NBER WORKING PAPER SERIES

THE NEGATIVE CONSEQUENCES OF LOSS-FRAMED PERFORMANCE INCENTIVES

Lamar Pierce
Alex Rees-Jones
Charlotte Blank

Working Paper 26619
http://www.nber.org/papers/w26619

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020

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 initial 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 during the conduct of this research. For helpful comments, we thank Greg Besharov, Jon de Quidt, Alex Imas, Ian Larkin, Ted O'Donoghue, and seminar participants at Cornell, the University of Maryland, the Society for Judgment and Decision Making Annual Meeting, and the Institutions and Innovation Conference. The opinions expressed in this paper are solely the authors', and do not necessarily reflect the views of any individual or institution listed above, nor of the National Bureau of Economic Research.

The Negative Consequences of Loss-Framed Performance Incentives
Lamar Pierce, Alex Rees-Jones, and Charlotte Blank
NBER Working Paper No. 26619
January 2020
JEL No. D03,D81,J22,J31

ABSTRACT

Behavioral economists have proposed that loss-averse employees increase productivity when bonuses are "loss framed"—prepaid then clawed back if targets are unmet. We theoretically document that loss framing raises incentives for costly risk mitigation and for inefficient multitasking, potentially leading to large negative performance effects. We empirically document evidence of these concerns in a nationwide field experiment among 294 car dealers. Dealers randomized into loss-framed (but financially identical) contracts sold 5% fewer vehicles than control dealers, generating a revenue loss of $45 million over 4 months. We discuss implications regarding the use of behavioral economics to motivate both employees and firms.

Lamar Pierce
Olin Business School
Washington University in St Louis
One Brookings Drive Box 1133
Saint Louis, MO 63130
pierce@wustl.edu

Alex Rees-Jones
Cornell University
Economics Department
404 Uris Hall
Ithaca, NY 14850
and NBER
arr34@cornell.edu

Charlotte Blank
Maritz
1375 North Highway Drive
Fenton, MO 63026
charlotte.blank@maritz.com

The Negative Consequences of Loss-Framed Performance Incentives

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

Behavioral economists have proposed that loss-averse employees increase productivity when bonuses are "loss framed"—prepaid then clawed back if targets are unmet. We theoretically document that loss framing raises incentives for costly risk mitigation and for inefficient multitasking, potentially leading to large negative performance effects. We empirically document evidence of these concerns in a nationwide field experiment among 294 car dealers. Dealers randomized into loss-framed (but financially identical) contracts sold 5% fewer vehicles than control dealers, generating a revenue loss of $45 million over 4 months. We discuss implications regarding the use of behavioral economics to motivate both employees and firms.
JEL: D03, D81, J22, J31
Keywords: Loss Aversion, Field Experiments, Worker Incentives, Franchise Contracts

I. Introduction

Of the large number of heuristics and biases documented in the behavioral economics literature, few have received as much attention as loss aversion. A dense experimental literature has rigorously examined many of the predictions of loss aversion, and in general has provided supportive results, leading Barberis (2013) to conclude that prospect theory (with its key element of loss aversion) is widely viewed as the best available description of risk evaluation in experimental settings.

In recent years, these successes have motivated interest in exploring the ways in which loss aversion might be harnessed by organizations. Working in partnership with behavioral economists, businesses have sought to use the observation that workers are influenced by loss aversion[1] as the basis for improvements to incentive schemes. In a now-classic initial demonstration, Hossain and List (2012) conducted a natural field experiment among Chinese factory workers and found that loss-framed bonus payments led to higher productivity.

* Pierce: Olin Business School, Washington University in St. Louis, pierce@wustl.edu. Rees-Jones: Cornell University, arr34@cornell.edu. Blank: Maritz, LLC, charlotte.blank@maritz.com. 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 initial 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 during the conduct of this research. For helpful comments, we thank Greg Besharov, Jon de Quidt, Alex Imas, Ian Larkin, Ted O'Donoghue, and seminar participants at Cornell, the University of Maryland, the Society for Judgment and Decision Making Annual Meeting, and the Institutions and Innovation Conference.

[1]See, e.g., Camerer et al. (1997); Fehr and Goette (2007); Farber (2008); Crawford and Meng (2011); Thakral and Tô (2019).

Results like these have contributed to the development of a lay theory that loss-framed bonus payments are desirable for motivating performance (see, e.g., Fryer Jr et al., 2012; Brooks, Stremitzer and Tontrup, 2012; Hong, Hossain and List, 2015; DellaVigna and Pope, 2018; Chung and Narayandas, 2017). The underlying logic of this lay theory is that loss aversion increases the perceived desirability of bonuses and thus motivates more effort to earn them. Several papers have made the important counterpoint that anticipation of such incentives could lead workers to avoid accepting loss-framed contracts (Luft, 1994; Imas, Sadoff and Samek, 2016; de Quidt, 2017). Ultimately, however, the experimental tests presented in these papers contain little evidence supporting this worry, with some evidence instead suggesting that loss framing is viewed as desirable by workers. Based on this literature, a prudent firm could conclude that there is little downside to the incorporation of loss framing into their bonus system.[2]

In this paper, we demonstrate that inducing loss aversion has substantial potential for negative performance effects in manners unexplored in this prior literature. In the theoretical component of this paper, we formalize the existing lay theory of the value of loss-framed bonuses. We elucidate the cases in which the existing intuitions are correct and the cases in which they fail. In the empirical component of this paper, we report the results of a large-scale field experiment in which loss-framed bonuses were deployed by a firm seeking to motivate sales. We find that this framing substantially harmed performance through subtle but important means.

To study this issue theoretically, we model a principal incentivizing a worker's acceptance of private costs of production through bonuses based on ex-post performance. In this environment, we study how the desirability of different strategies changes when loss aversion is induced. We first consider the case where the principal cares about only a single dimension of production. In such an environment, if production is deterministic, the common wisdom holds: inducing loss aversion weakly increases the perceived value of the worker's productive activities relative to private costs, and thus weakly increases effort. In stochastic environments, however, inducing loss aversion also substantially modifies the worker's risk tolerance, potentially leading to perverse effects. A sufficiently loss-averse worker would accept an arbitrarily large reduction in expected production for an arbitrarily small reduction in properly defined exposure to losses. This result arises because increasing loss aversion accentuates the kink in utility occurring at the reference point, which generates first-order risk aversion. Motivating loss avoidance in such a manner is rarely assumed to be in the interests of the principal. We next consider a generalization of this model in which the principal cares about multiple dimensions of production, with these dimensions potentially representing different products or different time windows. In such an environment, bonus systems can incentivize undesirable allocation of effort across dimensions, even in the absence of "behavioral" concerns (see, e.g., Holmstrom and Milgrom, 1991). We show that inducing loss aversion exacerbates

[2]While we have emphasized the apparent lack of concern about downsides of loss framing, we note that not all existing empirical tests demonstrate an upside. For example, the experimental work of de Quidt et al. (2017) finds precisely estimated null effects of loss framing, the positive effects in Hossain and List (2012) are generally small and imprecisely estimated, and the positive effects seen in DellaVigna and Pope (2018) are statistically insignificant. Furthermore, the work of Brooks, Stremitzer and Tontrup (2017) cautions that a poorly chosen target can set undesirable norms.

these incentives, again to the detriment of the principal.

Our development of this theoretical model was guided by the results of a field experiment, the analysis of which we document in the remainder of the paper. Motivated at least in part by the literature highlighted above, a major car manufacturer chose to modify their sales bonus system in an attempt to harness loss aversion. This manufacturer's existing system contained bonuses for two groups of car models. For each model group, the manufacturer offered large bonuses for exceeding a monthly sales target. By prepaying both bonuses, and clawing each back if its target was not achieved, the manufacturer hoped to incentivize additional effort and sales. The researchers recommended an experimental design to estimate the treatment effect of the policy change prior to systemwide adoption.[3] During the experimental intervention, \$66 million in bonus payments was randomized to either pre- or post-payment.

Analyzing results from the 294 dealers who chose to participate in the experiment, we find that the prepayment scheme resulted in substantial losses in revenue. Across the 140 car dealerships who faced the prepayment scheme in the initial four-month treatment window, we estimate that the prepayment scheme caused a 5% reduction in unit sales, ultimately resulting in over \$45 million in lost revenue.[4] Applying our estimates to the entire dealer network, we forecast that the nationwide deployment of this policy would have resulted in an annual revenue loss of over one billion dollars—a substantial amount, even in a major industry like the one we study. Furthermore, we show that these losses can be largely attributed to an accentuation of undesirable multitasking incentives: the reduction in sales that we find is driven by the model group incentivized with a smaller bonus.

Our paper shows that although loss framing may indeed motivate additional effort, this effort may be directed toward actions that harm overall performance. These findings help explain the rarity of loss-framing and penalty contracts in practice (Baker, Jensen and Murphy, 1988; Lazear, 1991)—a rarity that previously may be considered surprising given the substantial evidence suggesting loss aversion among workers, and given the finding that loss-framing is not shunned by agents themselves (Imas, Sadoff and Samek, 2016; de Quidt, 2017).

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 considerations (List, 2002, 2003, 2004*a*,*b*), our results clearly demonstrate 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—due to the small management team of many dealerships, 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

[3]Absent randomization, the effect of automotive policies is notoriously difficult to evaluate given the sensitivity of sales to, e.g., fluctuations in gas prices (Busse, Knittel and Zettelmeyer, 2013).

[4]This value is calculated based on 140 dealers losing 2.3 vehicles per month over the four month treatment window, applying a conservative average sales price of \$35,000 from Kelley Blue book data.

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 considerations may apply to the analysis of firms in quantitatively important ways.[5]

Our results also contribute to a better understanding of the forces governing the automobile market. Manufacturers face substantial legal restrictions governing their ability to control sales incentives: they cannot sell directly to consumers, and are highly restricted in the means by which they can vary prices offered to dealers. Consequently, manufacturers have few options other than the incentive contracts that we study to address misaligned incentives in the retail channel. The bonuses studied here should not be viewed as comparatively unimportant "deal sweeteners"—as in the common usage of the term bonus—but instead as the primary means of compensating dealers in the hopes of aligning incentives in the market. Our results are therefore of large-scale economic importance independent of their bearing on behavioral economic theory.

The remainder of this article proceeds as follows. Section II presents a theoretical examination of the positive and negative consequences of loss-framed incentives. Section III describes the setting, design, and results of our field experiment. Section IV concludes.

II. A Theoretical Examination of the Consequences of Loss Framing

In this section we present a model that clarifies how inducing loss-averse evaluation of incentives can help or hinder productivity.

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 and benefits of sales activities. To mitigate the divergence in incentives, and in particular the incentive to under-invest 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.

Using this model, we consider the wisdom of 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 be against the principal's best interests under relatively common conditions.

While we use the terminology of "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 means to motivate among his workers.

[5]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). See Malmendier (2018) for a review of the behavioral corporate finance literature.

A. Seller's Decision Problem

The agent, or *seller*, has a window of time over which outcomes are evaluated (e.g., one month). This seller's utility is determined by $u(\phi(s), c, \boldsymbol{z})$, where $\phi(s)$ denotes the payment received for completing s sales and c denotes the net private costs that the seller incurs in the sales process. The incentive scheme $\phi(s)$ is meant to promote the pursuit of sales, so $\phi(s_1) \geq \phi(s_2)$ if $s_1 > s_2$. Utility is assumed to be increasing in payments ($\phi(s)$) and decreasing in costs (c). The vector $\boldsymbol{z}$ denotes other choice-relevant variables. This term serves no role in our analyses, and is included merely to assure that our results permit for substantial heterogeneity in utility functions.

Over the window of time considered, the seller faces a long sequence of individual decisions of 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 consequences of the choice of a selling strategy is the determination of the joint distribution of sales (s), private costs (c), and other choice relevant variables ($\boldsymbol{z}$). Denote this distribution by $f(s, c, \boldsymbol{z}|\sigma)$.

The seller's decision problem is therefore to choose $\sigma \in \Sigma$ to maximize expected utility, given by $\mathbb{E}[u|\sigma] = \int u(\phi(s), c, \boldsymbol{z})\, df(s, c, \boldsymbol{z}|\sigma)$.

