groups. This model group had substantially lower ex ante average sales volume, commensurate with a smaller bonus and thus a smaller potential loss. Moving forward, we will refer to this model group as the *small-bonus model group*. The model group with higher sales volume, which has a larger bonus and thus a larger potential loss, is estimated to have had a

¹ In a purely rational model, from the perspective of car dealers, the effect of early payment is simply the provision of additional liquidity. If liquidity constraints do not bind, this should be irrelevant for productivity, and if liquidity constraints do bind, then prepayment could enable productivity-enhancing investment.

near-zero and statistically insignificant change in monthly sales. Moving forward, we will refer to this model group as the *large-bonus model group*.

Models of loss aversion make additional predictions of *when* these contracts would most change behavioral responses. Most importantly, the behavioral response is naturally predicted to depend on proximity to the performance target. If loss framing leads a dealer to care more about exceeding the target, then a loss-framing intervention should have its strongest effects when a dealer is at risk of falling just short of its target. In contrast, a dealer who would otherwise substantially over- or undershoot its target should be minimally influenced by this intervention.

To assess these predictions, we present a novel modification of the standard difference-in-differences approach that estimates the impact of treatment on a distribution rather than a mean. This approach, which we call difference-in-kernel-density-differences (DKDD), estimates the location of missing or excess mass in a distribution by comparing the pre- and posttreatment distributions in the treatment group to the pre- and posttreatment distributions in the control group.

Applying this technique, we examine the effect of the switch to the prepaid incentives on the distribution of monthly sales of each model group. Consistent with our earlier estimate of an average reduction in the small-bonus model group's sales, we find that this model group's sales are generally shifted from higher to lower values. More interestingly, we find that the previously reported null effect on average sales for the large-bonus model group was masking substantial and statistically detectable changes to the distribution. Dealers became significantly less likely to end the month with large-bonus model group sales narrowly below the target. Instead of ending the month having just barely triggered a loss, some dealers increased their sales over the target. Other dealers decreased their sales, with some evidence that they attended to selling the small-bonus model group instead. Taken together, these two groups lead to offsetting average effects, demonstrating why the initial difference-in-differences estimates yielded a null effect on average large-bonus model group sales despite statistically detectable changes in their behavior.

The patterns of results that we document resemble common findings on the downsides of multitasking incentives. An organizing idea from this literature is that, when multiple dimensions are separately incentivized, agents may benefit—perhaps at the expense of a principal—by neglecting one dimension to attend to another dimension with higher private returns. This tension is present in the foundational model of Holmstrom and Milgrom (1991) and has been extensively studied in research in both policy and firm settings (see, e.g., Baker and Hubbard 2003; Dumont et al. 2008; Chen, Li, and Lu 2018; Kim, Sudhir, and Uetake 2021). However, to our knowledge, the manner in which these incentives interact with the loss-framing interventions suggested by Hossain and List (2012) has not been previously explored.² Our belief that the logic of this multitasking literature might explain the otherwise surprising

² A partial exception to this claim is Balmaceda (2018), who presents results on multitasking incentives in the expectations-based reference-dependence model of Köszegi and Rabin (2006). This model features loss aversion, but it does not have scope for a framing effect induced by prepayment because reference points are endogenously determined by expectations.

findings of our experiment led us to conduct a theoretical reexamination of the predicted role of loss framing.

To study this issue theoretically, we model a principal incentivizing an agent's acceptance of private costs of production through bonuses based on ex post performance, potentially across multiple measured dimensions. We model agents as having a set of possible production strategies available, with each strategy characterized by its induced joint distribution of dimension-specific production and private costs. In this environment, we study how the desirability of different strategies changes when loss framing is applied.

Strategies that are incentivized by loss framing are straightforward to characterize. Loss framing increases the perceived incentive to avoid strategies that are high in a particular measure of exposure to loss. This means that, in fully general environments, loss framing does not have an unambiguously signed effect on productivity. If there are strategies available that reduce this measure of exposure to loss while also decreasing expected productivity, the incentive to pursue these strategies increases. These strategies include, but are not limited to, standard ways of gaming a multidimensional contract.

Our model differs from those previously used to study loss-framed incentives in more ways than just including multiple dimensions. To facilitate comparisons to prior research, we next reevaluate the nested unidimensional case to see if the prevailing wisdom holds there. In general, it does not: we again find that the performance impact of loss framing remains ambiguous and determined by the measure of exposure to loss. However, we illustrate a special case of broad relevance in which loss framing can be understood to be unambiguously helpful: the case where the agent's choice of distributions of output induced by his strategy can be ordered by first-order stochastic dominance. When this property holds, loss framing is predicted to weakly increase expected productivity. When this property does not hold, effects are ambiguous. The key insight driving this distinction is that loss framing not only steepens incentives but also modifies risk tolerance. If actions can be taken that reduce the risk of a loss at the cost of expected production, loss framing incentivizes such actions. When outcome distributions are ordered by first-order stochastic dominance, there is no trade-off between loss probability and expected productivity, and thus this potential for perverse incentives is eliminated. When outcome distributions are not so ordered, this trade-off can become important.

In both the uni- and multidimensional cases, a simple intuition drives our results: inducing loss aversion through loss framing motivates loss avoidance. While this insight borders on tautology, notice that it is different from the typical claim that inducing loss aversion motivates effort. In cases where agents' only means to protect themselves from losses is to increase the distribution of productivity (whether through effort or other means), then loss framing is unambiguously predicted to steepen the incentives to take those actions. If, however, more elements of the agent's approach (such as risk tolerance or multitasking strategy) can be modified to influence the probability of losses, then steepening incentives will also influence those decisions. Gaming of incentive schemes through such modifications is typically viewed to be undesirable from the perspective of a principal, and this influence on incentives to game is the negative consequence of loss-framed performance incentives.

Our findings contribute to several lines of research. First, and most directly, our findings directly inform the literature on loss-framed performance incentives. This literature was largely initiated by Hossain and List (2012), who made a conceptual argument that loss framing should enhance the productivity of loss-averse workers and tested that prediction in a comparatively small field experiment. This paper motivated a substantial body of follow-up research from both the lab (see, e.g., Brooks, Stremitz, and Tontrup 2012; Imas, Sadoff, and Samek 2016; de Quidt et al. 2017; de Quidt 2017; DellaVigna and Pope 2018; Ahrens, Bitter, and Bosch-Rosa 2023) and the field (see, e.g., Hong, Hossain, and List 2015; Chung and Narayandas 2017; Brownback and Sadoff 2020; Bulte, List, and van Soest 2020, 2021; Fryer et al. 2022). As emphasized in the recent meta-analysis of Ferraro and Tracy (2022), a notable contrast exists between this lab and field literature: field tests on average suggest near-zero effects, whereas lab tests on average suggest large positive effects. Ferraro and Tracy (2022) document that this contrast may be at least partly attributed to publication bias and underpowered designs leading to an exaggerated average published effect among lab studies.

Our paper provides two critical inputs to this literature. First, we provide an unusually large-scale and unusually highly incentivized test of the impact of loss-framed performance incentives. This test yields perhaps surprising results, demonstrating previously undocumented potential for reliance on loss framing to be counterproductive. Second, we establish that the central prediction that is stated and tested within this literature—that loss framing increases productivity among loss-averse agents—is not unambiguously a prediction in relatively general models without further conditions. We provide more refined conditions that provide guidance on when such productivity gains are guaranteed and point to factors that could lead the common statement of the prediction to fail. Our theoretical findings may help further organize the different effects documented across settings. The lab settings that have most reproducibly found a positive effect of loss framing feature relatively uncomplicated effort decisions that seem particularly amenable to the assumption that strategies are ordered by first-order stochastic dominance. In contrast, the field settings studied examine more complex decisions, which likely involve more elaborate trade-offs between risk of loss and productivity. While we do not believe that this is the only explanation for the contrast, we do believe that it contributes.

Our paper also contributes to the broader literature on applications of behavioral economics in policy settings. In recent years, there has been a great deal of interest in harnessing behavioral economics in general, and loss framing specifically, as a policy tool. As examples, loss framing interventions have been deployed with the goal of influencing tax compliance (Hallsworth et al. 2017), teaching effort (Fryer et al. 2022), or disposable bag use (Homonoff 2018). Despite some success stories, studies of large-scale policy interventions of this type are still relatively rare, and results on the efficacy of these interventions are mixed. A reasonable reader of this literature can still question whether these interventions can be deployed “at scale” with expectations of meaningful behavioral response. We provide two inputs to this debate. First, we document the deployment of a loss-framing intervention at a scale rarely attempted in the behavioral economics literature and find that it did lead to meaningful behavioral response. While ultimately the induced behavioral response

was undesirable, the fact that behavioral economic factors were activated suggests hope for harnessing them for policy purposes. Second, our study provides a rare demonstration of such interventions having a behavioral economic impact when the targeted units are *firms* and not *individuals*. We return to further discussion of these contributions in the conclusion.

The remainder of this article proceeds as follows. In Section I we describe the setting, design, and results of our field experiment. In Section II we present a theoretical examination of the positive and negative consequences of loss-framed incentives. Section III concludes, and the Supplemental Appendix contains a variety of additional analyses.

I. An RCT on Loss Framing in Automobile Sales

In this section, we present our field experiment. We begin by providing institutional details necessary for understanding the market, followed by a detailed description of the experiment that was deployed. We then analyze the effect of prepayment on average sales before proceeding to examine its effect on the distribution of sales.

A. Setting

Our empirical setting is the new automobile market—a market governed by extensive contracting between automobile manufacturers and automobile dealers.

State laws in the United States significantly restrict manufacturers from directly selling vehicles to consumers, effectively forbidding vertical integration. As a result, manufacturers sell vehicles to independent dealers, who then sell or lease vehicles to consumers. Dealers are organized into designated market areas (DMAs) such as St. Louis, Missouri, or Philadelphia, Pennsylvania, within which multiple independent dealers typically compete.³ The sales activity of these dealers is our focus in this paper.

Independent dealers operate under franchise agreements that significantly shape the automobile market. Existing franchise laws restrict manufacturers' ability to open or close dealers, heavily limiting manufacturers' ability to control interbrand competition or react to changes in demand. Such laws also highly limit the ability of manufacturers to treat dealers differently, limiting the means by which they may present tailored sales incentives through dealer-specific prices (and more broadly limiting ability to price discriminate).

Although many of the approximately 17,000 new vehicle dealers in the United States are now being consolidated by private equity and publicly traded corporations such as Autonation and Lithia, the industry has historically been fragmented for several reasons (Roberts 2018). First, manufacturers have typically blocked dealer sales that put multiple competing dealers under common ownership. Second, many franchise laws prohibit owning more than a certain number of dealerships per state. With the lack of entry and exit and the limits on consolidation, the dealers for any

³ See Lafontaine and Scott Morton (2010) or Murry and Schneider (2016) for detailed information on the history and function of car dealership franchises.

given manufacturer represent a diverse mix of entities organized as public and private national corporations, private equity, family-owned dealers and groups, and other independently owned dealers. Across these disparate players, there are substantially varying managerial approaches and degrees of efficiency.

As a consequence of this market structure, manufacturers must use somewhat indirect means to influence dealers' sales incentives. Franchise laws discussed above constrain manufacturers to sell vehicles to dealers for a fixed invoice price. Dealers then have discretion to price the vehicles independently before selling to consumers at a negotiated price (Busse and Silva-Risso 2010; Bennett 2013). Since manufacturers rarely change published retail and invoice prices within a given year, other mechanisms are used to affect prices and quantities within their dealer network. Manufacturers provide cash incentives to both dealers and customers in order to reduce vehicles' effective price if there is excess inventory (Busse, Silva-Risso, and Zettelmeyer 2006). Similarly, they subsidize loans and leases through captive finance arms (Pierce 2012). To motivate sales volume, manufacturers typically use direct incentive programs that reward units sold. These programs are intended to partially address misaligned incentives for dealers to sell lower volume at higher prices, since the manufacturer benefits from volume but cannot capture value from higher prices in a given year (Busse, Silva-Risso, and Zettelmeyer 2006).⁴ Volume-based incentives also directly promote increasing the manufacturer's market share, which is valuable for preserving the manufacturer's competitive position within the industry and is directly rewarded in executive compensation (Ritz 2008; Pierce 2012). These volume-based incentive programs are the policies of interest in this paper.