B. Inducing Loss Aversion

The principal considers a modification to the incentive scheme (e.g., by applying a clawback incentive) that induces loss aversion relative to a target level of sales (R). If loss aversion is induced, the seller's utility function is modified to be $u^{\Lambda}(s, c, \boldsymbol{z}) = u(\phi(s), c, \boldsymbol{z}) - \Lambda \cdot (\phi(R) - \phi(s)) \cdot I(\phi(s) < \phi(R))$. Note that I 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 the reference point from 0 to R under the assumption of piecewise-linear gain/loss utility.[6]

In order to assess the impact of inducing loss aversion on incentives, we will consider how its introduction influences choices between two selling strategies. The influence

[6]Formally, assume total utility is governed by $f_1(\phi(s), c, \boldsymbol{z}) + \eta f_2(\phi(s) - \phi(R))$, with $f_2(\phi(s) - \phi(R)) = (\phi(s) - \phi(R)) + (\lambda - 1) \cdot (\phi(s) - \phi(R)) \cdot I(\phi(s) < \phi(R))$. This model includes arbitrary "standard" utility represented by f_1, augmented with the additional consideration of f_2 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 $u(\phi(s), c, \boldsymbol{z})$ to incorporate both standard utility as well as utility derived from gains/loss evaluation without excess weighting of losses: $f_1(\phi(s), c, \boldsymbol{z}) + \eta(\phi(s) - \phi(R))$. 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.

will ultimately be governed by a simple statistic that summarizes the losses a particular selling strategy makes possible. Define the *loss exposure of strategy* σ as

$$L(\sigma) = \int (\phi(R) - \phi(s)) \cdot I(\phi(s) < \phi(R))\, df(s, c, \boldsymbol{z}|\sigma) \tag{1}$$

Using this definition, we may completely characterize the types of strategies that are incentivized by the induction of loss aversion. When characterizing the strategies that are incentivized, we will refer to loss aversion making strategy σ_1 more attractive relative to σ_2 if $\mathbb{E}[u^\Lambda|\sigma_1] - \mathbb{E}[u^\Lambda|\sigma_2] > \mathbb{E}[u|\sigma_1] - \mathbb{E}[u|\sigma_2]$. We say that strategy σ_1 is preferred to strategy σ_2 if $\mathbb{E}[u|\sigma_1] > \mathbb{E}[u|\sigma_2]$, with the relevant definition of individual utility u applied.

Proposition 1 (The impact of inducing loss aversion)**.** *Consider two selling strategies, σ_1 and σ_2. Inducing loss aversion makes σ_1 more attractive relative to σ_2 if and only if $L(\sigma_1) < L(\sigma_2)$. Furthermore, if the additional weight placed on losses (Λ) is sufficiently large, then σ_1 is preferred to σ_2 if and only if $L(\sigma_1) \leq L(\sigma_2)$.*

PROOF:

Note that expected utility with loss aversion induced may be expressed as $\mathbb{E}[u^\Lambda|\sigma] = \mathbb{E}[u|\sigma] - \Lambda \cdot L(\sigma)$. From this formulation, it follows immediately that the difference in expected utility resulting from σ_1 and σ_2 is $\mathbb{E}[u^\Lambda|\sigma_1] - \mathbb{E}[u^\Lambda|\sigma_2] = (\mathbb{E}[u|\sigma_1] - \mathbb{E}[u|\sigma_2]) - \Lambda(L(\sigma_1) - L(\sigma_2))$. The first term of this expression is simply the utility difference experienced absent the induction of 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.

To establish the second claim, consider an arbitrary pair of strategies drawn from Σ, denoted σ_i and σ_j. Restricting attention to pairs satisfying $L(\sigma_i) = L(\sigma_j)$, note that it trivially holds that σ_i is preferred to σ_j if and only if $L(\sigma_i) \leq L(\sigma_j)$. Turning attention to all other pairs, assign labels such that $L(\sigma_i) < L(\sigma_j)$. Note that $\mathbb{E}[u^\Lambda|\sigma_i] > \mathbb{E}[u^\Lambda|\sigma_j]$ if and only if $\Lambda > \frac{\mathbb{E}[u|\sigma_i] - \mathbb{E}[u|\sigma_j]}{L(\sigma_i) - L(\sigma_j)} \equiv T_{i,j}$. Thus, if $\Lambda > \max_{i,j} T_{i,j}$, for arbitrarily drawn pairs we are assured that σ_i is preferred to σ_j if and only if $L(\sigma_i) \leq L(\sigma_j)$. The second claim then follows. $\square$

Proposition 1 establishes clearly 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 through the induction of loss aversion.

C. Consequences of Inducing Loss Aversion for the Choice of Selling Strategy

Using the framework presented and the results of Proposition 1, we may now assess the classes of selling strategies that are incentivized by the induction of loss aversion.

We begin by establishing a version of the "common wisdom" of the value of loss-averse incentive schemes.

Corollary 1 (Motivation of dominant sales strategies)**.** *Consider two strategies, σ_1 and σ_2. Assume that $f(s|\sigma_1)$ first-order stochastically dominates $f(s|\sigma_2)$. The induction of* *loss aversion makes σ_1 more attractive relative to σ_2, and σ_1 is preferred to σ_2 if the additional weight placed on losses Λ is sufficiently large.*

PROOF:

Note that the assumption that $f(s|\sigma_1)$ first-order stochastically dominates $f(s|\sigma_2)$ implies that $L(\sigma_1) < L(\sigma_2)$. The claims therefore follow immediately from Proposition 1. $\square$

The implication of Corollary 1 is straightforward. Consider a case where there are activities a seller could take that lead to first-order stochastically dominant distributions of sales achieved. This aligns well with informal notions of "working harder." The principal will always prefer that the seller pursue this course of action, but private costs may dissuade the seller from taking it. In such a case, the induction of loss aversion helps motivate the desired sales strategy, and is guaranteed to successfully lead the seller to prefer it if he is sufficiently loss averse.

We believe that Corollary 1 captures the intuition that leads both researchers and practitioners to state that loss aversion serves as a positive force for motivating performance. And indeed, this interpretation is correct in several common decision environments. When there is no uncertainty in the sales that arise from any available sales strategies, a higher-sales strategy always satisfies the first-order stochastic dominance condition of Corollary 1. Similarly, if a worker's decision is to choose the probability that a binary outcome yielding a fixed bonus is earned (with greater probability requiring higher effort costs), a higher-output strategy again satisfies the conditions of Corollary 1. Imas, Sadoff and Samek (2016) and de Quidt (2017) model workers facing this latter situation, which rationalizes why they prove that loss framing has unambiguously positive effects on effort provision in the absence of selection considerations.

In many other environments, however, Corollary 1 does not apply. In practice, sellers often must make decisions between strategies unordered by first-order stochastic dominance, trading off risk and reward. We turn to characterizing the potentially undesirable incentives induced in those decisions.

Corollary 2 (Motivation for costly avoidance of loss exposure)**.** *Consider two strategies, σ_1 and σ_2, for which $\mathbb{E}[s|\sigma_1] < \mathbb{E}[s|\sigma_2]$ and $L(\sigma_1) < L(\sigma_2)$. The induction of loss aversion makes σ_1 more attractive relative to σ_2, and σ_1 is preferred to σ_2 if the additional weight placed on losses Λ is sufficiently large.*

PROOF:

Follows immediately from Proposition 1. $\square$

Corollary 2 illustrates a significant, negative consequence of the utilization of loss framing. In many environments, we imagine the principal to be risk-neutral, aiming to incentivize the agent to pursue selling strategies with the highest expected return. As we see here, the induction of loss aversion can run counter to that goal. While increasing the seller's perceived marginal returns in the loss domain does increase motivation for productivity, it also accentuates the kink in utility occurring at the reference point. This kink generates first-order risk aversion, generating a strong incentive to pursue lower-expected-return strategies that entail lower risks of loss.[7] For sufficiently large loss exposure, the seller will decline selling strategies with arbitrarily large risk premia in higher expected sales merely to avoid marginal increases in exposure to losses.

D. Incentives with Multidimensional Performance

The analysis of the prior section assumes that there is a single dimension of performance evaluation. In most applications, however, performance is multidimensional. For example, the seller may have to sell multiple products, or might sell the same product in several different windows of time. We now turn to evaluating the consequences of dimension-specific induction of loss aversion in such cases.

Let (s^d, c^d) denote the amount of sales and private costs arising from sales activities in dimension $d \in D$. Let ϕ^d denote the potentially dimension-specific incentive scheme. Let $(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{z})$ denote the full vector of utility inputs across all dimensions. The seller's decision is the choice of a *multidimensional selling strategy*, which constitutes a complete contingent management plan of sales activities over all dimensions. Continue to denote an individual selling strategy as σ and the (finite) set of possible selling strategies as Σ. As before, the decision-relevant consequences of the choice of a multidimensional selling strategy is the determination of the joint distribution of $(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{z})$.

The seller's multidimensional utility function is defined as $U(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{z}) = \sum_{d\in D} u^d(\phi^d(s^d), c^d, \boldsymbol{z}^d)$, and is modified to $U^\Lambda(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{z}^d) = \sum_{d\in D} u^d(\phi^d(s^d), c^d, \boldsymbol{z}^d) - \Lambda(\phi^d(R^d) - \phi^d(s^d)) \cdot I(\phi(s^d) < \phi(R^d))$ when loss aversion is induced. The multidimensional expected utility associated with strategy $\boldsymbol{\sigma}$ is given by $\mathbb{E}[U|\boldsymbol{\sigma}] = \int U(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{z})\, df(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{z}|\boldsymbol{\sigma})$.

In this framework, we may define the *multidimensional loss exposure of strategy* σ as

$$(2) \qquad L^M(\sigma) = \int \sum_{d\in D} (\phi^d(R^d) - \phi^d(s^d)) \cdot I(\phi^d(s^d) < \phi^d(R^d))\, df(\boldsymbol{s}, \boldsymbol{c}, \boldsymbol{z}|\sigma)$$

In our assumed utility function, and thus in our definition of loss exposure, we assume that losses across each dimension are weighted equally. Note that this assumption is of little conceptual importance, and similar results would follow (with more cumbersome notation) if we allowed dimension-specific weights on losses.

In this more complex environment, we may establish close analogs to Proposition 1 and its resulting corollaries.

Proposition 2 (The multidimensional impact of inducing loss aversion)**.** *Consider two multidimensional selling strategies,* σ_1 *and* σ_2*. Inducing loss aversion makes* σ_1 *more attractive relative to* σ_2 *if and only if* $L^M(\sigma_1) < L^M(\sigma_2)$*. Furthermore, if the additional weight placed on losses* (Λ) *is sufficiently large, then* σ_1 *is preferred to* σ_2 *if and only if* $L^M(\sigma_1) \leq L^M(\sigma_2)$.

[7] The prior work of Armantier and Boly (2015) and de Quidt (2017) also notes that changing the reference point can have ambiguous effects on effort provision due to changes in risk tolerance. In these analyses, the potential for ambiguous effects is not attributed to loss aversion, but instead to another component of prospect theory: diminishing sensitivity. Diminishing sensitivity generates risk aversion over gains and risk inclination over losses, immediately yielding changes in risk attitudes as the reference point moves. In our model we intentionally exclude diminishing sensitivity and illustrate that loss framing can be undesirable even in its absence.

PROOF:

Note that, closely following in Proposition 1, multidimensional expected utility in the presence of loss aversion may be expressed as $\mathbb{E}[U^\Lambda|\sigma] = \mathbb{E}[U|\sigma] - \Lambda \cdot L^M(\sigma)$. From this formulation, it follows immediately that the difference in expected utility resulting from σ_1 and σ_2 is $\mathbb{E}[U^\Lambda|\sigma_1] - \mathbb{E}[U^\Lambda|\sigma_2] = (\mathbb{E}[U|\sigma_1] - \mathbb{E}[U|\sigma_2]) - \Lambda(L^M(\sigma_1) - L^M(\sigma_2))$. The last term is positive if and only if $L^M(\sigma_1) < L^M(\sigma_2)$, establishing the first claim of the proposition.