The car manufacturer that we study has multiple vehicle models sold at over a thousand dealers across all 50 states. Since our access to data is governed by a nondisclosure agreement, we will refer to this manufacturer as "CarCo."⁵ Like many manufacturers, CarCo uses incentive programs for specific vehicle models to motivate sales volume at the majority of its dealers. The specific program considered in this paper has two separate incentive schemes, each focused on a particular set of models. We refer to these sets of models as the large-bonus model group and the small-bonus model group, as described in the introduction. Both groups are typically, but not always, retailed together at the same dealerships: 88 percent of dealers in the 2017 program sold both model groups. Dealers in the focal incentive program sell over 90 percent of all vehicles in the two model groups.

Like most dealer incentive programs, CarCo's program rewards dealers for reaching monthly targets.⁶ These targets are set for each model group by taking average sales in that calendar month for the previous four years. All vehicles in a specific

⁴This incentive misalignment is partially mitigated by the profits from servicing cars they have sold, which is ultimately influenced by sales volume.

⁵Our legal agreement allowed CarCo to correct the paper for factual accuracy and anonymity, but we had full rights to publish regardless of our results and conclusions.

⁶For an illustrative example of how these shape policies and behavior at one dealer (not necessarily a CarCo franchisee), listen to Episode 513 of *This American Life* at <https://www.thisamericanlife.org/513/129-cars>. We note that supply and demand disruptions due to the COVID-19 pandemic led to the suspension of many of these quota-based systems.

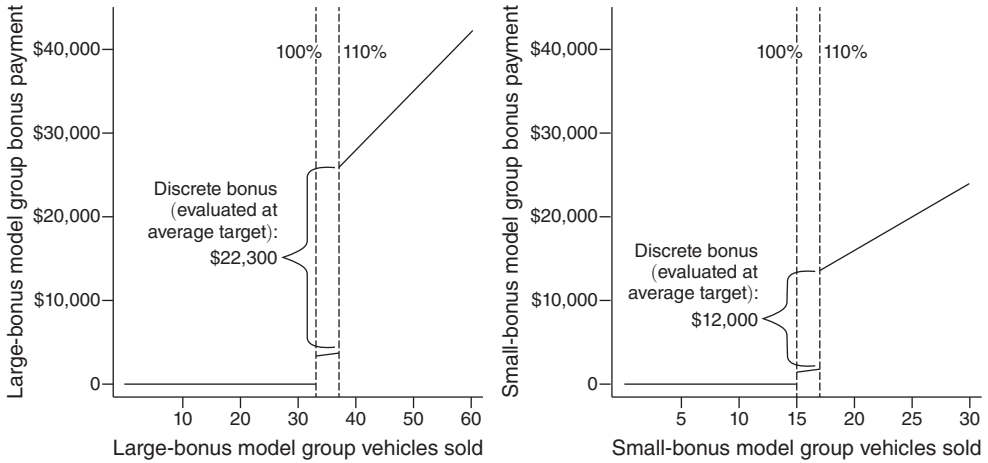


FIGURE 1. DEALER INCENTIVE PLAN FOR THE TWO MODEL GROUPS

Notes: This figure shows the bonus structure for both model groups. Each monthly model-group bonus is earned independently. Targets are set based on the dealer's average sales over the previous four years in that calendar month, determining the "100 percent" target indicated in the figure. If sales exceeded 110 percent of that target, a large fixed bonus is conveyed along with additional marginal incentives. To help illustrate the typical application of this bonus system, we have plotted this figure applying the average targets taken from Table 1.

model group that are sold to consumers count toward the target equally.⁷ For example, a given dealer's April 2018 target would be the average of their April sales from 2014 through 2017 rounded up to the nearest integer. Quota-based systems like these are common across industries, with variation in the number of targets and the time period allowed to reach those targets (Chung, Narayandas, and Chang 2019; Misra and Nair 2011; Chung, Steenburgh, and Sudhir 2014).

Figure 1 illustrates CarCo's incentive program. Dealers selling below their monthly target earn no bonus. Dealers selling between 100 percent and 110 percent of their monthly target receive \$100 for each vehicle sold in that month. Dealers selling 110 percent or more receive \$700 per vehicle for the large-bonus model group or \$800 per vehicle for the small-bonus model group. The higher per vehicle bonus in the small-bonus model group primarily reflects slowing sales for those models, but it also serves to counteract the group's smaller discrete bonuses at 110 percent based on lower average sales volume. Bonus qualification operates separately for each model group, such that a dealer could qualify for one or both group-specific bonuses. Because the per vehicle bonuses apply to *all* vehicles from that group sold in that month—not just marginal vehicles over the target—exceeding the 110 percent threshold conveys a large fixed bonus.⁸ To illustrate, applying the average targets reported in Table 1, selling the single vehicle on the margin of the 110 percent

⁷Historically, certain models have received "double points" toward the target, but this did not happen during our time period.

⁸Interviews with dealers consistently indicated that they view the 100 percent goal as mostly irrelevant—or a small consolation prize for missing the key 110 percent level.

threshold would yield \$22,300 or \$12,000 in a given month in the large-bonus and small-bonus model groups, respectively. For the largest dealers, this marginal car can be worth over \$200,000.⁹

Participation in this program is voluntary and requires payment of a fixed monthly fee, but the majority of dealers choose to participate.¹⁰ The almost universal participation in this program is largely due to the often intense price competition between dealers in the same DMA. As one dealer explained, a competitor will price below invoice, hoping to make all their profit off the monthly incentive program. A nonparticipating dealer thus cannot compete on price without losing money.

Although exogenous variation in demand plays a large role in target achievement, car dealers have multiple ways in which they can increase sales that can be broadly categorized as “effort.” First, they can increase managerial oversight or incentives for salespeople. Second, they can increase their advertising spending. Third, they can price more aggressively, accepting lower (and at times negative) margins to close deals. Fourth, they can work more aggressively to qualify buyers for financing, which effectively lowers the vehicle price.¹¹

Dealers also have several ways in which they can react to monthly sales targets that involve gaming the incentive system (Larkin 2014; Oyer 1998; Courty and Marschke 2004). First, dealers can attempt to move customers across calendar months (or direct their salespeople to do so). If the dealer is near the crucial 110 percent threshold, they can attempt to accelerate the purchase decision of a customer by offering a lower price or prioritizing the paperwork and financing. If a given sale is irrelevant for target attainment, dealers can attempt to delay its closing until after the month-end so it helps toward next month’s target attainment. Second, dealers can potentially influence the types of cars they sell by either directing customers to particular models or prioritizing customers with a preference for particular models. Dealers may attempt to sell specific cars within a model group based on expected ease of sale (to earn volume bonuses) versus their markup (to yield direct profits). They may also seek to advance one entire model group over the other based on the relative likelihood of reaching their discrete 110 percent bonuses in that month. Dealers rarely focus all attention on selling one model group from the beginning of the month because the relative likelihoods (and thus expected values) of reaching the targets are not immediately clear. Customers may strongly prefer one model over others, so pushing them hard to purchase from the other model group risks losing a

⁹Such dramatic discontinuities are rarely features of optimal incentives in common related models, and as a result, many economists view these contracts as surprising. We note, however, that both Herweg, Müller, and Weinschenk (2010) and Gill and Stone (2010) offer cases where highly discontinuous or even binary incentive contracts can be rationalized among reference-dependent agents. In practice, the popularity of contracts like these are often driven by the relative ease of assessing incentives. In order to perceive strong incentives for effort, an agent merely needs to assess whether they have reached a salient sales benchmark. In contrast, qualitative incentives are substantially more complex to evaluate in the nonlinear contracts commonly arising in economic theory. Some practitioners view avoiding this complexity to be a major desideratum of contract design.

¹⁰Most nonparticipants are very small dealers. The program contract operates on an annual basis, such that the incentive structure or features cannot be changed within a calendar year. CarCo will occasionally offer optional features (e.g., model-specific bonuses) within the year that some dealers will choose to accept.

¹¹Note that this final option is not desirable if it exposes a captive lender to undue default risk because of misrepresentation of creditworthiness or if it ruins the credit of customers with strong brand loyalty, thereby barring them from future new car purchases (Jansen et al. 2023).

sale to other dealers. Dealers also are sensitive to CarCo's ability to reallocate inventory away from dealers that ignore one particular model group.

Although these gaming behaviors might mitigate the risk of missing the large payoffs from hitting thresholds, they can be costly to the manufacturer due to their risk of driving customers to other manufacturers' vehicles. More directly, increased gaming is undesirable for the manufacturer because it aims to increase bonus payments without necessarily increasing sales volume. Like Oyer (1998) and Larkin (2014), we note that the marginal cost of increased gaming does not imply that the incentive structure is necessarily suboptimal since the motivational benefits from the convex structure may justify gaming costs.

B. Experimental Design

The experiment in this paper originated with a CarCo brand manager's interest in using loss framing to increase the efficacy of their dealer incentive program. Several CarCo executives had been discussing launching a version of their incentive program in which they manipulated the timing of bonus payments to induce loss framing. Monthly bonus payments had historically been paid in the month following the incentivized sales, allowing the program manager time to validate the qualification of each vehicle reported as sold. As an alternative, CarCo considered giving dealers a large up-front bonus each month, then clawing back any unearned money at the end of the month. CarCo leaders expected this front-loaded payment to be both highly appealing and motivational for dealers, who would appreciate the advanced cash flow and work hard to retain it. The brand managers thus proposed to change the incentive plan to universally implement advanced payments.

The researchers, working in conjunction with the brand managers, proposed an RCT that would guide whether this modification should be deployed. Control dealers would keep their existing postpayment system. Treated dealers would be prepaid the bonus associated with exactly meeting their 110 percent sales target. If this sales target were not achieved by the end of the month, the difference between their earned and prepaid bonus would be clawed back. If the sales target were instead exceeded, the dealer would receive an additional bonus payment to transfer the excess earned incentives.

Although there were reasons to believe loss framing might improve performance, several considerations raised doubts. First, the goal of advanced payment is to motivate additional marginal effort to hit the monthly target. However, given the existing extreme motivations to hit the target, it was possible that most cost-effective marginal activities were already incentivized. Second, there were concerns that increased incentive gaming might substitute for effort at the margin. Incentive gaming is widely recognized as a concern in contracts of this nature (Oyer 1998; Steenburgh 2008; Larkin 2014; Benson 2015; Ederer, Holden, and Meyer 2018), particularly those with nonlinear returns to performance. Finally, given the heterogeneity of organizational forms, management style, and sophistication, it was unclear how many dealers would respond "behaviorally." If enough firms viewed prepayment as free capital with no associated motivational effects, and if that free

capital did not enable productivity-enhancing investments, then the loss-framed contracts would be inherently costly to CarCo.

Due to fairness concerns, CarCo wished to avoid having experimental participants in direct competition facing different treatment assignments. This led us to block-assign treatment by DMA—a classification of geographic areas created by Nielsen that classifies local media markets (and thus regions for advertising). Figure 2 presents a map of DMAs in the lower 48 states, their treatment assignments, and the dealers within them.

Anticipating our use of a difference-in-differences design, we randomized the treatment status of DMAs in a manner that guarantees similar pre-trends in sales. We first calculated DMA-specific sales trends using logged unit sales data between January 2014 and February 2017. We followed Athey and Imbens (2017) in stratifying by region and DMA size and then using a bipartite matching procedure to generate the set of matched pairs that minimized the sum of distances between mate sales trends (Lu et al. 2011). Mates were then randomly assigned to the two conditions.

Several institutional features influenced our research design. First, again guided by fairness considerations, we were required to ensure that each participating dealer spend equal time in each experimental condition. Consequently, participants were randomly assigned to treatment or control for an initial four-month treatment window, and then conditions were flipped for a second four-month treatment window. We believe that the comparison between these two groups in the initial four months provides a clean estimate of treatment effects. In contrast, the comparison between groups after treatments are flipped may be influenced by their prior experimental assignment. In the text of this paper, we focus attention on treatment effects estimated in the initial treatment window, and when we refer to the “treatment” or “control” group, we refer to assignment in that period. Analysis of the second window is relegated to the Supplemental Appendix (but provides similar evidence of negative treatment effects).

Second, the franchise agreement required dealer approval for any change to the incentive program. We therefore invited dealers to opt in to participation after assigning DMAs to condition, but without revealing that assignment to the dealers. Dealers were initially invited to participate via postings on the program tracking website. The postings described the program and asked them to either opt in or opt out of the pilot program. The posting explained that participating dealers would be randomly assigned to receive prepayment for four months starting either in May 2017 or September 2017. Dealers opting out would remain in the existing postpayment system. Multiple reminders were posted before CarCo regional sales directors attempted to directly contact the dealers in person or by phone. Of the 1,227 dealers in the incentive program, only the 294 (24 percent) that chose to participate were included in the experiment and subsequent analysis. Additionally, 336 explicitly opted out of the program, while 597 failed to respond. Interviews with CarCo managers and dealerships indicated that many nonparticipants did not want the accounting hassle of prepayment.

We present extensive analysis of selection into the study in Supplemental Appendix A.3, comparing these groups with both CarCo data from before the experiment and dealer-level data from the National Establishment Time Series (NETS). NETS

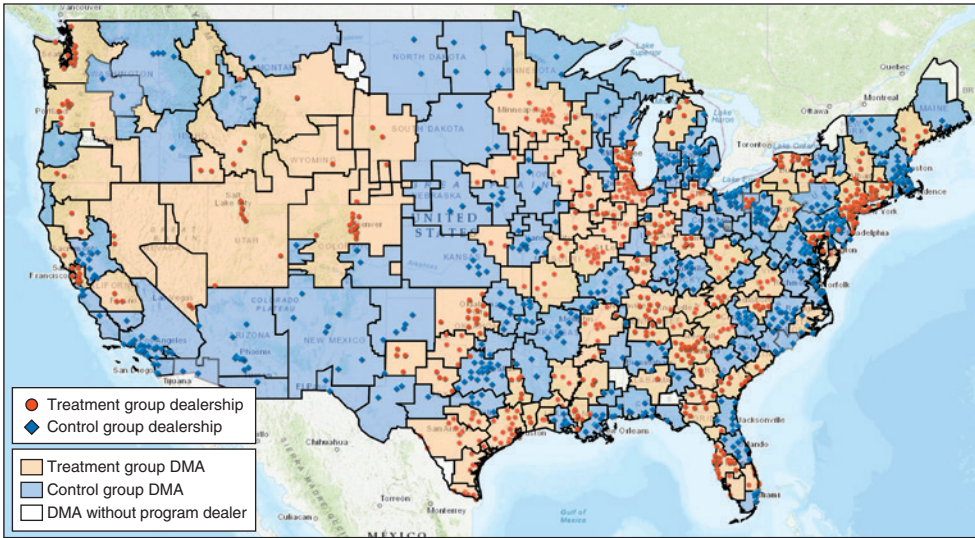


FIGURE 2. BLOCK ASSIGNMENT OF DEALERS BY DMA

Notes: This figure shows all program dealers in the lower 48 states with their block-assigned treatment group. Clear spaces reflect DMAs without program dealers.

data show no statistically identifiable differences among the three groups in dealer age, employee count, creditworthiness or liquidity, ownership structure, or membership in a dealer group. Participants and opt-out dealers were statistically indistinguishable in average monthly sales (48.5 versus 43.2, $p = 0.16$), but nonrespondents were considerably smaller than both (31.1, $p < 0.01$). Nonrespondents were less likely to carry both model groups (86 percent) than participants (90 percent) and opt-out dealers (93 percent). Participation was markedly different across regions ($\chi^2 = 101.89$, $p = 0.000$), which CarCo attributed largely to regional sales representatives using different follow-up strategies to encourage participation after the standardized initial email.¹² Overall, selection into participation based on observables is modest, and due to our experimental design, selection does not threaten internal validity. However, the potential for selection on unobservables remains—for example, based on anticipation of susceptibility to loss framing or on self-assessments of elasticity to incentives—and thus caution is warranted when extrapolating our findings to new populations.

Finally, CarCo required us to change the group assignment of two DMAs due to concerns about competition spreading between neighboring DMAs. Recall that fairness concerns motivated the DMA-level blocking of treatment to avoid close competitors being assigned to different treatments. Upon observing treatment assignments, CarCo representatives identified two pairs of neighboring DMAs where they believed cross-DMA competition was a concern and where treatment assignments were

¹²There are no differences in treatment group assignment by region ($\chi^2 = 2.46$, $p = 0.652$).

different. To address this concern, one of the DMAs within each pair of neighbors had its treatment assignment exchanged with its DMA mate from the bipartite matching procedure. The net result is four DMAs having their treatment assignment modified.

While treatment reassignments have the potential to confound experiments, several considerations suggest that this instance is comparatively benign. First, note that CarCo's intervention was not to ensure or prohibit treatment from being deployed to specific DMAs or dealers. All experimental participants experienced four months under prepaid treatment and postpaid control with only the timing randomized, and CarCo's intervention was to ensure that the *timing* of treatment was the same across these particular neighbors. Second, because we flip assignment of the target DMAs and their mates in the bipartite match, the potential of this flip to proxy for some feature of the DMA is mitigated to the extent that our bipartite match mates are similar. Given these considerations, we believe these flips should be treated analogously to a constraint in the randomization procedure that guarantees the same treatment assignments for these neighbors, as compared to a more selection-inducing form of experimental noncompliance.¹³

Based on this construction, our experimental sample consists of 294 dealers representing 116 DMAs in 47 states. Of the 294 dealers, 140 from 55 DMAs were assigned to the treatment group, while 154 from 61 DMAs were assigned to the control group.¹⁴ All but 29 of these dealers sold both model groups: 18 sold only the large-bonus model group, and 11 sold only the small-bonus model group. Collectively, these dealers represent monthly sales of over 15,000 vehicles and \$600 million in revenue.

The experiment initiated on May 1, 2017, with all treatment group dealers receiving their first-month bonus prepayment. The treatment group received an average total monthly prepaid bonus of \$37,395, with averages of \$25,077 and \$13,330 for the large- and small-bonus model groups, respectively.¹⁵ For those dealers who failed to meet the 110 percent target and therefore suffered clawbacks, those clawbacks averaged \$28,306, with averages of \$26,290 and \$13,564 for the large- and small-bonus model groups, respectively.

Despite researcher concerns, no dealer during our study was unable to repay unearned prepayments due to overspending and insufficient cash on hand. In addition, no dealer asked to be removed from the program.

C. Response to Pretreatment Incentives

Before proceeding to an analysis of the effects of our program, we document dealers' behavior in the presence of the baseline, postpaid incentive scheme.

¹³ For completeness, in Supplemental Appendix Section A.5, we reproduced our main analyses while dropping the reassigned DMAs (i.e., the 2 identified by CarCo and their match mates). Results are broadly similar, although the statistical significance of our pooled estimates is weakened. Comparatively statistically strong results remain for the small-bonus model group (with *p*-values for the OLS and PPML specifications of 0.059 and 0.035, respectively). And the pattern of results found in the DKDD analysis is essentially identical (as would be expected given that so few observations are being removed from the distributions used in this procedure).

¹⁴ See Supplemental Appendix A.1 for the density of dealers within DMAs across conditions.

¹⁵ The two model group averages do not sum to the total average because some dealers sell only one model group.

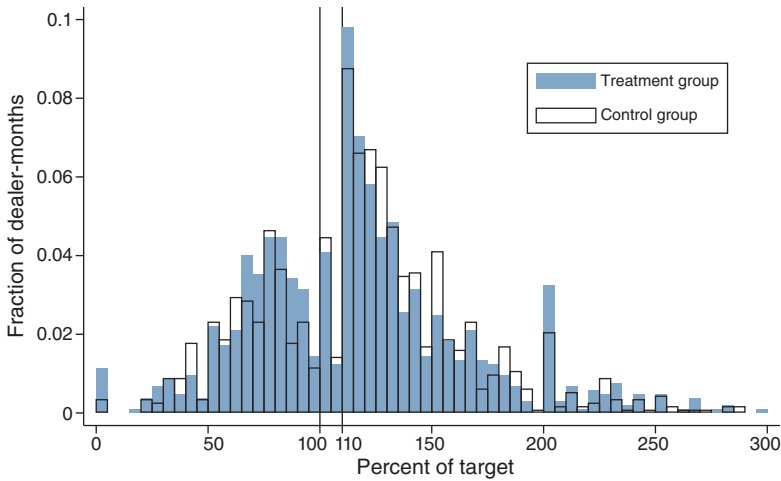


FIGURE 3. PRETREATMENT MONTHLY SALES (RELATIVE TO TARGET) BY TREATMENT ASSIGNMENT

Notes: This figure shows the distribution of monthly sales in the four months prior to the experiment, expressed as a percentage of the assigned sales target. Model group results are included separately for each dealer-month. The two vertical lines represent the discrete bonus thresholds at 100 percent and 110 percent.

Figure 3 shows the distribution of each monthly dealer sales outcome as a percentage of its respective target in the four months prior to the experiment. Dealers that sell both model groups have each represented separately. This figure demonstrates the importance of the 110 percent target in driving sales. The incentive contracts shown in Figure 1 feature a large discontinuity in both levels (a “notch”) and slopes (a “kink”) occurring at the target, with both features leading to a prediction of excess mass in the vicinity of the target.¹⁶ Such excess mass is starkly visible in this figure, with substantial asymmetries in mass observed around both the 100 percent and 110 percent target implying elastic response to incentives. Note that, in the pre-period data examined in this figure, there is no statistically detectable difference in either the average or the distribution of the plotted variable across treatment and control groups (t -test: $p = 0.702$, Kolmogorov-Smirnov: $p = 0.122$). This suggests pretreatment balance on sales relative to targets (despite some imbalance on absolute sales that we discuss in the next section).

D. Summary Statistics

We now begin an analysis of the sales activities in both the pretreatment and posttreatment windows. Table 1 provides descriptive statistics on the monthly number of vehicles sold, target levels, and the rate of attaining given targets. While our ultimate analyses will be somewhat more sophisticated than the mere comparisons of means presented in this table, our results are foreshadowed by such comparisons.

¹⁶See Kleven (2016) for a thorough review of bunching-based identification strategies.

TABLE 1—SUMMARY STATISTICS

	Treatment group			Control group			By participation status		
	Pre	Post	Total	Pre	Post	Total	In	Out	All
<i>Large-bonus model group</i>									
Vehicles sold	29.52 (21.44)	31.41 (23.72)	30.46 (22.61)	39.39 (52.61)	41.06 (52.13)	40.23 (52.35)	35.49 (41.03)	27.23 (29.18)	29.24 (32.65)
Target sales	26.98 (19.36)	32.15 (22.03)	29.57 (20.89)	31.45 (33.37)	38.59 (41.69)	35.03 (37.92)	32.38 (30.98)	27.24 (24.64)	28.49 (26.41)
Hit 110% target	0.56 (0.50)	0.43 (0.50)	0.50 (0.50)	0.65 (0.48)	0.50 (0.50)	0.57 (0.49)	0.54 (0.50)	0.46 (0.50)	0.48 (0.50)
Hit 100% target	0.62 (0.49)	0.48 (0.50)	0.55 (0.50)	0.70 (0.46)	0.55 (0.50)	0.62 (0.49)	0.59 (0.49)	0.53 (0.50)	0.54 (0.50)
<i>Small-bonus model group</i>									
Vehicles sold	15.14 (14.07)	14.16 (13.47)	14.65 (13.78)	17.22 (20.64)	18.80 (25.19)	18.01 (23.03)	16.37 (19.16)	11.07 (11.68)	12.35 (14.04)
Target sales	12.39 (9.91)	14.69 (11.92)	13.54 (11.02)	13.81 (13.37)	16.86 (17.00)	15.34 (15.36)	14.46 (13.45)	10.88 (9.90)	11.74 (10.96)
Hit 110% target	0.64 (0.48)	0.44 (0.50)	0.54 (0.50)	0.64 (0.48)	0.48 (0.50)	0.56 (0.50)	0.55 (0.50)	0.48 (0.50)	0.50 (0.50)
Hit 100% target	0.70 (0.46)	0.50 (0.50)	0.60 (0.49)	0.70 (0.46)	0.53 (0.50)	0.62 (0.49)	0.61 (0.49)	0.56 (0.50)	0.57 (0.49)

Notes: Summary statistics of monthly sales performance by model group, treatment, and participation status. Means and standard deviations presented in all cells. The first two rows present monthly sales and monthly target thresholds. The rows below summarize the probability of earning the larger fixed bonus for exceeding the 110 percent threshold and the smaller fixed bonus for exceeding the 100 percent threshold. The seemingly identical pre-period small-bonus model group percentages for the two participant groups is due to rounding, and they are slightly different at the third decimal point.