To establish the second claim, consider an arbitrary pair of multidimensional selling strategies drawn from Σ, denoted σ_i and σ_j. Restricting attention to pairs satisfying $L^M(\sigma_i) = L^M(\sigma_j)$, note that it trivially holds that σ_i is preferred to σ_j if and only if $L^M(\sigma_i) \leq L^M(\sigma_j)$. Turning attention to all other pairs, assign labels such that $L^M(\sigma_i) < L^M(\sigma_j)$. Note that $\mathbb{E}[U^\Lambda|\sigma_i] > \mathbb{E}[U^\Lambda|\sigma_j]$ if and only if $\Lambda > \frac{\mathbb{E}[U|\sigma_i] - \mathbb{E}[U|\sigma_j]}{L^M(\sigma_i) - L^M(\sigma_j)} \equiv T_{i,j}$. Thus, if $\Lambda > \max_{i,j} T_{i,j}$, for arbitrarily drawn pairs we are assured that σ_i is preferred to σ_j if and only if $L^M(\sigma_i) \leq L^M(\sigma_j)$. The second claim then follows. □

Proposition 2 establishes that the same logic established in Proposition 1 transfers to the multidimensional environment. Rather than considering a single dimension's exposure to loss, the key statistic now becomes the summation of dimension-specific exposures to loss. With access to this theorem, the following corollary follows immediately.

Corollary 3 (Incentive for multidimensional gaming)**.** *Consider two multidimensional strategies, σ_1 and σ_2, for which $\mathbb{E}[\sum_{d \in D} s^d|\sigma_1] < \mathbb{E}[\sum_{d \in D} s^d|\sigma_2]$, and for which $L^M(\sigma_1) < L^M(\sigma_2)$. The induction of loss aversion makes σ_1 more attractive relative to σ_2, and σ_1 is preferred to σ_2 if the additional weight placed on losses Λ is sufficiently large.*

Like Corollary 2, Corollary 3 merely highlights a worrying special case of the proposition. Corollary 3 establishes that, for sufficiently loss-averse individuals, the desire to minimize exposure to losses is sufficient to motivate the pursuit of strategies that result in lower total expected sales. Unlike in Corollary 2, the manner in which this result arises does not require the presence of uncertainty, but merely the ability to choose to forego large gains in one dimension to avoid small losses in another.

These results on the consequences of multidimensional assessment of performance conceptually align with the experimental findings of Rubin, Samek and Sheremeta (2018), which explore the impact of loss aversion in a domain with separate incentives for quantity and quality.

E. Illustrative Examples

To help convey intuitions, we present three simple examples that illustrate the key findings above.

In all examples, we consider a seller with the utility function $u(\phi(s), c) = \phi(s) - c$. We assume the incentive contract is a simple piece-rate providing \$1 per sale: $\phi(s) = s$.

In each example, we will consider the incentives to pursue two discrete strategies. Strategy 2 will always be preferred by a risk-neutral principal. We will consider the consequences of the risk-neutral principal inducing loss aversion relative to a target of 10 sales. We assume that the seller's excess weight placed on losses is $\Lambda = 2$.

Example 1: Increased Sales in Deterministic Setting

In this example, the seller has two potential selling strategies. Strategy 1 yields 8 sales with certainty, but requires the seller to incur 6 units of effort costs. Strategy 2 yields 12 sales with certainty, but requires the seller to incur 11 units of effort costs.

While strategy 2 yields higher sales and thus higher incentive payments, these payments come at the expense of higher effort costs incurred by the seller. The seller's utility is higher pursuing strategy 1:

$$u(\phi(s_1), c_1) = 8 - 6 = 2 > 1 = 12 - 11 = u(\phi(s_2), c_2). \tag{3}$$

Now consider the consequence of the principal inducing loss aversion. This modifies the utility evaluation to:

$$u^{\Lambda}(\phi(s_1), c_1) = 8 - 6 - \Lambda \cdot 2 = -2 < 1 = 12 - 11 - \Lambda \cdot 0 = u(\phi(s_2), c_2). \tag{4}$$

The worker would therefore choose strategy 2 with loss aversion induced.

Loss aversion's induction of greater effort in this example follows from a direct application of Corollary 1. In a deterministic environment such as this, the sole consequence of inducing loss aversion is increasing the perceived benefits of incurring additional effort costs to produce amounts considered gains.

Example 2: Reduced Sales in Stochastic Setting

In this example, the seller has two potential selling 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 the induction of loss aversion, the seller's utility is higher pursuing strategy 2:

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

Now consider the consequence of the principal inducing loss aversion relative to a sales target of 10. This modifies the utility evaluation to:

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

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 ($\frac{1}{8}$) yields $L(\sigma_2) = \frac{6}{8}$, generating the inequality above. The worker would therefore choose strategy 1 with loss aversion induced.

Loss aversion's induction of risk avoidance in this example follows from a direct application of Corollary 2. In this environment, the consequence of inducing loss aversion is generating aversion to the loss-generating outcomes that are possible under strategy 2, even though strategy 2 yields higher sales in expectation.

Example 3: Reduced Sales in Multidimensional Setting

In this example, we modify the setting to include a second dimension of production. We will refer to the two dimensions as good 1 and good 2.[8] Both goods are subject to the dimension-specific utility function $u(\phi(s), c) = \phi(s) - c$, and to the same incentive contract: $\phi(s) = s$. Throughout this example, we will assume 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 the induction of loss aversion, the seller's utility is higher pursuing strategy 2:

$$U(\boldsymbol{s}_1, \boldsymbol{c}_1) = 10 + 10 = 20 < 21 = 9 + 12 = U(\boldsymbol{s}_2, \boldsymbol{c}_2)$$

Now consider the consequence of the principal inducing loss aversion relative to a sales target of 10 for each good. This modifies the utility evaluation to:

$$U^{\Lambda}(\boldsymbol{s}_1, \boldsymbol{c}_1) = 10 + 10 - \Lambda \cdot L^M(\boldsymbol{\sigma}_1) = 20 > 19 = 9 + 12 - \Lambda \cdot L^M(\boldsymbol{\sigma}_2) = U^{\Lambda}(\boldsymbol{s}_2, \boldsymbol{c}_2)$$

Note that strategy 1 has no loss exposure, so $L^I(\boldsymbol{\sigma}_1) = 0$. Strategy 2, however, results in a loss of 1 for good 1. Aversion to that loss leads the seller to prefer strategy 1, even though it results in lower total sales and lower total bonus payments.

Loss aversion's negative consequences in this example follow from a direct application of Corollary 3. In this case, the aim to avoid losses drives the seller to attend to domain-specific targets at the cost of aggregate sales maximization.

F. Summary

Researchers and practitioners have taken for granted that an immediate implication of loss framing is increased incentive for performance. We have articulated ways in which this statement does, and does not, hold. As we illustrate, the effect of inducing loss framing is completely characterized by introducing an incentive to avoid losses. While this statement may appear nearly tautological, we note that it meaningfully differs from the common claim in prior literature that loss framing motivates effort. The addition of a strong incentive for risk mitigation, as well as increased incentives for cross-incentivized-dimension gaming, can lead the seller to pursue a variety of activities that negatively impact a risk-neutral principal. As a result, in a number of common settings concerning worker compensation, the induction of loss aversion comes with substantial downsides. In the next section, we demonstrate the consequences of this induction in such a setting.

III. An Empirical Examination of Loss Framing in Automobile Sales

In this section we present the field experiment that motivated our theoretical exercise. We begin by providing institutional details necessary for understanding the market, fol-

[8] Note that good two may equivalently be thought of as a second (undiscounted) period of sales for good 1.

lowed by a detailed description of the experiment that was deployed. We then analyze the effect of the induction of loss aversion on sales, and the manners in which selling behavior was modified to achieve these aggregate outcomes.

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 prohibit 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.[9] The sales activity of these dealers is our topic of focus in this paper.

Independent dealers are operated under franchise agreements that significantly shape the automobile market. Existing franchise laws restrict a manufacturer's ability to open or close dealers, heavily limiting manufacturer's 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 dealers per state. With the lack of entry and exit and the limits on consolidation, the dealers for any given manufacturer represent a wide mix of ownership that includes 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.[10] Dealers then have discretion to price the vehicles independently before selling to consumers at a negotiated price (Bennett, 2013; Busse and Silva-Risso, 2010). 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 effectively reduce vehicle price if

[9]See Lafontaine and Scott Morton (2010) or Murry and Schneider (2016) for detailed information on the history and function of car dealership franchises. Murry (2018) explains why competition among these dealers leads to underinvestment in advertising.

[10]See Cachon and Olivares (2010) and Olivares and Cachon (2009) for discussions of the challenges of managing inventory under this system.

there is excessive 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). These direct incentive programs are the policy of interest in this paper.

Our car manufacturer of interest has multiple vehicle models sold at over a thousand dealers across all fifty states. Since our access to data is governed by a non-disclosure agreement, we will refer to this manufacturer as CarCo.[11] Like many manufacturers, CarCo uses incentive programs for specific vehicle models to motivate sales volume at the vast majority of its dealers. The specific program considered in this paper has two separate incentive schemes, each focused on particular set of models. We refer to these sets of models as Model Group A and Model Group B. Both groups are typically, but not always, retailed together at the same dealerships: 88% of dealers in the 2017 program sold both model groups.

Like most dealer incentive programs, CarCo's program rewards dealers for reaching monthly goals.[12] These quota-based systems 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, 2013). Dealers receive specific monthly sales goals for each model group based on their average sales in that calendar month in the previous four years. All vehicles that are sold to consumers count toward the target, with all vehicles of that model group counting equally.[13]

Figure 1 illustrates CarCo's incentive program. Dealers selling below their monthly target earn no bonus. Dealers selling between 100% and 110% of their monthly target receive \$100 for each vehicle sold in that month. Dealers selling 110% or more receive \$700 per vehicle for Model Group A or \$800 per vehicle for Model Group B. 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% threshold conveys a large fixed bonus.[14] To illustrate, applying the average targets reported in Table 1, selling the single vehicle on the margin of the 110% threshold would yield $\$19,300$ or $\$9,200$ in a given month, respectively. For the largest dealers, this marginal car can be worth over $\$200,000$.

Participation in this program is voluntary, but the vast majority of dealers choose to

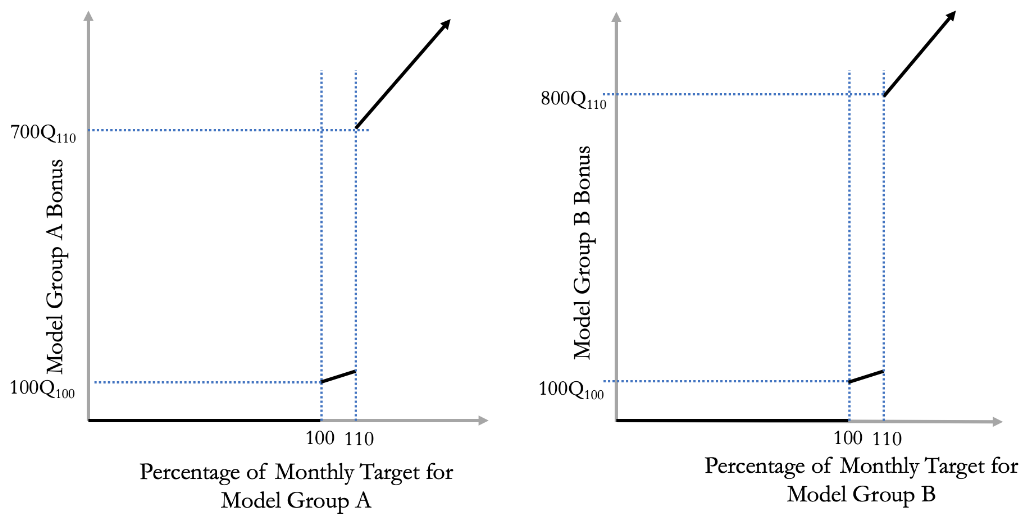
[11]Our legal agreement allowed CarCo to correct the paper for factual accuracy and anonymity, but we had full rights to publish our results.