Turning attention first to the large-bonus model group data for monthly vehicles sold, we note that the control group saw a modest increase in sales when comparing the pre- and posttreatment windows (39.39 versus 41.06, an increase of 1.67). These control group increases primarily reflect seasonal demand fluctuations that are also reflected in the higher targets. For comparison, the treatment group saw a slightly larger but roughly similar increase (29.52 versus 31.41, an increase of 1.89; difference-in-differences: 0.22). An analogous comparison of the rate of achieving the 110 percent target suggests that treatment increases the chance of earning the large fixed component of the bonus by 2 percentage points.¹⁷ In short, simple comparisons of means suggest that the treatment was associated with quite modest increases in the large-bonus model group's sales and bonus-condition attainment.

Turning attention next to the small-bonus model group data for monthly vehicles sold, we note that the control group again saw a modest increase in sales when comparing the pre- and posttreatment windows (17.22 versus 18.80, an increase of 1.58). For comparison, the treatment group saw a decrease in sales (15.14 versus 14.16, a decrease of 0.98; difference-in-differences: −2.56). An analogous comparison of the rate of achieving the 110 percent target suggests that treatment decreases the

¹⁷Control: 0.65 versus 0.50, a decrease of 0.15. Treatment: 0.56 versus 0.43, a decrease of 0.13. Difference-in-differences: 0.02.

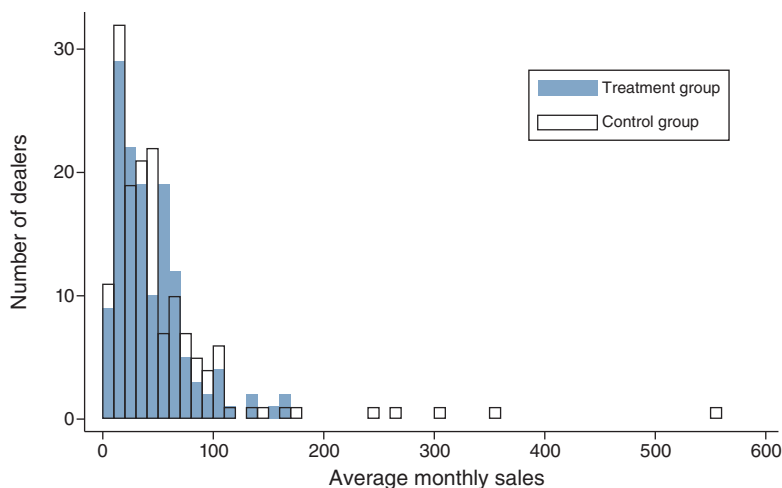


FIGURE 4. PRE-PERIOD AVERAGE MONTHLY CAR SALES BY TREATMENT GROUP

Notes: This figure shows a histogram of average monthly sales for each dealer (summing across both model groups) in the four months prior to the experiment. Five of the six largest dealers are in one DMA (and thus are block assigned). This block assignment accounts for the unbalanced right tail of control group observations.

chance of earning the large fixed component of the bonus by 4 percentage points.¹⁸ In short, a simple comparison of means suggests a comparatively large negative effect of loss framing for the small-bonus model group's sales and bonus-condition attainment.

Examination of Table 1 also illustrates two key issues that will prompt robustness analyses.

First, even prior to treatment, the means of key variables (e.g., sales) are significantly different between the treatment and control groups. These differences may generate some concern about systematic problems with random assignment, but they can be attributed entirely to the presence of a small number of influential outliers in the control group. Five of the six largest dealers, who sell over five times the monthly volume of the average dealer, are in one DMA and thus block assigned. The exclusion of this one DMA renders the two groups statistically indistinguishable in pre-period monthly sales. The largest dealer, which has ten times the monthly sales of the average dealer, was by chance also assigned to the control group. Figure 4 shows the distribution of average dealer monthly sales in the window prior to experimental intervention, illustrating broadly similar distributions with the exception of these outliers. These issues partially motivate the inclusion of dealer fixed effects in our main analyses. Additionally, in Supplemental Appendix Section A.4, we reproduce our main results while excluding this outlier DMA and its matched mate and show that this exclusion minimally affects our reported results.

¹⁸Control: 0.64 versus 0.48, a decrease of 0.16. Treatment: 0.64 versus 0.44, a decrease of 0.20. Difference-in-differences: -0.04.

Second, sales and target attainment show clear trends in the control group. This nonstationarity is expected due to the significant seasonality in automobile sales and illustrates the importance of our difference-in-differences design.

E. Impact of Loss Framing on Average Sales

In this section, we examine the average effect of loss framing during the first four months of the experiment, when the 140 treatment group dealers received the prepayment treatment while 154 control group dealers retained the preexisting postpayment scheme. Sales behavior in the four months prior to experimental assignment serves as the pre-period in our estimation data.

We consider two regression specifications:

$$(1) \quad E[Y|d, m, Treated_d, Treatment\ period_m] \\ = \beta \times Treated_d \times Treatment\ period_m + \mu_m + \nu_d$$

and

$$(2) \quad \ln(E[Y|d, m, Treated_d, Treatment\ period_m]) \\ = \beta \times Treated_d \times Treatment\ period_m + \mu_m + \nu_d.$$

Regression specification (1) is a standard OLS implementation of a difference-in-differences design.¹⁹ The conditional expectation of the total number of cars sold by dealer d in month m is determined by the interaction between an indicator of whether the dealer was assigned to the treatment group ($Treated_d$) and an indicator for whether the current month m is in the window when treatment is active ($Treatment\ period_m$), month-specific fixed effects (μ_m), and dealer-specific fixed effects (ν_d). In this framework, β measures the treatment effect of interest in units of additional cars sold.

Regression specification (2) is a standard Poisson regression implementation of a difference-in-differences approach.²⁰ As illustrated in the equation, the same set of predictors and fixed effects are included. The key distinction is that, in this framework, coefficient β measures the treatment effect in log units of sales. When discussing these results, we will rely on the approximation that $\ln(1 + \beta) \approx \beta$ for small β and will thus convert the log-point estimates into effects expressed in percentages.

Throughout our analyses, all standard errors are clustered at the DMA level.

Table 2 presents these results. Columns 1 and 2 present results for total sales across both model groups, where the first column reports the OLS specification and the second column reports the Poisson specification. Columns 3 and 4 and 5 and 6

¹⁹ Note that the necessary parallel trends assumption for this model is supported by the stratified matching randomization in our experimental design, and it is visually presented in Supplemental Appendix A.2.

²⁰ We apply Poisson regression rather than OLS with a logged dependent variable because some dealerships have months with zero sales.

TABLE 2—DIFFERENCE-IN-DIFFERENCES ESTIMATES OF TREATMENT EFFECT ON SALES

	OLS (1)	PPML (2)	OLS (3)	PPML (4)	OLS (5)	PPML (6)
<i>Treated_d × Treatment period_m</i>	−2.31 (1.34)	−0.038 (0.022)	0.14 (0.94)	0.019 (0.025)	−2.55 (1.40)	−0.154 (0.060)
Model group	Pooled	Pooled	L-B	L-B	S-B	S-B
Month fixed effects	X	X	X	X	X	X
Dealer fixed effects	X	X	X	X	X	X
Observations	2,352	2,352	2,261	2,261	2,216	2,216

Notes: This table presents regressions predicting the dealer/month-specific number of vehicles sold. Odd-numbered columns present OLS estimates of equation (1). Even-numbered columns present Poisson Pseudo-Maximum Likelihood (PPML) estimates of equation (2). We first present an analysis pooling the two model groups together (columns 1 and 2), followed by an analysis of large-bonus model group (columns 3 and 4) and small-bonus model group (columns 5 and 6). Standard errors are clustered at the DMA level.

present these same regressions restricted to the large-bonus and small-bonus model groups, respectively.

The average treatment effect in column 1 implies that 2.31 fewer cars were sold per month under loss framing. Note that, due to the comparatively few clusters in our analysis arising from DMA-level block assignment, these estimates are not extremely precise. Despite this imprecision, the null hypothesis of no effect can be rejected at the 10 percent but not the 5 percent α -level ($p = 0.086$). The Poisson model in column 2 produces similar results, with a 3.8 percent decrease in sales and similar statistical significance ($p = 0.081$).

We next examine the differential effect this policy had on the two model groups under consideration. Note that in columns 3 and 4 of Table 2, we find no strong indications of an effect of the policy on sales for the large-bonus model group. Point estimates suggest an impact of 0.14 car sales per month, or 1.9 percent, but effects are far from any common significance thresholds ($p = 0.882$ and $p = 0.463$, respectively). In contrast, the estimated treatment effects for the small-bonus model groups are statistically stronger and quantitatively large: −2.55 cars sold per month ($p = 0.071$) or −15.4 percent ($p = 0.010$).

In light of the standard prediction that loss-framed contracts motivate additional productivity, these findings may be viewed as surprising. There is little indication of this positive effect, and indeed all but extremely modest positive effects fall outside of the confidence intervals of our pooled estimates.²¹ Instead, our estimates tentatively point toward overall *negative* impacts of loss framing, although with substantial imprecision in our estimates and significance only at the lower-than-usual 10 percent α -level. Our estimates additionally suggest that the overall negative effect on joint sales is driven by a large reduction in small-bonus model group sales. Large-bonus model group sales, in contrast, saw a small and not meaningfully off-setting gain. This pattern will be more fully explored in the following section.

²¹ Based on our estimates from columns 1 and 2, we may reject any positive treatment effect in excess of 0.33 cars a month or 0.47 percent of sales at a 5 percent α -level.

F. The Impact of Loss Framing on Distribution of Sales

Having assessed the impact of loss framing on average sales, we now turn to assessing its impact on the distribution of monthly sales. The pre-experiment baseline distribution represented in Figure 3 shows that under the original postpayment scheme, dealer attention toward the 110 percent target was reflected in missing sales density just below the target and excess density above. Importantly, it also showed indistinguishable distributions between the two experimental groups. Our goal now is to evaluate how these distributions were affected by loss framing. By analyzing the full distribution, we may assess more nuanced predictions regarding the impact of loss framing: most centrally, that changes in behavior induced by the treatment should occur among dealers “close” to hitting the monthly target.

DKDD Estimation.—In order to infer the patterns of excess and missing mass induced by loss framing, we develop a methodology for density estimation closely related to standard difference-in-differences approaches. Conceptually, this approach may be thought of in two steps: (i) estimating the evolution of the distribution of final sales between the pre-period and the treatment window and (ii) estimating the difference in such evolutions between the treatment and control group. Under a generalization of the typical parallel-trend assumption—now requiring parallel evolution of full distributions rather than means—this provides an estimate of the causal impact of treatment on the distribution of final sales achieved.

Formally, consider a univariate, independent sample $\{x_n^{(g,t)}\}_{n \in N}$. In this notation, subscript n denotes the observation (out of a total set of N), the superscript g denotes two groups, and the superscript t denotes two time periods. This variable is distributed according to group-and-time-specific unknown densities $f^{(g,t)}$. Let the group-and-time-specific number of observations be denoted by $N^{(g,t)}$.

In the absence of group-specific intervention, densities are assumed to evolve in a manner that satisfies the assumption $f^{(g,2)} - f^{(g,1)} = f^\Delta$ for all groups g . f^Δ thus denotes the manner in which mass is shifted in the density function over time. However, if treatment is applied to one group (denoted T, with the other denoted C for control), the distribution that arises is represented as $f^{(T,2)} = f^{(T,1)} + f^\Delta + f^T$. The term f^T denotes an additional redistribution of mass induced by the treatment, and estimation of this term is the goal of this exercise.

Given these assumptions, a close analog to the common difference-in-differences estimator for means immediately arises:

$$(3) \quad f^T = [f^{(T,2)} - f^{(T,1)}] - [f^{(C,2)} - f^{(C,1)}].$$

Conceptually, we may examine the impact of treatment by examining how the distribution in the treatment group changed over time, differencing out the changes that may be attributable to the mere passage of time as inferred by the changes occurring in the control group.

Our formal estimate of f^T arises through the application of the analog principle (Goldberger 1968; Manski 1986), substituting finite-sample estimates of

these densities for the population densities themselves. Given a kernel function K and a bandwidth h , define the kernel density estimator to be $\hat{f}_h^{(g,t)}(x) = \frac{1}{hN^{(g,t)}} \sum_{n=1}^{N^{(g,t)}} K\left(\frac{x - x_n^{(g,t)}}{h}\right)$.²² Given these definitions, the DKDD estimator of f^T , evaluated at point x with bandwidth h , is given by

$$(4) \quad \hat{f}_h^T(x) = \left[\hat{f}_h^{(T,2)}(x) - \hat{f}_h^{(T,1)}(x) \right] - \left[\hat{f}_h^{(C,2)}(x) - \hat{f}_h^{(C,1)}(x) \right].$$

The consistency of equation (4) as an estimator for f^T follows immediately from the well-known consistency of the individual kernel estimators and Slutsky's theorem.

In the next section, we apply this method to infer f^T as induced by loss framing. Whenever presented, confidence intervals will be generated through a bootstrap procedure, recalculating $\hat{f}_h^{(g,t)}(x)$ from 10,000 simulated samples generated by sampling by DMA cluster (with replacement).

DKDD Estimates of the Distributional Impact of Loss Framing.—Figure 5 presents our estimates of the treatment effect of loss framing on the distribution of final sales achieved, measured relative to the clawback threshold (i.e., the sales necessary to reach the 110 percent sales target). In this figure, positive (negative) values indicate that treatment increases (decreases) the probability of ending the month with sales at the x -axis location.

The left panel presents the DKDD estimate for the small-bonus model group. Results are clearly in line with the earlier estimate of a negative impact of loss framing on average sales. Two patterns are particularly notable. First, the positive estimates in the region to the left of the clawback threshold indicate that there is an increase in the probability that dealers fail to achieve their 110 percent sales target (although, as indicated by the plotted confidence regions, at any individual point these positive estimates are statistically relatively weak). Second, the negative estimates just to the right of the clawback threshold indicate that dealers are less likely to achieve sales just past their 110 percent sales target. Other than this sharp decline in the probability of ending the month with sales just over the threshold, this figure does not display sharp evidence of excess or missing mass at particular points and instead might better be characterized as a general and nonlocalized shift from higher to lower values. This shift downward is consistent with the 4 percent decline in 110 percent sales target attainment reported in Section ID.

The right panel presents results for the large-bonus model group. Recall that we reported a null effect of loss framing on average sales for this model group. In this figure, we see that the previous null finding was masking statistically significant effects across the distribution of sales. Most notably, there is a large decline in the

²² While we have chosen to use kernel-based methods, the approach outlined above can be applied with a variety of alternative approaches to density estimation. Our DKDD estimator may be considered a special case in a broader class of “difference-in-density-differences” estimators. See Kuhn and Yu (2021) for an application of this approach to histogram bins rather than kernel density estimates.

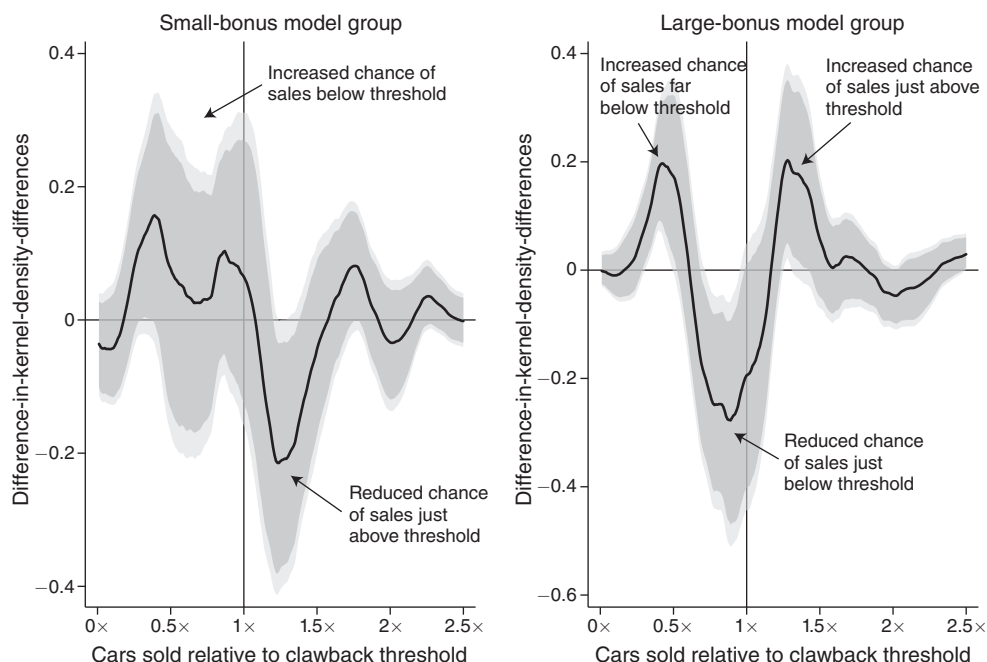


FIGURE 5. THE IMPACT OF LOSS FRAMING ON THE DISTRIBUTION OF SALES

Notes: This figure shows DKDD estimates of the impact of loss framing on the distribution of sales achieved. The x-axes show the amount of sales relative to the clawback threshold. The two subfigures present estimates derived from each model group. Confidence regions, based on 10,000 bootstrap iterations resampled by DMA, are shaded. The 90 percent confidence region is shaded darkly, and the 95 percent confidence region is shaded lightly. Kernel: Epanechnikov; bandwidth: .1x; sample sizes: 2,175 (left panel) and 2,234 (right panel).

probability that the treated dealers end the month in the region just under the clawback threshold relative to the control group. The redistribution of mass indicates that dealers who would otherwise have just triggered a clawback changed their behavior in one of two ways. Some dealers proceeded in a manner that resulted in more cars being sold, resulting in the observed excess mass of sales amounts just over the bonus threshold. This excess mass drives the 2 percent increase in target attainment reported in Section ID and is consistent with these dealers directing more attention, effort, or resources to these sales to avoid triggering the loss of the large bonus. However, some dealers proceeded in a manner that resulted in fewer cars being sold, resulting in the observed excess mass of sales amounts substantially below the bonus threshold. In this figure, these two groups are approximately equally sized, demonstrating why they average to a null effect when combined in the earlier difference-in-differences analysis.

We have suggested that some of our results might be explained by loss framing leading to a reassessment of multitasking behavior. In essence, the loss-framing intervention could lead a dealer to redeploy effort or resources from the sale of the small-bonus model group to the sale of the large-bonus model group. Such a story might be natural if dealers become more attentive to, or concerned about, reaching the necessary sales target for the large-bonus model group when loss framing is

active. Such a story could explain the overall negative results for the small-bonus model group through cannibalization of its sales.

To partially probe this possibility, we assess how estimated treatment effects change when these potential incentives for cannibalization of the small-bonus model group are in effect “shut down.” To do so, we isolate a minority of dealer-months in which the dealer fell far short of the large-bonus model group target.²³ We will assume that, in these cases, the dealers understood that diverting resources from small-bonus model group sales would not help push their large-bonus model group sales over the clawback threshold. If this is the case, then the effect of loss framing would not operate through motivating that diversion of resources. We operationalize the notion of “falling far short” of the large-bonus model group target by flagging dealers who sold less than or equal to 50 percent of their large-bonus model group clawback threshold—a mere 11 percent of dealer-months fall in this group.

Figure 6 presents the relevant analyses. As a baseline, the left panel presents our DKDD estimate of the impact of loss framing on small-bonus model group sales when dealers with minimal large-bonus model group sales are excluded. Compared to results from the full sample (previously presented in Figure 5), the pattern of effects is essentially unchanged. However, among the minority of responses with minimal large-bonus model group sales, a substantially different pattern emerges. The right panel presents our DKDD estimate of the impact of loss framing on small-bonus model group sales among dealers who demonstrably were not in contention for bonus attainment for the large-bonus model group. Here, we see that loss framing was associated with a higher probability of ending the month with small-bonus model group sales in the vicinity of the clawback threshold and a lower probability of ending the month with sales substantially below that threshold. This pattern of results is consistent with a shift from lower to higher sales amounts, with an associated increase in target attainment. In short, when incentives to redirect resources to large-bonus model group sales are “shut down,” the impact of loss framing on small-bonus model group sales appears straightforwardly positive and in line with standard intuitions.

G. Summary and Interpretation of Empirical Results

Taking stock, we have shown that the overall effect of introducing loss framing was a marginally significant *reduction* of average monthly sales. This negative effect was driven by a reduction of average sales for the small-bonus model group, whereas no effect was detected for the large-bonus model group. However, examining the manner in which distributions of sales changed, we observe that dealers facing loss framing were significantly less likely to end the month having narrowly missed avoiding clawback of the bonus for the large-bonus model group. Some of these dealers are estimated to have instead sold more vehicles from this model group, consistent with increasing or diverting effort and resources toward this model

²³ We include the small minority of dealers who sell only the small-bonus model group in this group.

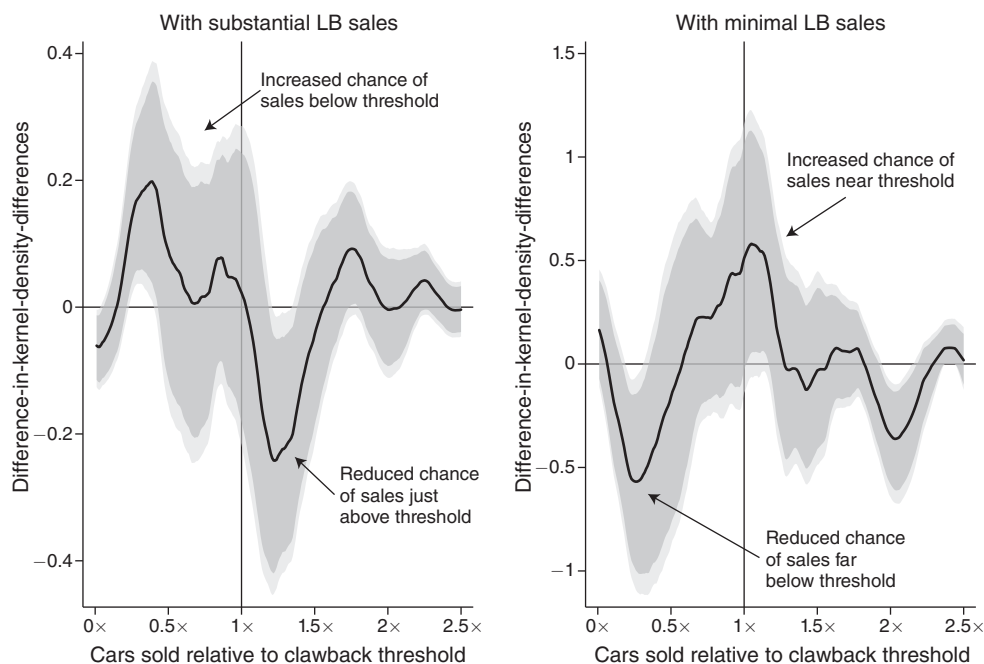


FIGURE 6. THE IMPACT OF LOSS FRAMING ON SMALL-BONUS MODEL GROUP SALES, CONDITIONAL ON LARGE-BONUS MODEL GROUP SALES

Notes: This figure shows DKDD estimates of the impact of loss framing on the distribution of small-bonus model group sales achieved. The x-axes show the amount of sales relative to the clawback threshold. The two subfigures present estimates conditioning on whether that month's sales of the large-bonus model group were comparatively high (over 50 percent of the clawback threshold) or low (under 50 percent of the clawback threshold). Confidence regions, based on 10,000 bootstrap iterations resampled by DMA, are shaded. The 90 percent confidence region is shaded darkly, and the 95 percent confidence region is shaded lightly. Kernel: Epanechnikov; bandwidth: .1x; sample sizes: 1,931 (left panel) and 244 (right panel).

group's sale. Other dealers are estimated to have sold fewer vehicles from this model group. Furthermore, when isolating dealers who appear to not be pursuing target attainment for the large-bonus model group, treatment effects of loss framing on the small-bonus model group are straightforwardly desirable.

These results suggest that one consequence of loss framing was a rebalancing of effort and resources across the two model groups. On average, the model group with the smaller potential loss was neglected, potentially due to effort or resources being directed to help avoid the larger potential loss—and with this pattern reversing when the large potential loss is unavoidable. In Section II, we present and analyze a model that illustrates this possibility.