[12]For 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.

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

[14]Interviews with dealers consistently indicated that they view the 100% goal as mostly irrelevant, or a small consolation prize for missing the key 110% level.

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.

participate by paying a fixed monthly fee.[15] 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 non-participating dealer thus cannot compete on price without losing money, because they will not enjoy the month-end bonus from reaching the sales target.

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 exert more resources toward attracting and selling to customers, either through managerial oversight or incentivizing salespeople. Second, they can increase their advertising spending to attract more customers. Third, they can price more aggressively, accepting lower (and often negative) margins to close deals. Fourth, they can work more aggressively to qualify buyers for financing, which effectively lowers the vehicle price.[16]

Dealers also have several ways in which they can reach monthly sales targets that involve "gaming" the incentive system to reduce the risk of missing incentive thresholds

[15]Most non-participants are very small dealers. The program contract operates on an annual basis, such that 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.

[16]Note 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.

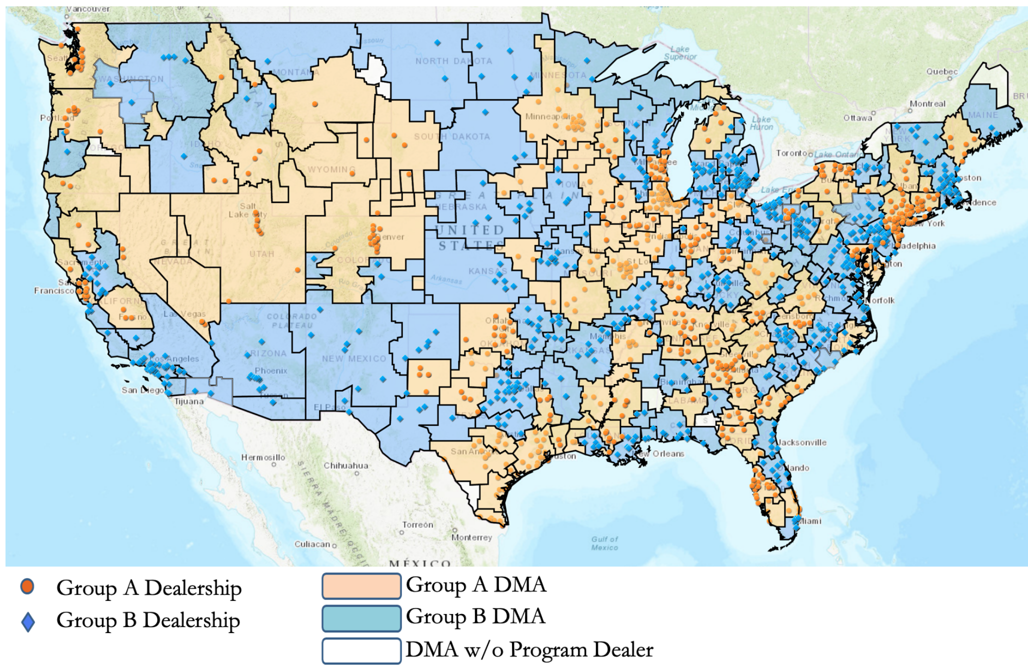
(Larkin, 2014; Oyer, 1998; Courty and Marschke, 2004). First, dealers can attempt to move customers across calendar months. If the dealer is near the crucial 110% threshold, they can attempt to accelerate the purchase decision of a customer by offering a lower price or prioritizing the paperwork and financing. If they are clearly not going to achieve the threshold in a given month, they can attempt to delay the closing of the sale until after the month-end to raise the likelihood of achieving the threshold in the next month. Second, dealers can potentially move customers across model groups if there are multiple vehicles consistent with the customer's preferences. This push to the customer's second-best alternative can be similarly motivated with price cuts, financing, or trade-in value. Although these gaming behaviors might mitigate the risk of missing the large payoffs from hitting thresholds, it can be costly to the manufacturer because it risks driving customers to the brands of other manufacturers. More directly, increased gaming is costly for the manufacturer because they must pay larger bonuses without increasing sales volume.

B. Experimental Design

The experiment in this paper originated with a CarCo brand manager's interest in using loss aversion 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 identification number (VIN) 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. This new incentive design was similar to those deployed to factory workers in Hossain and List (2012) or to teachers in Fryer Jr et al. (2012). 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 a pilot study that would assess the new scheme's efficacy. 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, there were reasons to believe 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 (Larkin, 2014; Oyer, 1998; Benson, 2015; Ederer, Holden and Meyer, 2018; Steenburgh, 2008), particularly those with non-linear returns to performance. Finally, given the heterogeneity of organizational forms, management style, and sophistication, it was unclear how many dealers would respond "behaviorally" and how many would treat advanced payments as an interest-free short-term loan. If indeed the firms treated prepayments as free capital, and that free capital did not enable productivity-enhancing investments, then the loss-framed contracts were inherently costly to CarCo.

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.

The initial experimental concept was a randomized controlled trial with all dealers block-assigned to two conditions by 199 DMAs.[17] 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 aspects of the franchise agreement forced adjustments in the research design. First, we were required to ensure that each participating dealer received equal treatment. 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

[17]Block assignment was important in the design because inter-dealer competition for customers might contaminate the control group.

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 online appendix (but provides similar evidence of negative treatment effects when appropriate control groups are formed).

Second, CarCo asked us to change the group assignment of two DMAs due to concerns about competition spreading between contiguous DMAs in the same state. In both cases, we exchanged the requested DMA with a DMA from the other group that most closely matched the pre-trends in the region.

Finally, shortly before implementation, CarCo attorneys determined that the franchise agreement required that any change to the incentive program for a given dealer necessitated dealer approval. We therefore created an opt-in process 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 either for four months starting in May, 2017 or September, 2017. Dealers opting out would 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 1,226 dealers in the incentive program, 294 (24%) chose to participate. Additionally, 335 explicitly opted out of the program, while 597 failed to respond.

The resulting 294 dealers, which represented 116 DMAs in 47 states, are shown in Figure 2. 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.[18] All but 29 of these deals sold both model groups: 18 sold only Model Group A vehicles and 11 sold only Model Group B. 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 the bonus that would be earned if their 110% target were achieved ($700 per vehicle for Model Group A and $800 per vehicle for Model Group B). At the end of the month, total sales for each dealer were calculated to determine the actual earned monthly bonus. If the dealer sold less than the 110% target for a given Model Group, the excess prepayment was debited directly from their account. If the dealer sold strictly more than that target, an additional payment was made. 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, so each dealer fully participated for the eight-month span.

C. Summary Statistics and Baseline Response to Incentives

Before proceeding to analysis of the effects of our program, we present simple summary statistics on the sales activities of firms by treatment assignment and participation status. We additionally provide an examination of the response to the incentive scheme

[18]See Appendix Figure A.1 for the density of dealers within DMAs across conditions.

Table 1—: Summary Statistics

	MODEL GROUP A								
	Treatment Group			Control Group			By Participation Status		
	Pre	Post	Total	Pre	Post	Total	In	Out	All
Vehicles Sold	29.52 (21.44)	31.41 (23.72)	30.46 (22.61)	39.19 (52.55)	41.05 (52.13)	40.12 (52.33)	35.45 (41.02)	27.26 (29.18)	29.25 (32.65)
100% Target Level	26.98 (19.36)	32.15 (22.03)	29.57 (20.89)	31.38 (33.30)	38.59 (41.69)	34.98 (37.88)	32.36 (30.96)	27.26 (24.65)	28.50 (26.41)
Hit 110% Target	0.56 (0.50)	0.43 (0.50)	0.50 (0.50)	0.64 (0.48)	0.50 (0.50)	0.57 (0.50)	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)	0.69 (0.46)	0.55 (0.50)	0.62 (0.49)	0.59 (0.49)	0.53 (0.50)	0.54 (0.50)
Below 75% Target	0.16 (0.37)	0.29 (0.45)	0.22 (0.42)	0.15 (0.35)	0.22 (0.41)	0.18 (0.39)	0.20 (0.40)	0.26 (0.44)	0.25 (0.43)
Below 50% Target	0.03 (0.16)	0.05 (0.22)	0.04 (0.19)	0.03 (0.17)	0.03 (0.18)	0.03 (0.18)	0.03 (0.18)	0.06 (0.24)	0.06 (0.23)

	MODEL GROUP B								
	Treatment Group			Control Group			By Participation Status		
	Pre	Post	Total	Pre	Post	Total	In	Out	All
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)
100% Target Level	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)
Below 75% Target	0.20 (0.40)	0.33 (0.47)	0.26 (0.44)	0.20 (0.40)	0.31 (0.46)	0.25 (0.43)	0.26 (0.44)	0.31 (0.46)	0.30 (0.46)
Below 50% Target	0.07 (0.26)	0.13 (0.34)	0.10 (0.30)	0.07 (0.26)	0.10 (0.31)	0.09 (0.28)	0.09 (0.29)	0.12 (0.33)	0.12 (0.32)

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 hitting earning the larger fixed bonus for exceeding the 110% threshold, the smaller fixed bonus for exceeding the standard target threshold, as well as the probability of having particularly bad months with sales below 75% or 50% of the target. The seemingly identical pre-period Model Group B percentages for the two participant groups is due to rounding, and they are slightly different at the third decimal point.

that existed before the treatment period.

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 mere comparisons of means presented in this table, we note that our results are ultimately foreshadowed by such comparisons.

Turning attention first to the Model Group A data for monthly vehicles sold, we note that the control group saw a modest increase in sales when comparing the pre- and post-treatment windows (39.19 vs. 41.05, an increase of 1.86). For comparison, the treatment group saw effectively the same increase (29.52 vs. 31.41, an increase of 1.89; difference-in-differences: 0.03). An analogous comparison of the rate of achieving the 110% target suggests that treatment increases the chance of earning the large fixed component of the bonus by 1 percentage point.[19] In short, simple comparisons of means suggest that the treatment was associated with quite modest increases in Model Group A sales and bonus-condition attainment.

Turning attention next to the Model Group B data for monthly vehicles sold, we note that the control group again saw a modest increase in sales when comparing the pre- and post- treatment windows (17.22 vs. 18.80, an increase of 1.58). For comparison, the treatment group saw a decrease in sales (15.14 vs. 14.16, a decrease of 0.98; difference-in-differences: -2.56). An analogous comparison of the rate of achieving the 110% target suggests that treatment decreases the chance of earning the large fixed component of the bonus by 8 percentage points.[20] In short, a simple comparison of means suggests a comparatively large negative effect of loss framing for Model Group B sales and bonus-condition attainment.

Examination of this table also illustrates two key issues that will need to be accommodated in the empirical analysis to come.

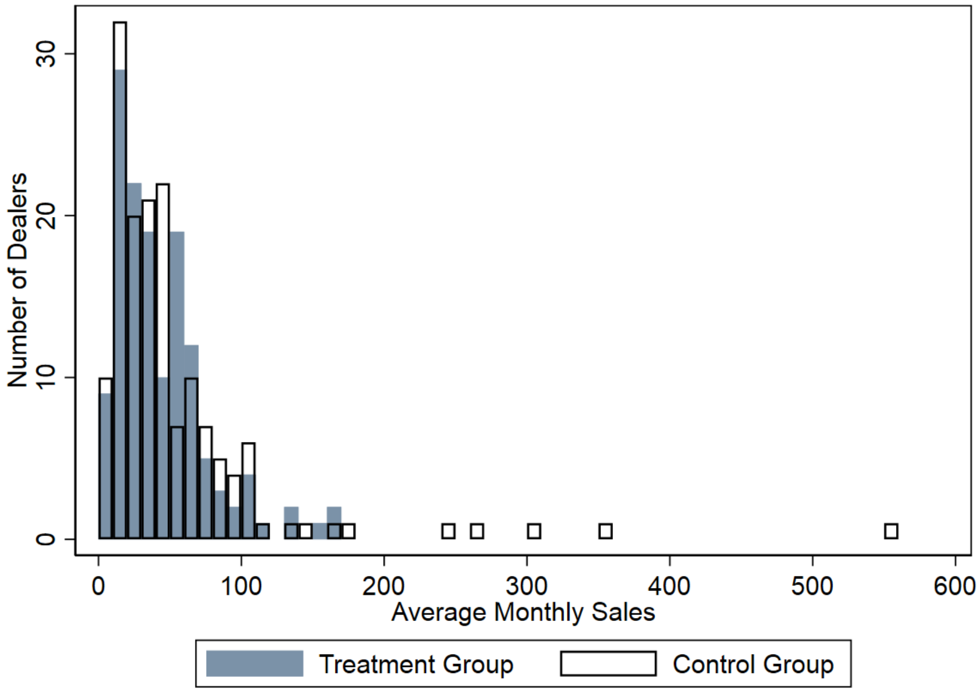
First, the clear trends seen in sales and in target attainment for the control group illustrate that these variables are not stationary in the absence of treatment. This non-stationarity is expected due to the significant seasonality in automobile sales, and suggests an important role for both a difference-in-differences design in general and flexible time controls more specifically.

Second, we note that the means of key variables (e.g., sales) show significant differences between treatment and control groups in the period prior to treatment. Such differences suggest the potential for different characteristics of dealers across treatment groups, motivating our inclusion of dealer fixed effects in some analyses to come. These differences may generate some concern about systematic problems with random assignment, but we note that 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 monthly sales. The largest dealer, which has ten times monthly sales of

[19]Control: 0.64 vs. 0.50, a decrease of 0.14. Treatment: 0.56 vs. 0.43, a decrease of 0.13. Difference-in-differences: 0.01.

[20]Control: 0.64 vs. 0.48, a decrease of 0.12. Treatment: 0.64 vs. 0.44, a decrease of 0.20. Difference-in-differences: -0.08.

Figure 3. : Histogram of Monthly Car Sales by Treatment Assignment



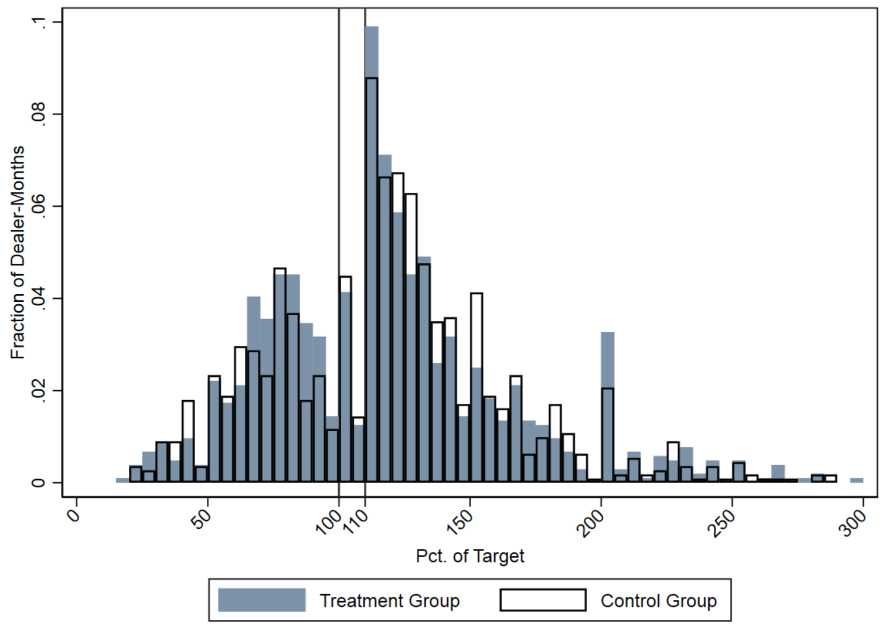
Notes: This figure shows histogram of average monthly sales (summing across both model groups) in the four-months prior to the experiment. Five of the six largest dealers are in one DMA (and thus are block-assigned). This block assignment accounts for the unbalanced right-tail of control group observations.

the average dealer, was by chance also assigned to the control group. Figure 3 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. While we do not exclude these observations from our main analysis, in Section III.G we demonstrate that our main results persist with their exclusion.

Figure 4 shows the distribution of monthly sales in the four months prior to the experiment. This figure clearly demonstrates the importance of the 110% 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.[21] Such excess mass is starkly visible in this figure, with substantial asymmetries in mass observed around both the 100% and 110% target. Note that, in the pre-period data examined in this figure, there is no statistically detectable difference in either average sales or in the distribution

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

Figure 4. : Monthly Sales (Relative to Target) by Treatment Assignment



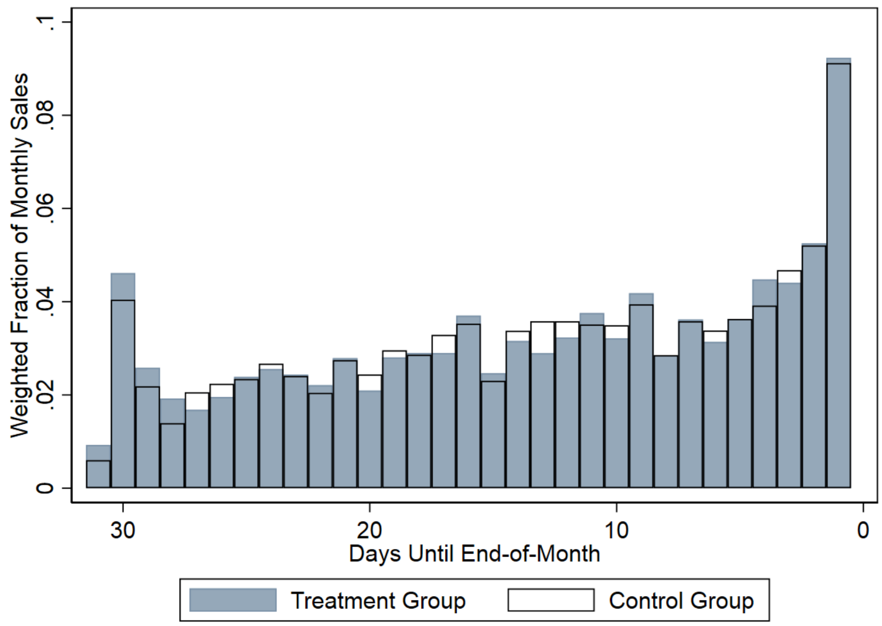
Notes: This figure shows distribution of monthly sales results in the four months prior to the experiment, expressed as a percentage of the assigned sales target. Model group results are included separately for each dealer-month. The two vertical lines represent the discrete bonus thresholds at 100%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

of sales across treatment and control groups (T-test: $p = 0.487$, Kolmogorov-Smirnov: $p = 0.176$).

Figure 5 shows the distribution of sales by day-of-the-month in each group during the pre-period, demonstrating some daily responsivity to the intertemporal incentives induced by these contracts.[22] As previously noted in, e.g., Larkin (2014), high-powered incentives measured over discrete time windows can result in high effort to close sales at the end of the window of measurement. Consistent with this consideration, we see that sales are particularly concentrated at the end of the month. Also consistent with Larkin (2014), we see increased sales at the beginning of the month, likely from when dealers delayed closing deals at the end of a month where the target was unreachable. As with our analysis of within-month responsivity to incentives, there is no statistically detectable difference in means between treatment and control groups (T-test: $p = 0.692$). There is a small distributional difference that is statistically distinguishable because of the large number $(57,005)$ of daily observations (Kolmogorov-Smirnov: $p = 0.033$).

[22]Each day-of-the-month is inverse-weighted by the number of days during the four-month period when dealers were open. Nearly all dealers were closed January 1-3 and on Sundays.

Figure 5. : Daily Sales Timing by Treatment Assignment



Notes: This figure shows the percentage of monthly sales occurring on each day of the month in the four months prior to the experiment. We correct for days when dealers are closed (Sundays and the New Year's holiday) by inverse-weighting sales by the number of open days during the four month period.

D. Impact of Loss Framing on Average Sales

The focus of our analyses is on 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 pre-existing post-payment scheme.

We initially estimate the average treatment effect on the participating dealers with OLS, regressing total units sold on dummy variables for the treatment group, the treatment period, and the interaction. All standard errors are clustered at the DMA level.

Table 2 presents results for the change in total sales. Columns 1–3 present results for total sales across both model groups, including no controls. The second column incorporates month dummies to control for flexible time trends, which absorbs the post-treatment variable. The third column incorporates dealer fixed effects, which absorbs treatment group. Columns 4–6 and 7–9 present these same regressions restricted to Model Groups A and B, respectively.

The average treatment effect in columns 1–3 implies that 2.3 fewer cars were sold per month under loss framing relative to the control group—a 5% decrease in total unit sales

Table 2—: Difference-in-Differences Estimates of Treatment Effect on Sales

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment Period X Assigned Prepay	-2.3	-2.3	-2.3	0.0	0.0	0.0	-2.6	-2.6	-2.6
	(1.25)	(1.25)	(1.33)	(0.90)	(0.90)	(0.96)	(1.31)	(1.31)	(1.40)
Treatment Period	3.2			1.9			1.6		
	(0.96)			(0.72)			(1.23)		
Assigned Prepay	-9.6	-9.6		-9.7	-9.7		-2.1	-2.1	
	(9.38)	(9.40)		(6.54)	(6.55)		(3.61)	(3.62)	
Constant	53.0	45.5	40.9	39.2	34.6	29.9	17.2	13.9	12.9
	(8.42)	(8.84)	(1.02)	(5.97)	(6.47)	(1.14)	(3.11)	(2.99)	(0.63)
Model Group	Pooled	Pooled	Pooled	A	A	A	B	B	B
Month Fixed Effects		X	X		X	X		X	X
Dealer Fixed Effects			X			X			X
Observations	2352	2352	2352	2264	2264	2264	2216	2216	2216
R^2	0.010	0.015	0.961	0.014	0.019	0.956	0.009	0.017	0.912

Notes: This table presents difference-in-differences estimates of the effect of the prepayment intervention on total vehicles sold. The dependent variable is the month-specific number of vehicles sold at the dealer level. The first three columns present analysis pooling the two model groups together, whereas the next two groups of three columns present analysis of two model groups. Standard errors are clustered at the DMA level.

relative to the pre-treatment monthly average of 48.5. 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 reasonable (but not conservative) α-levels (p-values range from 0.06 to 0.08).

To help assess the parallel-trend assumptions that are central to difference in differences designs, Figure 6 shows time-paths of sales over the first eight months of 2017.[23] This window includes the four-month pre-treatment window and the four months during which the treatment group was prepaid. Several patterns are evident from this figure. First, the observable pretrends in both groups are encouragingly similar. Second, the estimated treatment effect is evident in the larger gap between groups in months 5-8. We believe this figure provides some reassurance that our attempt to match pre-trends with our randomization procedure was successful, and some assurance that the treatment effects we study are clearly apparent in minimally processed data.