H. Robustness Considerations and Additional Results

In the Supplemental Appendix, we assess a variety of robustness considerations and present additional supporting results. These include assessment of the correlates of opting in to the experiment (see Supplemental Appendix A.3), the demonstration

that our primary estimates are robust to the exclusion of the influential outliers documented in Figure 4 (see Supplemental Appendix A.4), the demonstration that our primary estimates are robust to the exclusion of the four DMAs whose treatments were reassigned (see Supplemental Appendix A.5), the estimation of treatment effects in the second treatment window based on a synthetic control approach utilizing experimental nonparticipants (see Supplemental Appendix A.6), a summary of results from post-experiment interviews and surveys (see Supplemental Appendix A.7), and an assessment of the timing of sales (see Supplemental Appendix A.8). Ultimately, none of these analyses contradict the summary of results presented above, but at times they offer additional context and illustration of underlying dealer behavior.

II. A Theoretical Examination of Loss Framing

In this section, we present a model that clarifies how inducing loss-averse evaluation of incentives through loss framing can help or hinder productivity. We use this model to demonstrate the ways in which past literature has incompletely characterized the predicted impact of loss-framing interventions and use these results to consider situations in which prevailing wisdom is expected to hold or is at risk to fail.

In our model, an agent sells items on behalf of a principal. The principal benefits from the agent's sales, and the agent faces private costs (or benefits) from sales activities. To mitigate the divergence in incentives, and in particular the incentive to underinvest in sales activities due to private costs, the principal has provided the agent with a bonus contract that yields direct payments for the amount of sales completed. There are potentially multiple dimensions of sales activity, each with its own incentive contract.

Using this model, we consider the wisdom of applying loss framing, which we assume operates by inducing loss aversion relative to a target level of sales. We characterize sufficient conditions for cases in which alternative sales approaches will be made desirable as compared to the choice made in the absence of loss aversion. As we will illustrate, the alternative strategies that are incentivized can reduce expected sales under relatively common conditions.

While we use the terminology "sales" to be explicit about the link to worker activities in the empirical section of our paper, note that this labeling is in no way essential. This theory can be understood to apply to broad notions of (observable) performance that a principal aims to motivate among his workers.

A. Seller's Decision Problem

The agent, or *seller*, has a window of time over which outcomes are evaluated (e.g., one month). During this window, the seller completes s^d sales and incurs c^d in private costs of sales activity for each dimension $d \in D$. Let \mathbf{s} and \mathbf{c} denote the vectors of these values across all dimensions.

When assessing each dimension of sales activity, the seller's ex post utility is determined by $u^d(\phi^d(s^d), c^d)$, where $\phi^d(s^d)$ denotes the payment received for completing s^d sales and c^d denotes the net private costs that the seller incurs in the sales process. The incentive scheme $\phi^d(s^d)$ is meant to promote the pursuit of

sales, so $\phi^d(s_1^d) \geq \phi^d(s_2^d)$ if $s_1^d > s_2^d$. Utility is assumed to be increasing in payments ($\phi^d(s^d)$) and decreasing in costs (c^d).

Over the window of time considered, the seller faces a long sequence of individual decisions on how to manage the sales process. The seller's decision is the choice of a *selling strategy*, which constitutes a complete contingent management plan over the window in question. Denote an individual selling strategy as σ and the (finite) set of possible selling strategies as Σ . The decision-relevant consequence of the choice of a selling strategy is the determination of the joint distribution of sales (\mathbf{s}) and private costs (\mathbf{c}). Denote this joint distribution by $f(\mathbf{s}, \mathbf{c}|\sigma)$ and marginal distributions by, for example, $f(\mathbf{s}|\sigma)$.

The seller's decision problem is therefore to choose $\sigma \in \Sigma$ to maximize expected utility, given by $E[U|\sigma] = \int U(\mathbf{s}, \mathbf{c}) df(\mathbf{s}, \mathbf{c}|\sigma)$, where $U(\mathbf{s}, \mathbf{c}) = \sum_{d \in D} u^d(\phi^d(s^d), c^d)$.

B. Introducing Loss Framing

The principal considers a modification to the incentive contract that introduces loss framing (e.g., through a clawback scheme). We assume that the consequence of introducing loss framing is to induce loss aversion relative to a dimension-specific target level of sales (R^d). If loss aversion is induced, the seller's utility function is modified to be $U^\Lambda(\mathbf{s}, \mathbf{c}) = \sum_{d \in D} u^d(\phi^d(s^d), c^d) - \Lambda[\phi^d(R^d) - \phi^d(s^d)] \cdot \mathbf{1}\{\phi(s^d) < \phi(R^d)\}$. $\mathbf{1}$ denotes the indicator function, taking a value of 1 if the statement in parentheses is true and 0 otherwise. In this utility specification, for cases when the target is exceeded—a gain—utility is identical to the prior formulation, but for cases when the target is not met—a loss—an additional cost proportional to the lost incentive is imposed. The magnitude of this additional cost is governed by the excess weight placed on losses, $\Lambda \in \mathbb{R}^+$. This specification arises from the movement of reference points from 0 to R^d under the assumption of piecewise-linear gain/loss utility.²⁴

In order to assess the impact of loss framing, we will consider how its introduction influences choices between two selling strategies. Before presenting our result, we begin with a simple example that helps illustrate the potential for negative consequences.

²⁴Formally, assume dimension-specific utility is governed by $f_1^d(\phi^d(s^d), c^d) + \eta f_2^d(\phi^d(s^d) - \phi^d(R^d))$, with $f_2^d(\phi^d(s^d) - \phi^d(R^d)) = (\phi^d(s^d) - \phi^d(R^d)) + (\lambda - 1) \cdot (\phi^d(s^d) - \phi^d(R^d)) \cdot \mathbf{1}\{\phi^d(s^d) < \phi^d(R^d)\}$. This model includes arbitrary "standard" utility represented by f_1^d , augmented with the additional consideration of f_2^d allowing for a piece-wise linear version of prospect theory. In this formulation, the coefficient of loss aversion is λ , and the decision weight placed on prospect theory is η . Translating this model into the version in text, define $u^d(\phi^d(s^d), c^d)$ to incorporate both standard utility as well as utility derived from gains/loss evaluation without excess weighting of losses: $f_1^d(\phi^d(s^d), c^d) + \eta[\phi^d(s^d) - \phi^d(R^d)]$. Defining $\Lambda = \eta \cdot (\lambda - 1)$, the equivalence of the models then follows. Note that in this formulation, we have assumed that gains and losses are evaluated with respect to the difference in payments relative to reference level. Some applications of prospect theory would instead assume that gains and losses are evaluated with respect to the utility difference relative to the reference level. In practice, this assumption will matter little for our results: conceptually similar results can be generated under the alternate assumption. An advantage of the current approach is that, by assuming that the magnitude of loss is influenced by lost payments as opposed to lost utility, we need not impose much structure on the direct utility function u . In contrast, if utility evaluations served as an input to gain/loss evaluation, substantially more structure would need to be imposed.

Example 1: Reducing Sales to Avoid a Loss.—In this example, consider a seller who sells two types of goods, referred to as good 1 and good 2. Both goods are subject to the same incentive contract, $\phi(s) = s$, and the same dimension-specific utility function $u(\phi(s^d), c^d) = \phi(s^d) - c^d$. Throughout this example, we will assume that there are no private costs of sales: $c = 0$ for all dimensions and in all strategies.

In this example, the seller has two potential selling strategies. Strategy 1 yields 10 sales of each good. Strategy 2 yields 9 sales for good 1 and 12 sales for good 2.

Absent loss framing, the seller's utility is higher pursuing strategy 2:

$$U(s_1, c_1) = 10 + 10 = 20 < 21 = 9 + 12 = U(s_2, c_2).$$

Now consider the consequence of the principal applying loss framing relative to a sales target of 10 for each good. Further assume that $\Lambda = 2$. This modifies the utility evaluation to

$$U^\Lambda(s_1, c_1) = 10 + 10 = 20 > 19 = 9 + 12 - \Lambda \cdot 1 = U^\Lambda(s_2, c_2).$$

In this example the expected sales-maximizing strategy is chosen in the absence of loss framing, but it is not chosen when loss framing is introduced. Choosing strategy 2 over strategy 1 effectively entails trading the opportunity to sell 1 unit of good 1 for the opportunity to sell 2 units of good 2. For the purposes of maximizing expected sales, this is a desirable trade. However, when framed as a loss, forgoing that 1 unit of good 1 is viewed as more costly. For a sufficiently loss averse seller, trading a one-unit loss for a two-unit gain is undesirable.

This simple example captures the core intuition that arises in the theory to come. Stated most simply, applying loss framing incentivizes the seller to avoid losses. If strategies are available that avoid losses at the cost of expected sales, then applying loss framing increases the incentives to pursue those strategies.

C. Characterizing the Impact of Loss Framing

In order to formalize the claims made in interpreting the previous example, we require a means of characterizing when strategies assist in avoiding losses. Define the *loss exposure of strategy σ* as

$$(5) \quad L(\sigma) = \int \sum_{d \in D} [\phi^d(R^d) - \phi^d(s^d)] \cdot \mathbf{1}\{\phi^d(s^d) < \phi^d(R^d)\} df(s, c | \sigma).$$

In words, the loss exposure of a strategy is the expected magnitude of dimension-specific losses that arises from the distribution of outcomes that the strategy induces, summed across dimensions.

Using this definition, we may completely characterize the types of strategies that are incentivized by loss framing. When characterizing the strategies that are incentivized, we will refer to loss aversion making strategy σ_1 more attractive relative to σ_2 if $E[U^\Lambda | \sigma_1] - E[U^\Lambda | \sigma_2] > E[U | \sigma_1] - E[U | \sigma_2]$. For ordinal evaluation, we say that strategy σ_1 is preferred to strategy σ_2 if $E[U | \sigma_1] > E[U | \sigma_2]$, with the relevant definition of individual utility (U or U^Λ) applied.

PROPOSITION 1 (The Impact of Inducing Loss Framing): *Consider two selling strategies, σ_1 and σ_2 . Inducing loss framing makes σ_1 more attractive relative to σ_2 if and only if $L(\sigma_1) < L(\sigma_2)$. Furthermore, for sufficiently large values of Λ , σ_1 is preferred to σ_2 if $L(\sigma_1) < L(\sigma_2)$ and only if $L(\sigma_1) \leq L(\sigma_2)$.*

PROOF:

Note that expected utility with loss framing may be expressed as $E[U^\Lambda|\sigma] = E[U|\sigma] - \Lambda \cdot L(\sigma)$. From this formulation, it follows immediately that the difference in expected utility resulting from σ_1 and σ_2 is $E[U^\Lambda|\sigma_1] - E[U^\Lambda|\sigma_2] = (E[U|\sigma_1] - E[U|\sigma_2]) - \Lambda[L(\sigma_1) - L(\sigma_2)]$. The first term of this expression is simply the utility difference experienced absent loss aversion, and the second term is the difference in exposure to losses scaled by the excess weight on losses. This second term is positive if and only if $L(\sigma_1) < L(\sigma_2)$, establishing the first claim of the proposition.

We next establish the claim that, for sufficiently large Λ , σ_1 is preferred to σ_2 if $L(\sigma_1) < L(\sigma_2)$. Consider a pair of strategies with unequal loss exposure, denoted σ_i and σ_j , and assign labels such that $L(\sigma_i) < L(\sigma_j)$. Note that $E[U^\Lambda|\sigma_i] > E[U^\Lambda|\sigma_j]$ if and only if $\Lambda > (E[U|\sigma_i] - E[U|\sigma_j]) / [L(\sigma_i) - L(\sigma_j)] \equiv T_{i,j}$. Thus, if we define “sufficiently large” to mean $\Lambda > \max_{i,j} T_{i,j}$, for arbitrarily drawn pairs we are assured that $\Lambda > (E[U|\sigma_i] - E[U|\sigma_j]) / [L(\sigma_i) - L(\sigma_j)]$ and thus that σ_i is preferred to σ_j . The claim then follows.

The final claim, that for sufficiently large Λ , σ_1 is preferred to σ_2 only if $L(\sigma_1) \leq L(\sigma_2)$, follows immediately from the claim just proved. ■

Proposition 1 fully characterizes the impact of inducing loss aversion in performance incentives. When loss aversion is induced, incentives to reduce exposure to losses are increased. If a given strategy has a lower degree of loss exposure, a sufficiently loss-averse seller can be motivated to choose it if the principal imposes loss framing. In situations where the set of strategies involve trade-offs between loss exposure and expected sales, this drives the potential for negative consequences like those illustrated in the leading example above.