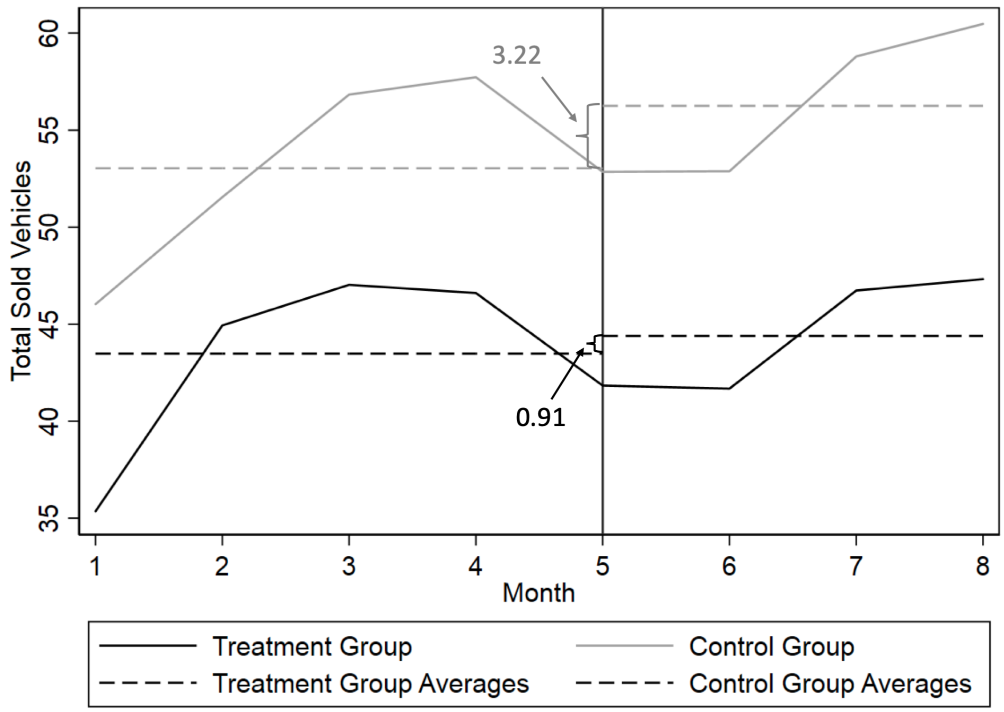
We next examine the differential effect this policy had on the two model groups under consideration. Note that in columns 4–6 of Table 2, we find no effect of the policy on sales for Model Group A. Estimates suggest an impact of 0.03 car sales per month, with this effect being statistically indistinguishable from zero. In contrast, the estimated treatment effect for Model Group B is -2.6 cars sold per month—a reduction of 17% relative to the pre-treatment average (p-values range from 0.05 to 0.07).

These findings provide support for the concern that cross-dimension gaming might contribute to a negative impact of loss framing. In our data, the overall negative effect on joint sales is driven by a large reduction in Model Group B sales, with a small (and not meaningfully offsetting) gain in Model Group A sales. This pattern is predicted by the considerations described in Section II. Recall from the discussion in Section III.A that Model Group A delivers a substantially larger fixed bonus for passing the 110% target than Model Group B. Because a failure to attain the Model-Group-A target results in a larger loss, incentives for loss avoidance raise the importance of selling Group A vehicles relative to those from Group B. The results presented thus far are consistent with some degree of diversion of Model Group B opportunities to Model Group A (with a high rate of lost customers in these attempts, and potentially with costly risk avoidance further masking the postive shift of sales to Model Group A).

To test an additional prediction of cross-dimension gaming, we examined how the estimated treatment effect evolved over the course of the month. As previously discussed in the context of Figure 5, gaming of the monthly evaluation of bonus qualification is thought to contribute to the high rate of sales occurring at the end of the month. The induction of loss aversion could accentuate this force. To test for this possibility, we re-estimated regressions (6) and (9) from Table 2, restricting the data to each possible number of days until the end of the month. Results are presented in Figure 7. As seen in the figure, we do not find evidence that treatment effects operate at a specific time of the month; however, given that the average dealer sells fewer than one car of either brand per day, on average, the size of the confidence intervals in this figure illustrate that we

[23] Similar figures by model group and by participation status are presented in Appendix A.2.

Figure 6. : Monthly Trends in Sold Vehicles for Participating Dealers



Notes: This figure shows average total dealer sales by month for both dealer groups. Vertical lines represent the first month in each of the two treatment periods. Horizontal dashed lines represent the average monthly sales across each four month period. Note that the period specific averages do not equal the sum of model-specific averages in Table 1 because some dealers only carry one model-group.

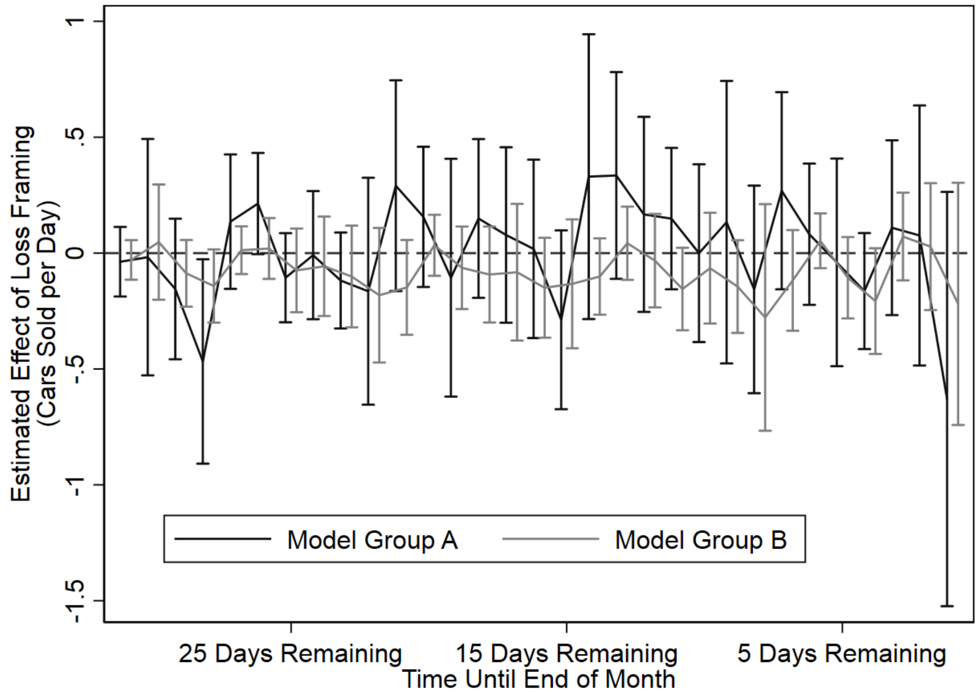
are underpowered to rule out substantial intertemporal effects.

In sum, we see evidence that the loss framing of bonus payments reduced total sales. This reduction was driven by sales decreases in Model Group B, consistent with concerns about cross-dimension gaming motivations being influenced by loss aversion. More nuanced tests of the prediction of cross-dimension gaming will be presented in Section III.E.

E. Impact of Loss Framing on Distribution of Sales

Having established that loss framing was associated with an average decline in overall sales, we now turn to assessing how this intervention affected the distribution of sales. Different loss-averse reactions—e.g., increased effort, increased risk-avoidance, or increased cross-dimension gaming—generate different predictions regarding these distributional effects. Examination of the changes in these distributions therefore provides a means of assessing the mechanisms underlying the negative consequences we have observed.

Figure 7. : Day-Specific Estimated Treatment Effects on Cars Sold



Notes: This figure shows the estimated treatment effect from regressions (6) and (9) in Table 2, with the sample restricted to include only the day of the month indicated on the X-axis.

Method for Inferring Impact of Treatment: Difference in Kernel Density Differences Estimation (DKDD)

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: (1) estimating the evolution of the distribution of final sales between the pre-period and the treatment window, and (2) 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:

$$f^{T} = (f^{(T,2)} - f^{(T,1)}) - (f^{(C,2)} - f^{(C,1)}). \tag{5}$$

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 in 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 $\hat{f}^{(g,t)}(x) = \frac{1}{N^{(g,t)}} \sum_{i=1}^{N^{(g,t)}} K(\frac{x - x_n^{(g,t)}}{h})$. Given these definitions, the Difference in Kernel Density Differences (DKDD) estimator of f^{T}, evaluated at point x, is given by

$$\hat{f}_h^{T}(x) = (\hat{f}_h^{(T,2)}(x) - \hat{f}_h^{(T,1)}(x)) - (\hat{f}_h^{(C,2)}(x) - \hat{f}_h^{(C,1)}(x)). \tag{6}$$

The consistency of equation 6 as an estimator for f^{T} follows immediately from the well-known consistency of the individual Kernel estimators, combined with Slutsky's theorem.

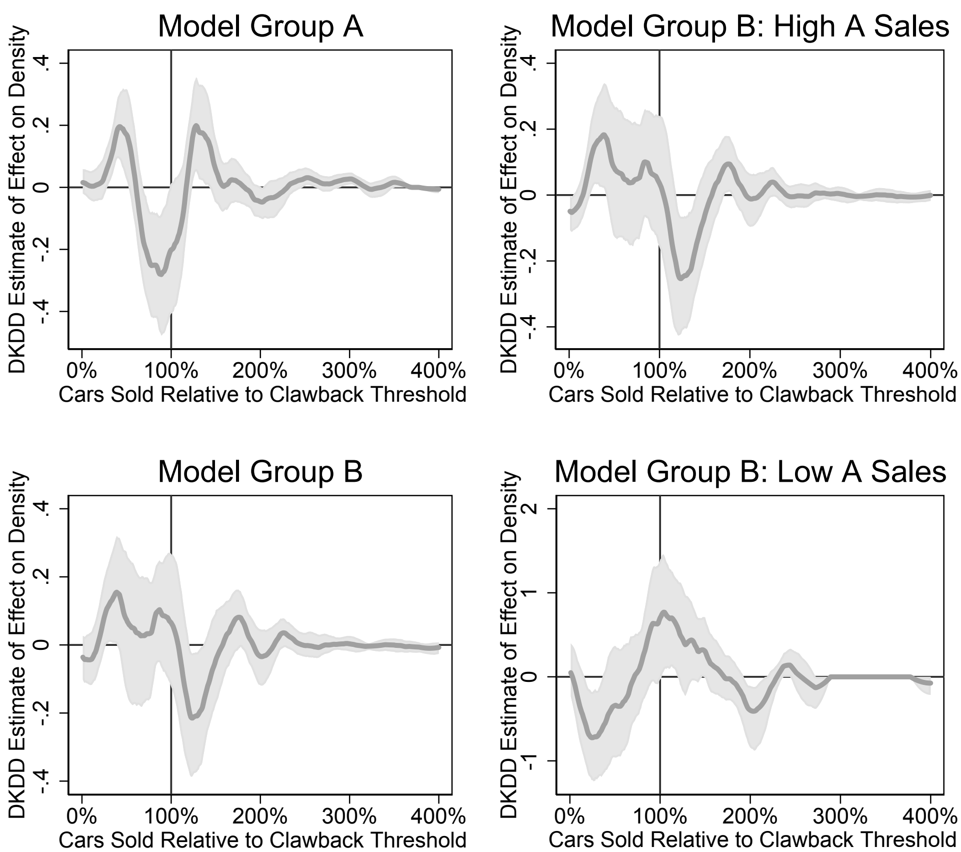
In the next section, we apply this method to infer f^{T} as induced by loss framing. Whenever presented, confidence intervals will be generated through a bootstrap procedure, recalculating $\hat{f}^{(g,t)}(x)$ from 1,000 simulated samples generated by sampling by DMA cluster (with replacement).

DKDD Estimates of the Impact of Loss Framing

Figure 8 presents our estimates of the treatment effect of loss framing on the distribution of final sales achieved.

Turning attention to the first panel of the first column, presenting estimates for Model Group A, we see clear evidence of the tell-tale signs of a response to loss framing. When loss aversion was induced, dealers pursued selling strategies that left them substantially less likely to face a small loss. Some of this missing mass appears to have shifted to small gains, consistent with some degree of a desirable increase in sales efforts. However, some of this missing mass appears to have shifted to comparatively *low* sales amounts,

Figure 8. : The Impact of Loss Framing on the Distribution of Sales



Notes: This figure shows difference in kernel density differences estimates of the impact of loss framing on the distribution of sales achieved. The first column contains estimates derived from each model group. The second column contains estimates derived from Model Group B, conditioning on whether that months sales of Model Group A were comparatively high (over 50% of the 110% target) or low (under 50% of the 110% target). The x-axes show the amount of sales relative to the target that needed to be achieved to avoid clawback. Standard errors are generated by conducting 1,000 bootstrap iterations, resampling by BAC. Kernel: Epanechnikov; Bandwidth: 10%.

rationalizable only by a demoralizing effect of facing a loss or by active substitution of potential sales opportunities into different model groups or different time windows. Overall, however, it appears the region of missing mass was split approximately evenly between increases and decreases in sales, consistent with the near-zero net effect of the intervention estimated in the standard difference-in-differences analysis.

Turning attention to the second panel of the first column, a more worrying picture emerges. Consistent with our regression results, we see clear evidence of effective cannibalization of Model Group B. Mass is shifted from comparatively higher to comparatively lower amounts of sales. While workers still did try to sell near-exactly the target amount (as seen by the spike in mass near the target value), possibilities to sell in excess of that amount appear to have been systematically spurned. Since loss framing induces incentives to reduce expected exposure to losses, and since larger losses occur by not attaining the Model Group A target, such findings arise naturally from the model presented in Section II.

The second column provides further support of the claim that Model Group B's shift to lower performance is partially driven by diversion of sales opportunities to Model Group A. In these panels, we recreate our DKDD estimates conditioning on that month's sales performance for Model Group A. In the top panel, we see that in months with comparatively high Model Group A sales (defined as being over 50% of their target), a picture very similar to the one just considered is seen. We estimate higher probability mass over the region of sales values that do not yield a bonus and missing mass associated with values just past the bonus threshold. In contrast, however, the lower panel shows that in months with comparatively low Model Group A sales—that is, months in which marginal sales of Model Group A are unlikely to yield a bonus—a more encouraging pattern is seen. In this case, dealers are less likely to end the month having not achieved a bonus, and are more likely to end the month in the near vicinity of the bonus target. These results are consistent with loss aversion motivating additional effort to complete Model Group B sales *as long as there was not competing motivation to attain the Model Group A target*.

F. Selection Into Loss-Framed Incentives

In this paper, we have focused on the negative effects of loss framing *conditional on participation in our experiment*. While prior research has assumed that such effects should be positive, several papers have emphasized the potential for negative effects arising by selecting who participates (Imas, Sadoff and Samek, 2016; de Quidt, 2017). To the extent that treatment assignment was random conditional on participation, these concerns should not affect our estimated treatment effects. However, for comparability to prior research and to completely examine existing accounts of the negative effects of loss framing, we now examine the predictors of selection into our experiment.

Despite the manufacturer's initial belief that this program would be highly desirable to dealers (due to providing early cash flow), comparatively few dealers opted into our experiment: 294 dealers opted in, 335 actively opted out, and 598 opted out through non-response. This low rate of opt-in could potentially be interpreted as prima facie evidence that dealers did in fact anticipate loss aversion, and thus avoided a situation that might induce it. Note, however, that several alternative explanations of the low participation rate are present. First, dealers may have failed to participate purely because they did not know of this opportunity. However, given the manner in which the opportunity was advertised, we believe this was unlikely to be a major factor.[24] Second, dealers may have

[24]Multiple emails were sent to the incentive program's contact at each dealer. CarCo regional dealer representatives indicated that dealers almost certainly read at least one of the multiple invitation emails. Since the incentive program is

feared that they would be unable to make the necessary clawback payments if targets were not met. In practice, several fail-safes ensured that dealers would face relatively few severe negative consequences if this situation arose, but lack of knowledge of those fail-safes or remaining concerns could possibly drive behavior.[25] Third, dealers may have been averse to accepting the administrative costs that a change in accounting practices would require.[26] In sum, while the low rate of participation is consistent with the concern that loss-averse agents will select out of loss framed contracts, we cannot firmly establish that this is this is the mechanism driving nonparticipation.

To examine how selection into the experiment influences the composition of our sample, we began by testing for differences between the three groups of dealers (opt-in, opt-out, and non-respondent) across the variables available in CarCo's internal data.

To begin, we emphasize that we find no statistically distinguishable differences in treatment group assignment across the three participation groups ($\chi^2 = 3.00, p = 0.223$). Treatment was assigned to 48.3% of participants, 54.0% of opt-outs, and 48.6% of non-respondents.

Among CarCo's internal data, we see evidence of selection on observables across several dimensions. First, as emphasized in the discussion of our difference-in-differences results, pre-period sales volume is notably associated with treatment assignment. Although participants and opt-out dealers had statistically indistinguishable monthly sales volume (48.8 vs. 43.3, $p = 0.17$), non-respondents had substantially fewer sales (31.1, $p < 0.01$).[27] Second, and consistent with these volume differences, there are significant cross-group differences in whether or not the dealers carry both model groups ($\chi^2 = 10.50, p = 0.005$). Non-respondents (86%) are less likely to do so than both participants (90%) and opt-out dealers (93%). Third, we find substantial differences across regions in participation rates, and particularly in the rate of opt-outs through non-responses ($\chi^2 = 101.88$, $p = 0.000$). Although we can only speculate why this is the case, a likely possibility is that CarCo's regional sales representatives used different strategies to encourage participation after the standardized initial email solicitations. Despite these regional differences in participation, there are no differences in treatment group assignment by region ($\chi^2 = 1.25, p = 0.870$).

These differences reject the hypothesis that dealers' willingness to be exposed to loss-framing through participation in our experiment is completely random conditional on observables. While the random assignment of treatment preserves the validity of estimated treatment effects within this group, we sought to collect more dealer-level data to help us better understand the features of the dealers included in our study.

We generate our dealer-level data by combining CarCo's sales and incentive records

crucial for their sales profits, program communications are high priority. In addition, CarCo's regional dealer representatives followed up with non-respondents to encourage participation. Based on these considerations, we believe effectively all non-respondents knowingly defaulted to non-participation.

[25]Based on the researcher's concerns about the negative consequences of failure to repay, CarCo agreed to automatically unenroll any dealer failing repayment. Among participating dealers, this condition was never triggered.

[26]Regional dealer representatives informally indicated that many of the dealers claimed that their lack of participation was due to avoidance of what they viewed as an accounting "headache," combined with little need for the early cash flow.

[27]These averages do not equal the sum of model-group average sales in Table 1 because some dealers carry only one model group.

Table 3—: Group Means by Participation Status for Dealers with NETS Data

	Selection Group			Group Difference Test
	Participants	Opt-Out	Non-Respondents	
Year Founded	1976.49	1976.69	1974.24	$F = 0.99$
	(27.16)	(27.81)	(27.37)	[0.371]
Employees	55.13	50.95	51.93	$F = 0.72$
	(46.70)	(45.35)	(40.25)	[0.488]
2015 DUNS Score	2.47	2.41	2.41	$F = 0.84$
	(0.60)	(0.52)	(0.56)	[0.431]
Has DUNS Score	0.70	0.62	0.66	$\chi^2 = 2.80$
	(0.46)	(0.49)	(0.47)	[0.246]
2015 Min Paydex	73.16	74.20	73.19	$F = 1.24$
	(8.73)	(6.35)	(9.58)	[0.291]
Has Paydex Score	0.81	0.75	0.80	$\chi^2 = 2.09$
	(0.39)	(0.43)	(0.40)	[0.351]
Publicly-Held	0.05	0.25	0.24	$\chi^2 = 3.78$
	(0.23)	(0.44)	(0.43)	[0.151]
Part of Group	0.18	0.23	0.26	$\chi^2 = 5.73$
	(0.39)	(0.42)	(0.44)	[0.057]

Notes: Summary statistics are presented for three groups: those who chose participation, those who opted-out, and those who did not respond and therefore were non-participants by default. Means with standard deviations in parentheses for each group. F-statistics from ANOVA are presented for continuous variables. Chi-squared tests are for dichotomous variables. P-values are in brackets. Each line only includes those dealers with a populated NETS field.

with supplemental data from the National Establishment Time Series (NETS). NETS aggregates data on most establishments in the United States (Barnatchez, Crane and Decker, 2017), including car dealers, providing data on ownership structure, employment, and financial strength data from Dun & Bradstreet.[28] We successfully matched 877 non-participants and 276 participants via phone, address, name, and ownership data. We then examined differences in sales, ownership, age, employment, and financial health.[29]

Table 3 presents group means and statistical tests of differences across the three groups for the 1,153 dealers with NETS data. We see few observable differences between the three groups. In particular, no statistical differences are detected in the age or size of the firm, in the Dun & Bradstreet measures of financial health, in the Paydex measures of reliable bill payment, or in publicly-held status. One potential difference is seen in the "Part of Group" variable. Participants are more likely to be stand-alone dealers (18%) than opt-out (23%) or non-respondents (26%). These differences could arise if dealers that were part of larger organizations were less likely to have autonomy to change accounting systems or rules.

In summary, we see relatively modest evidence of observable selection into participation in our experiment. We acknowledge that we cannot rule out perhaps interesting unobservable differences in dealers by participation status, but we note these would be very unlikely to bias our treatment estimates given our random assignment procedure.

G. Robustness Considerations and Additional Results

In this section we provide additional evidence to support our empirical results and address identification concerns.

Robustness to Outliers

As previously discussed in Section III.C, notable differences exist in the average pre-intervention sales between the treatment and control groups. As we argued in our discussion of Figure 3, this difference can be attributed to the random assignment of an outlier DMA. A single DMA contained five of the six largest participating dealers, each with average monthly sales several multiples of the average of other dealers. Due to the natural concern that these influential outliers meaningfully drive results, we reestimate our difference in differences model excluding this DMA. With its exclusion, average sales in the pre-period are statistically indistinguishable ($p = 0.451$), with means of 43.5 and 45.5 for the treatment and control group, respectively. Regression results are presented in Table 4. Average monthly sales decrease by 1.6 vehicles per month—4% relative to pre-intervention averages—with nearly all losses coming from the Model Group B. The smaller coefficient is largely reflected by the smaller nominal increase in sales coming from the smaller average sales in a control group without most of the largest dealers.

[28]CarCo did not share internal data on dealership financial or ownership structure with us for legal reasons.

[29]The NETS data fields have varying levels of completeness. Employment data, for example, is complete for approximately 90% of the matched dealers. Dun & Bradstreet data, however, is complete for only about two-thirds.

Table 4—: Difference-in-Difference Estimates Excluding the Outlier DMA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment Period X Assigned Prepay	-1.6	-1.6	-1.6	-0.4	-0.4	-0.4	-1.4	-1.4	-1.4
	(1.05)	(1.06)	(1.13)	(0.84)	(0.84)	(0.90)	(0.68)	(0.69)	(0.73)
Treatment Period	2.5			2.3			0.4		
	(0.69)			(0.64)			(0.53)		
Assigned Prepay	-2.0	-2.0		-4.7	-4.7		0.9	0.9	
	(6.04)	(6.05)		(4.59)	(4.59)		(2.33)	(2.33)	
Constant	45.5	37.4	36.4	34.2	28.9	26.5	14.3	11.1	11.5
	(4.38)	(4.18)	(0.94)	(3.74)	(3.53)	(0.73)	(1.44)	(1.39)	(0.47)
Model Group	Pooled	Pooled	Pooled	A	A	A	B	B	B
Month Fixed Effects		X	X		X	X		X	X
Dealer Fixed Effects			X			X			X
Observations	2296	2296	2296	2208	2208	2208	2160	2160	2160
R^2	0.001	0.009	0.957	0.005	0.012	0.954	0.001	0.014	0.890

Notes: This table presents difference-in-differences estimates of the effect of the prepayment intervention on total vehicles sold for the sample excluding the outlier DMA in control group that included five of the six largest dealers by volume. Two other dealers were dropped from this DMA, resulting in 287 remaining. See Figure 3 for details. The dependent variable is the month-specific number of vehicles sold at the dealer level. The first three columns present analysis pooling the two model groups together, whereas the next two groups of three columns present separate analysis of two model groups. Standard errors are clustered at the DMA level.

Interview and Survey Evidence

In this paper, we have focused on the observed reaction to the treatment using objective sales data. As a complement to this evidence, we conducted informal interviews and formal surveying of participants after the experiment. Fourteen dealers were interviewed, and sixty-two dealers (21%) responded to the online survey. Results are reported in Appendix Section A.3. Overall, the interviews and surveys suggested reasonably universal perceived importance of the incentive program, but more heterogeneous reactions to the loss framing treatment. Some dealers indicated perceptions that were clearly in line with loss-averse evaluation, while others provided assessments of the program more in line with neoclassical considerations. The interviews also indicated both awareness of, and efforts directed towards, the need to optimize efforts across the two model groups, consistent with the exacerbation of multitasking issues that we have discussed.