D. Reassessing Unidimensional Results

The results of the prior section establish that, in a relatively minimalist and general model, the performance effects of loss framing can be negative. This finding may be viewed as surprising in light of existing literature. It is in opposition to commonly applied intuitions regarding loss framing, and it would seem to contradict some formal results. This begs the question: What feature of our modeling approach drives these differences in conclusions? Since prior works have focused on unidimensional environments, and given our focus on the exacerbation of incentives for inefficient multitasking, one might be tempted to infer that the inclusion of additional dimensions is purely responsible for the differences. In this section, we will document, however, that conceptually similar results arise in unidimensional environments.

Example 2: Reduced Sales in a Unidimensional Stochastic Setting.—In this example, a seller is considering his approach to selling a single type of good. As in example 1, assume that the incentive contract is $\phi(s) = s$ and the utility function is $u(\phi(s), c) = \phi(s) - c$ (with superscript d suppressed in this unidimensional case). Again, assume that $\Lambda = 2$.

The seller has two potential strategies. Strategy 1 yields 10 sales with certainty. Strategy 2 results in an amount of sales drawn from the integers 7–14, each with an equal probability of occurring. This yields an expectation of 10.5 sales. Both strategies require the seller to incur 5 units of effort costs.

Absent loss framing, the seller's utility is higher pursuing strategy 2:

$$E[u(\phi(s_1), c_1)] = 10 - 5 = 5 < 5.5 = 10.5 - 5 = E[u(\phi(s_2), c_2)].$$

If the principal imposes loss framing relative to a sales target of 10, this modifies the utility evaluation to

$$\begin{aligned} E[u^\Lambda(\phi(s_1), c_1)] &= 10 - 5 - \Lambda \cdot L(\sigma_1) \\ &= 5 > 4 = 10.5 - 5 - \Lambda \cdot L(\sigma_2) = E[u^\Lambda(\phi(s_2), c_2)]. \end{aligned}$$

Note that strategy 1 has no loss exposure, so $L(\sigma_1) = 0$. Strategy 2, however, has the potential for losses of size 3, 2, or 1 if the realization of sales is 7, 8, or 9, respectively. Weighting each of these losses by its probability of occurring ($1/8$) yields $L(\sigma_2) = 6/8$, generating the inequality above. The worker would therefore choose strategy 1 with loss aversion induced.

In this example, loss framing again leads the seller to pursue a strategy that does not maximize expected sales. As with example 1, the strategies available involve a trade-off between loss exposure and expected sales. In example 1, avoidance of loss exposure was pursued through “gaming” of sales pursuit across dimensions. In this example, avoidance of loss exposure was pursued through gaming of tolerated risk. This illustrates that similar concerns described above can extend into the unidimensional environments considered in the literature as long as risk/reward trade-offs are present.

E. A Modified Statement of the Conventional Wisdom

Example 2 demonstrates that the conceptual issues considered in this paper extend even to unidimensional environments that have been the focus of this literature. The element that is essential for a potential of negative consequences of loss framing is not multidimensionality but rather the presence of strategies that trade off greater loss exposure for greater expected productivity. In the unidimensional environment, this leads us to a proposition that we believe provides an accurate but qualified refinement of the prevailing wisdom regarding loss framing.

PROPOSITION 2 (Motivation of Dominant Sales Strategies): *Consider two (unidimensional) strategies, σ_1 and σ_2 . Assume that $f(s|\sigma_1)$ first-order stochastically dominates $f(s|\sigma_2)$. Inducing loss framing cannot make σ_2 more attractive relative to*

σ_1 , and for some potential reference values, inducing loss framing makes σ_1 more attractive relative to σ_2 . For those reference values, σ_1 is preferred to σ_2 if the additional weight placed on losses Λ is sufficiently large.

PROOF:

The assumption that $f(s|\sigma_1)$ first-order stochastically dominates $f(s|\sigma_2)$ implies that $L(\sigma_1) \leq L(\sigma_2)$, with the inequality strict for some potential reference values. The claims therefore follow immediately from Proposition 1. ■

Proposition 2 implies that loss framing is predicted to have an unambiguous (weakly) positive effect on productivity when the available strategies are ordered by first-order stochastic dominance. This aligns exactly with the intuitions just described: when strategies are so ordered, there exists no trade-off between loss exposure and expected sales.

Proposition 2 additionally helps to reconcile our results with existing theoretical claims. For example, prediction 1 of Imas, Sadoff, and Samek (2016) and prediction 1 of de Quidt (2017) both establish unambiguously positive effects of loss framing. The models in these papers involve effort choices where greater effort would indeed induce a first-order stochastically dominant outcome distribution. And indeed, we believe that this is a reasonable modeling assumption in these papers' experimental contexts, which present straightforward "real effort tasks" that are intentionally designed to be clean measures of effort provision. To the extent that most of the lab-experimental literature uses tasks like these, the conventional wisdom may perhaps be generally expected to hold in these papers. A notable exception occurs in the work of Ahrens, Bitter, and Bosch-Rosa (2023), who document negative effects of loss framing among players of a coordination game facing strategic uncertainty. This experimental environment involves the type of loss exposure/productivity trade-offs that can generate negative effects in our unidimensional case, and indeed the authors present their findings as a validation of our theory.

Beyond demonstrating the ways in which this past literature is typically correct, however, this analysis also demonstrates ways in which past literature has been limited. By focusing on environments where the only way to avoid losses is to increase effort, the lab-experimental literature has directed attention to cases where the channels for negative consequences do not exist. In the field settings where loss-framed interventions have been attempted—for example, among factory workers, salespeople, or teachers—productivity is clearly a substantially more complex object with more scope for meaningful trade-offs. This creates the potential for negative productivity consequences like those we have documented to rise to first-order importance. This perhaps helps explain why these field studies have not systematically seen the large positive effects that have been observed, on average, in the lab literature (as documented in the meta-analysis of Ferraro and Tracy 2022).

III. Discussion

In recent years, the findings of behavioral economics have been broadly disseminated to the general public. This surge of public engagement has contributed to a

wave of interest in designing policy interventions that harness behavioral economic forces. While the potential for using loss framing to sharpen performance incentives has received a great deal of attention, formal field evaluation of such programs is currently limited to a relatively small number of cases (see Ferraro and Tracy 2022 for a critical review). Furthermore, formal theoretical examination of the effect of these incentives has been limited to their application to somewhat narrow and stylized circumstances.

In this paper, we have sought to critically assess the desirability of loss framing for organizations and policy makers. In an unusually large field experiment, in which \$66 million in bonus payments was subject to random assignment of loss framing, we have found strong indications that loss aversion influenced sales activity in a manner contrary to our typical expectations. When distilling our findings, some caution is merited due to the marginal significance of some of our main results. However, the overall patterns in our data, and especially those revealed in our DKDD analysis, provide clear reasons to believe that loss framing was ultimately harmful to sales efforts. Furthermore, despite common claims to the contrary, we have shown that such negative results are theoretically expected.

Our findings clearly illustrate a position summarized in Card, DellaVigna, and Malmendier (2011): that field experiments can often benefit from the development of guiding theory, even *ex post*.²⁵ The initial experimental findings of Hossain and List (2012) motivated substantial interest in loss-framing interventions. However, in papers that followed, some researchers faced difficulties when attempting to generate similar positive treatment effects.²⁶ Through a formalized reexamination of the underlying conceptual premise of this literature, we are able to suggest factors that could contribute to the mixed efficacy of these interventions and are able to make more precise predictions about the situations in which loss framing may be counterproductive.

We note that although automotive retailing in the United States is an idiosyncratic (yet economically important) setting, the characteristics that generate our results are in no way unique to that setting. One survey reported that 72 percent of firms include bonus pay in their sales compensation systems. Of those firms, 76 percent use performance quotas (Joseph and Kalwani 1998). Discrete bonus systems are used across numerous private and public sectors (e.g., Tzioumis and Gee 2013; Fryer 2013; Asch 1990), most of which face the potential for gaming and multi-tasking problems. Our theoretical model explains why loss framing poses the same hazards in these settings as it does in auto dealer incentives.

Although the presence of loss aversion is undesirable in our setting, we note that this does not undercut the demonstration that loss aversion is *present*. Until relatively recently, demonstrations of loss aversion were largely restricted to lab settings, allowing reasonable researchers to question the field relevance of the phenomenon. More recently, perhaps due to the increase in data availability, we have seen

²⁵ This point is particularly emphasized in fields where generalizability is a common criticism of experimentation (Di Stefano and Gutierrez 2019).

²⁶ And indeed, the original work of Hossain and List (2012) did not find unambiguous positive effects in all situations studied.

a proliferation of demonstrations of loss aversion in fundamentally economic field environments. Loss aversion has now been shown to influence job search behavior (DellaVigna et al. 2017), labor supply (see Camerer et al. 1997 and the many papers it inspired), house prices (Genesove and Mayer 2001; Andersen et al. 2022), tax compliance (Engström et al. 2015; Rees-Jones 2018), and more (for a thorough review, see O'Donoghue and Sprenger 2018). We add to this growing list of field applications and provide a demonstration with an unusually large estimated financial impact.

Because the dealers that we study are self-contained businesses themselves, an important novelty of our results is our demonstration that behavioral models apply not only to individual workers but also to firms. While it is often argued that the experience, stakes, and competitive forces present among market actors should eliminate the role of behavioral economic considerations (List 2002, 2003, 2004a,b), our results suggest that this is not so in one large and important market. This applicability may be partially explained by the largely private ownership structure of car dealerships that is particularly dominant in our focal manufacturer's dealer network (98 percent). Because many dealerships have small management teams, they may operate more like individuals than the abstract ideal of a firm. Furthermore, franchising laws heavily limit entry and consolidation in this market, limiting some of the competitive pressure that is often assumed to discipline behavioral tendencies. Despite these caveats, we note that the ownership characteristics and competitive frictions in our setting are common in much of the world (Bloom and Van Reenen 2007; Bloom, Sadun, and Van Reenen 2012) and that some frictions to idealized competition exist in many other markets. We therefore view our paper as strongly suggesting that behavioral economic considerations apply to the analysis of firms in quantitatively important ways²⁷—and thus that policy interventions guided by behavioral economics might be productively deployed to such populations.

REFERENCES

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller.** 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493–505.
- Ahrens, Steffen, Lea Bitter, and Ciril Bosch-Rosa.** 2023. "Coordination under Loss Contracts." *Games and Economic Behavior* 137: 270–93.
- Andersen, Steffen, Cristian Badarinza, Lu Liu, Julie Marx, and Tarun Ramadorai.** 2022. "Reference Dependence in the Housing Market." *American Economic Review* 112 (10): 3398–440.
- Asch, Beth J.** 1990. "Do Incentives Matter? The Case of Navy Recruiters." *ILR Review* 43 (3): 89–106.
- Athey, Susan, and Guido W. Imbens.** 2017. "The Econometrics of Randomized Experiments." In *Handbook of Economic Field Experiments*, Vol. 1, edited by Abhijit Vinayak Banerjee and Esther Duflo, 73–140. Amsterdam: Elsevier.
- Baker, George P., and Thomas N. Hubbard.** 2003. "Make versus Buy in Trucking: Asset Ownership, Job Design, and Information." *American Economic Review* 93 (3): 551–72.
- Baker, Malcolm, Xin Pan, and Jeffrey Wurgler.** 2012. "The Effect of Reference Point Prices on Mergers and Acquisitions." *Journal of Financial Economics* 106 (1): 49–71.
- Balmaceda, Felipe.** 2018. "Optimal Task Assignments with Loss-Averse Agents." *European Economic Review* 105: 1–26.

²⁷ For additional papers emphasizing the role of loss aversion in corporate settings, see Loughran and McDonald (2013); Ljungqvist and Wilhelm Jr (2005); Baker, Pan, and Wurgler (2012); Dittmann, Maug, and Spalt (2010).

- Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker. 2017. "An Assessment of the National Establishment Time Series (NETS) Database." Federal Reserve Bank of Minneapolis System Working Paper 17–29.
- Benartzi, Shlomo, John Beshears, Katherine L. Milkman, Cass R. Sunstein, Richard H. Thaler, Maya Shankar, Will Tucker-Ray, William J. Congdon, and Steven Galing. 2017. "Should Governments Invest More in Nudging?" *Psychological Science* 28 (8): 1041–55.
- Bennett, Victor Manuel. 2013. "Organization and Bargaining: Sales Process Shoice at Auto Dealerships." *Management Science* 59 (9): 2003–18.
- Benson, Alan. 2015. "Do Agents Game Their Agents' Behavior? Evidence from Sales Managers." *Journal of Labor Economics* 33 (4): 863–90.
- Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." *Quarterly Journal of Economics* 122 (4): 1351–408.
- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012. "The Organization of Firms across Countries." *Quarterly Journal of Economics* 127 (4): 1663–705.
- Brooks, Richard R. W., Alexander Stremitz, and Stephan Tontrup. 2012. "Framing Contracts: Why Loss Framing Increases Effort." *Journal of Institutional and Theoretical Economics* 168 (1): 62–82.
- Brownback, Andy, and Sally Sadoff. 2020. "Improving College Instruction through Incentives." *Journal of Political Economy* 128 (8): 2925–72.
- Bulte, Erwin, John A. List, and Daan van Soest. 2020. "Toward an Understanding of the Welfare Effects of Nudges: Evidence from a Field Experiment in the Workplace." *Economic Journal* 130 (632): 2329–53.
- Bulte, Erwin, John A. List, and Daan van Soest. 2021. "Incentive Spillovers in the Workplace: Evidence from Two Field Experiments." *Journal of Economic Behavior & Organization* 184: 137–49.
- Busse, Meghan R., and Jorge M. Silva-Risso. 2010. "'One Discriminatory Rent' or 'Double Jeopardy': Multicomponent Negotiation for New Car Purchases." *American Economic Review* 100 (2): 470–74.
- Busse, Meghan, Jorge Silva-Risso, and Florian Zettelmeyer. 2006. "\$1,000 Cash Back: The Pass-Through of Auto Manufacturer Promotions." *American Economic Review* 96 (4): 1253–70.
- Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *Quarterly Journal of Economics* 112 (2): 407–41.
- Card, David, Stefano DellaVigna, and Ulrike Malmendier. 2011. "The Role of Theory in Field Experiments." *Journal of Economic Perspectives* 25 (3): 39–62.
- Chen, Yvonne Jie, Pei Li, and Yi Lu. 2018. "Career Concerns and Multitasking Local Bureaucrats: Evidence of a Target-Based Performance Evaluation System in China." *Journal of Development Economics* 133: 84–101.
- Chung, Doug J., and Das Narayandas. 2017. "Incentives versus Reciprocity: Insights from a Field Experiment." *Journal of Marketing Research* 54 (4): 511–24.
- Chung, Doug J., Das Narayandas, and Dongkyu Chang. 2020. "The Effects of Quota Frequency: Sales Performance and Product Focus." *Management Science* 67 (4): 2151–70.
- Chung, Doug J., Thomas Steenburgh, and K. Sudhir. 2014. "Do Bonuses Enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-Based Compensation Plans." *Marketing Science* 33 (2): 165–87.
- Courty, Pascal, and Gerald Marschke. 2004. "An Empirical Investigation of Gaming Responses to Explicit Performance Incentives." *Journal of Labor Economics* 22 (1): 23–56.
- DellaVigna, Stefano, Attila Lindner, Balázs Reizer, and Johannes F. Schmieder. 2017. "Reference-Dependent Job Search: Evidence from Hungary." *Quarterly Journal of Economics* 132 (4): 1969–2018.
- DellaVigna, Stefano, and Devin Pope. 2018. "What Motivates Effort? Evidence and Expert Forecasts." *Review of Economic Studies* 85 (2): 1029–69.
- Di Stefano, Giada, and Cédric Gutierrez. 2019. "Under a Magnifying Glass: On the Use of Experiments in Strategy Research." *Strategic Organization* 17 (4): 497–507.
- Dittmann, Ingolf, Ernst Maug, and Oliver Spalt. 2010. "Sticks or Carrots? Optimal CEO Compensation When Managers Are Loss Averse." *Journal of Finance* 65 (6): 2015–50.
- Dumont, Etienne, Bernard Fortin, Nicolas Jacquemet, and Bruce Shearer. 2008. "Physicians' Multitasking and Incentives: Empirical Evidence from a Natural Experiment." *Journal of Health Economics* 27 (6): 1436–50.
- Ederer, Florian, Richard Holden, and Margaret Meyer. 2018. "Gaming and Strategic Opacity in Incentive Provision." *RAND Journal of Economics* 49 (4): 819–54.
- Engström, Per, Katarina Nordblom, Henry Ohlsson, and Annika Persson. 2015. "Tax Compliance and Loss Aversion." *American Economic Journal: Economic Policy* 7 (4): 132–64.

- Ferraro, Paul J., and J. Dustin Tracy.** 2022. "A Reassessment of the Potential for Loss-Framed Incentive Contracts to Increase Productivity: A Meta-Analysis and a Real-Effort Experiment." *Experimental Economics* 25 (5): 1441–66.
- Fryer, Roland G., Jr.** 2013. "Teacher Incentives and Student Achievement: Evidence from New York City Public Schools." *Journal of Labor Economics* 31 (2): 373–407.
- Fryer, Roland G., Jr., Steven D. Levitt, John List, and Sally Sadoff.** 2022. "Enhancing the Efficacy of Teacher Incentives through Framing: A Field Experiment." *American Economic Journal: Economic Policy* 14 (4): 269–99.
- Genesove, David, and Christopher Mayer.** 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market." *Quarterly Journal of Economics* 116 (4): 1233–60.
- Gill, David, and Rebecca Stone.** 2010. "Fairness and Desert in Tournaments." *Games and Economic Behavior* 69 (2): 346–64.
- Goldberger, Arthur S.** 1968. *Topics in Regression Analysis*. New York: Macmillan.
- Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev.** 2017. "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *Journal of Public Economics* 148: 14–31.
- Herweg, Fabian, Daniel Müller, and Philipp Weinschenk.** 2010. "Binary Payment Schemes: Moral Hazard and Loss Aversion." *American Economic Review* 100 (5): 2451–77.
- Holmstrom, Bengt, and Paul Milgrom.** 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, & Organization* 7 (S): 24–52.
- Homonoff, Tatiana A.** 2018. "Can Small Incentives Have Large Effects? The Impact of Taxes versus Bonuses on Disposable Bag Use." *American Economic Journal: Economic Policy* 10 (4): 177–210.
- Hong, Fuhai, Tanjim Hossain, and John A. List.** 2015. "Framing Manipulations in Contests: A Natural Field Experiment." *Journal of Economic Behavior & Organization* 118: 372–82.
- Hossain, Tanjim, and John A. List.** 2012. "The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations." *Management Science* 58 (12): 2151–67.
- Imas, Alex, Sally Sadoff, and Anya Samek.** 2016. "Do People Anticipate Loss Aversion?" *Management Science* 63 (5): 1271–84.
- Jansen, Mark, Lamar Pierce, Jason Snyder, and Hieu Nguyen.** 2023. "Product Sales Incentive Spillovers to the Lending Market: Evidence from Subprime Auto Loan Defaults." *Management Science*. <https://doi.org/10.1287/mnsc.2023.4935>.
- Joseph, Kissan, and Manohar U. Kalwani.** 1998. "The Role of Bonus Pay in Salesforce Compensation Plans." *Industrial Marketing Management* 27 (2): 147–59.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler.** 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives* 5 (1): 193–206.
- Kim, Minkyung, K. Sudhir, and Kosuke Uetake.** 2021. "A Structural Model of a Multitasking Salesforce: Incentives, Private Information and Job Design." *Management Science* 68 (6): 4602–30.
- Kleven, Henrik Jacobsen.** 2016. "Bunching." *Annual Review of Economics* 8: 435–64.
- Kőszegi, Botond, and Matthew Rabin.** 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121 (4): 1133–65.
- Kuhn, Peter J., and Lizi Yu.** 2024. "Kinks as Goals: Accelerating Commissions and the Performance of Sales Teams." *Management Science*. <https://doi.org/10.1287/mnsc.2023.01661>.
- Lafontaine, Francine, and Fiona Scott Morton.** 2010. "Markets: State Franchise Laws, Dealer Terminations, and the Auto Crisis." *Journal of Economic Perspectives* 24 (3): 233–50.
- Larkin, Ian.** 2014. "The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales." *Journal of Labor Economics* 32 (2): 199–227.
- List, John A.** 2002. "Testing Neoclassical Competitive Market Theory in the Field." *Proceedings of the National Academy of Sciences* 99 (24): 15827–30.
- List, John A.** 2003. "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics* 118 (1): 41–71.
- List, John A.** 2004a. "Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace." *Econometrica* 72 (2): 615–25.
- List, John A.** 2004b. "Testing Neoclassical Competitive Theory in Multilateral Decentralized Markets." *Journal of Political Economy* 112 (5): 1131–56.
- Ljungqvist, Alexander, and William J. Wilhelm Jr.** 2005. "Does Prospect Theory Explain IPO Market Behavior?" *Journal of Finance* 60 (4): 1759–90.
- Loughran, Tim, and Bill McDonald.** 2013. "IPO First-Day Returns, Offer Price Revisions, Volatility, and Form S-1 Language." *Journal of Financial Economics* 109 (2): 307–26.

- Lu, Bo, Robert Greevy, Xinyi Xu, and Cole Beck.** 2011. "Optimal Nonbipartite Matching and Its Statistical Applications." *American Statistician* 65 (1): 21–30.
- Manski, Charles F.** 1986. "Analog Estimation of Econometric Models." In *Handbook of Econometrics*, Vol. 4, edited by Robert F. Engle and Daniel L. McFadden, 2559–82. Amsterdam: Elsevier.
- Misra, Sanjog, and Harikesh S. Nair.** 2011. "A Structural Model of Sales-Force Compensation Dynamics: Estimation and Field Implementation." *Quantitative Marketing and Economics* 9 (3): 211–57.
- Murry, Charles, and Henry S. Schneider.** 2016. "The Economics of Retail Markets for New and Used Cars." In *Handbook on the Economics of Retailing and Distribution*, edited by Emek Basker, 343–367. Cheltenham, UK: Edward Elgar Publishing.
- O'Donoghue, Ted, and Charles Sprenger.** 2018. "Reference-Dependent Preferences." In *Handbook of Behavioral Economics: Foundations and Applications*, Vol. 1, edited by B. Douglas Bernheim, Stefano DellaVigna and David Laibson, 1–77. Amsterdam: North-Holland.
- Oyer, Paul.** 1998. "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality." *Quarterly Journal of Economics* 113 (1): 149–85.
- Pierce, Lamar.** 2012. "Organizational Structure and the Limits of Knowledge Sharing: Incentive Conflict and Agency in Car Leasing." *Management Science* 58 (6): 1106–21.
- Pierce, Lamar and Alex Rees-Jones.** 2020. *The Negative Consequences of Loss-Framed Performance Incentives*. AEA RCT Registry. <https://doi.org/10.1257/rct.5362-1.0>.
- Pierce, Lamar, Alex Rees-Jones, and Charlotte Blank.** 2025. *Data and Code for: "The Negative Consequences of Loss-Framed Performance Incentives."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI: <https://doi.org/10.3886/E198489V1>.
- Rees-Jones, Alex.** 2018. "Quantifying Loss-Averse Tax Manipulation." *Review of Economic Studies* 85 (2): 1251–78.
- Ritz, Robert A.** 2008. "Strategic Incentives for Market Share." *International Journal of Industrial Organization* 26 (2): 586–97.
- Robbins, Michael W., Jessica Saunders, and Beau Kilmer.** 2017. "A Framework for Synthetic Control Methods with High-Dimensional, Micro-Level Data: Evaluating a Neighborhood-Specific Crime Intervention." *Journal of the American Statistical Association* 112 (517): 109–26.
- Roberts, Adrienne.** 2018. "Car Dealerships Face Conundrum: Get Big or Get Out." *Wall Street Journal*, April 8, 2018. <https://www.wsj.com/articles/car-dealerships-face-conundrum-get-big-or-get-out-1523192401>.
- Steenburgh, Thomas J.** 2008. "Effort or Timing: The Effect of Lump-Sum Bonuses." *Quantitative Marketing and Economics* 6 (3): 235–56.
- Tzioumis, Konstantinos, and Matthew Gee.** 2013. "Nonlinear Incentives and Mortgage Officers' Decisions." *Journal of Financial Economics* 107 (2): 436–53.