Examination of Treatment Effect in Second Treatment Window

As described in Section III.B, an "equal treatment" requirement led us to design our intervention so that all experimental participants faced four months in treatment condition and four months in the control condition. We have focused our attention on the estimated impact of treatment in the first four-month treatment window, prior to conditions being flipped. Conceptually, we believe that analysis of treatment effects in the first window have the cleanest interpretation. When examining the second window, the prior treatment of the control group is expected to affect results, leading to a confounding of a difference-in-differences design. For completeness, we attempt to estimate treatment effects in this window using synthetic control methods (Abadie, Diamond and Hainmueller, 2010; Robbins, Saunders and Kilmer, 2017).

The aim of our synthetic control analysis is to compare the sales among dealers receiving treatment in the second window to that of comparable dealers who did not participate in the experiment. We constructed the synthetic control group by matching on dealer size (in units sold), dual-dealer status, and the pre-trends in sales that were used in our algorithm for random assignment. The estimated model, represented in Figure A.7, shows a treatment effect of loss framing of -12% ($p = 0.019$)[30]—substantially larger than our estimate of -5% in our primary analysis. For comparison, Figure A.8 applies this method to compare the dealers treated in the first window to a synthetic control formed from control-group dealers, This analysis yields an estimated treatment effect of -4% ($p = 0.009$)—generally comparable with our difference-in-differences estimate.

In short, though we believe our difference-in-differences estimates are conceptually preferable to these synthetic control estimates (as they do not require comparisons to a group that has endogenously opted out of treatment), we note that both methods provide evidence of a negative effect of loss framing.

[30]All estimates are derived using the R package microsynth, with p-values generated through the permutation method with 10,000 permutations applied.

IV. 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 organizational policies optimized for behavioral agents. While the potential for using loss framing in sales incentives has received a great deal of attention, formal field evaluation of such programs is currently limited to a small number of cases (e.g., Hossain and List, 2012; Fryer Jr et al., 2012; Chung and Narayandas, 2017). 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. In an unusually large field experiment, in which $66 million in bonus payments was subject to random assignment of loss framing, we have demonstrated that organizations are not wrong to suspect that loss aversion might influence the incentives' efficacy. We have also demonstrated, however, that organizations would be wrong to expect this to always be a desirable outcome. Our estimates imply that loss framing caused a loss of $45 million in revenue generated by the treatment group during the 4-month treatment window. Furthermore, we have shown that such losses are not unexpected: in environments where agents can pursue costly risk mitigation strategies, or when agents must multitask between multiple incentivized activities, inducing loss aversion is expected to exacerbate reasonably classic problems in principal-agent relationships.

These findings serve as a demonstration of two tenants of field experimentation. First, they clearly illustrate the the fundamental need for experimentation prior to the deployment of unvetted policies. Recall that, prior to the researchers' intervention, the car manufacturer studied in this paper had considered deploying the clawback policy nationwide. If this had occured, the problems that we highlight would not have been detected. This policy was rolled out during a period of growth in car sales, and thus a simple pre/post comparison would have suggested that the deployment of the policy was associated with increased sales (see Figure 6). That analysis misleadingly paints a brighter picture than ours: our difference-in-differences estimates imply that system-wide adoption of loss framing would have resulted in an annual revenue loss of over one billion dollars.

Second, 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. The initial experimental findings of Hossain and List (2012) served as a motivating force for the consideration of loss framing. However, in papers that followed, some researchers faced difficulties when attempting to generate similar positive treatment effects (see, e.g., de Quidt et al., 2017; DellaVigna and Pope, 2018).[31] Through a formalized reexamination of the underlying conceptual premise of this literature, we are able to make better sense of why these seemingly inconsistent results arise, and are able to make more precise predictions about the situations in which loss framing may be ineffective or counterproductive.

Although we have emphasized that our results suggest that the presence of loss aver-

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

sion 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, demonstrations of loss aversion in fundamentally economic field environments have begun to proliferate. Loss aversion has now been shown to influence job search behavior (DellaVigna et al., 2017), labor supply (Camerer et al., 1997; Fehr and Goette, 2007; Farber, 2008; Crawford and Meng, 2011; Thakral and Tô, 2019), house prices (Genesove and Mayer, 2001), tax compliance (Engström et al., 2015; Rees-Jones, 2018), and more (for a thorough review, see O'Donoghue and Sprenger, 2018). In addition to bolstering these field demonstrations, we demonstrate that similar considerations apply when the agents of interest are firms rather than individuals. We additionally contribute to a growing body of evidence on how behavioral principles such as loss aversion can motivate undesirable "gaming" behaviors (Gubler, Larkin and Pierce, 2016; Rees-Jones, 2018).

Given our results, a natural question remains: if loss aversion is present but not always desirable, under what conditions will the induction of loss aversion be part of an optimal contract? Ultimately the answer to such questions will be governed by the structure imposed on the normative preferences of both the principal and agents, as well as the distribution of loss aversion among both parties. In our setting, we believe both preferences and biases are quite heterogeneous—perhaps intractably so for unobjectionable structural estimation. This has motivated our choice to present our theory with minimal structure imposed, which proves sufficient for qualitative positive statements but which is too unstructured for quantitative normative analysis. While our work does not formally answer these questions, we can speculate on the core principles guiding such analyses. The core problem present in principal-agent contracting is the divergence in preferences between the principal and the agent. This contracting problem is often easiest when the principal's and agent's preferences are most aligned. To the extent that we view principals as being unbiased and risk neutral, we believe that inducing loss aversion among agents will typically exacerbate the differences between their preferences, which in general will be undesirable. This suggests that, given control over both the contract provisions and loss framing, the optimal incentive program will often not involve loss framing's induction. When institutional features restrict the principal from offering an optimal contract (as is the case in our setting), the precise conditions governing the optimal implementation of loss framing remain ill explored. We believe the exploration of such questions could meaningfully guide future organizational policies like the one we have studied.

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.

Armantier, Olivier, and Amadou Boly. 2015. "Framing of incentives and effort provision." *International Economic Review*, 56(3): 917–938.

Athey, Susan, and Guido W Imbens. 2017. "The econometrics of randomized experiments." In *Handbook of Economic Field Experiments*. Vol. 1, 73–140. Elsevier.

Baker, George P, Michael C Jensen, and Kevin J Murphy. 1988. "Compensation and incentives: Practice vs. theory." *The Journal of Finance*, 43(3): 593–616.

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.

Barberis, Nicholas C. 2013. "Thirty years of prospect theory in economics: A review and assessment." *Journal of Economic Perspectives*, 27(1): 173–96.

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker. 2017. "An assessment of the National Establishment Time Series (NETS) database." *Opportunity Inclusive Growth Institute: System Working Paper*, 17–29.

Bennett, Victor Manuel. 2013. "Organization and bargaining: Sales process choice at auto dealerships." *Management Science*, 59(9): 2003–2018.

Benson, Alan. 2015. "Do agents game their agents' behavior? Evidence from sales managers." *Journal of Labor Economics*, 33(4): 863–890.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and explaining management practices across firms and countries." *The Quarterly Journal of Economics*, 122(4): 1351–1408.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012. "The organization of firms across countries." *The Quarterly Journal of Economics*, 127(4): 1663–1705.

Brooks, Richard R. W., Alexander Stremitzer, and Stephan Tontrup. 2012. "Framing contracts: Why loss framing increases effort." *Journal of Institutional and Theoretical Economics*, 168(1): 62–82.

Brooks, Richard R. W., Alexander Stremitzer, and Stephan Tontrup. 2017. "Stretch it but don't break it: The hidden cost of contract framing." *The Journal of Legal Studies*, 46(2): 399–426.

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): 1253–1270.

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 R, Christopher R Knittel, and Florian Zettelmeyer. 2013. "Are consumers myopic? Evidence from new and used car purchases." *American Economic Review*, 103(1): 220–56.

Cachon, Gérard P, and Marcelo Olivares. 2010. "Drivers of finished-goods inventory in the US automobile industry." *Management Science*, 56(1): 202–216.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor supply of New York City cabdrivers: One day at a time." *The Quarterly Journal of Economics*, 112(2): 407–441.

Card, David, Stefano DellaVigna, and Ulrike Malmendier. 2011. "The role of theory in field experiments." *Journal of Economic Perspectives*, 25(3): 39–62.

Chung, Doug J, and Das Narayandas. 2017. "Incentives versus reciprocity: insights from a field experiment." *Journal of Marketing Research*, 54(4): 511–524.

Chung, Doug J, Das Narayandas, and Dongkyu Chang. 2019. "The effects of quota frequency: Sales performance and product focus." *Harvard Business School Working Paper No. 17-059*.

Chung, Doug J, Thomas Steenburgh, and K Sudhir. 2013. "Do bonuses enhance sales productivity? A dynamic structural analysis of bonus-based compensation plans." *Marketing Science*, 33(2): 165–187.

Courty, Pascal, and Gerald Marschke. 2004. "An empirical investigation of gaming responses to explicit performance incentives." *Journal of Labor Economics*, 22(1): 23–56.

Crawford, Vincent P, and Juanjuan Meng. 2011. "New York City cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income." *American Economic Review*, 101(5): 1912–32.

DellaVigna, Stefano, and Devin Pope. 2018. "What motivates effort? Evidence and expert forecasts." *The Review of Economic Studies*, 85(2): 1029–1069.

DellaVigna, Stefano, Attila Lindner, Balázs Reizer, and Johannes F. Schmieder. 2017. "Reference-dependent job search: Evidence from Hungary." *The Quarterly Journal of Economics*, 132(4): 1969–2018.

de Quidt, Jonathan. 2017. "Your loss is my gain: a recruitment experiment with framed incentives." *Journal of the European Economic Association*, 16(2): 522–559.

de Quidt, Jonathan, Francesco Fallucchi, Felix Kölle, Daniele Nosenzo, and Simone Quercia. 2017. "Bonus versus penalty: How robust are the effects of contract framing?" *Journal of the Economic Science Association*, 3(2): 174–182.

Dittmann, Ingolf, Ernst Maug, and Oliver Spalt. 2010. "Sticks or carrots? Optimal CEO compensation when managers are loss averse." *The Journal of Finance*, 65(6): 2015–2050.

Ederer, Florian, Richard Holden, and Margaret Meyer. 2018. "Gaming and strategic opacity in incentive provision." *The RAND Journal of Economics*, 49(4): 819–854.

Engström, Per, Katarina Nordblom, Henry Ohlsson, and Annika Persson. 2015. "Tax Compliance and Loss Aversion." *American Economic Journal: Economic Policy*, 7(4): 132–164.

Farber, Henry S. 2008. "Reference-dependent preferences and labor supply: The case of New York City taxi drivers." *American Economic Review*, 98(3): 1069–82.

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

Fryer Jr, Roland G, Steven D Levitt, John List, and Sally Sadoff. 2012. "Enhancing the efficacy of teacher incentives through loss aversion: A field experiment." NBER Working Paper No. 18237.

Genesove, David, and Christopher Mayer. 2001. "Loss aversion and seller behavior: Evidence from the housing market." *The Quarterly Journal of Economics*, 116(4): 1233–1260.

Goldberger, A.S. 1968. *Topics in Regression Analysis.* Macmillan.

Gubler, Timothy, Ian Larkin, and Lamar Pierce. 2016. "Motivational spillovers from awards: Crowding out in a multitasking environment." *Organization Science*, 27(2): 286–303.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." *Journal of Law, Economics, Organization*, 7: 24–52.

Hong, Fuhai, Tanjim Hossain, and John A. List. 2015. "Framing manipulations in contests: A natural field experiment." *Journal of Economic Behavior Organization*, 118: 372 – 382. Economic Experiments in Developing Countries.

Hossain, Tanjim, and John A List. 2012. "The behavioralist visits the factory: Increasing productivity using simple framing manipulations." *Management Science*, 58(12): 2151–2167.

Imas, Alex, Sally Sadoff, and Anya Samek. 2016. "Do people anticipate loss aversion?" *Management Science*, 63(5): 1271–1284.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics*, 8(1): 435–464.

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.

Lazear, Edward P. 1991. "Labor economics and the psychology of organizations." *Journal of Economic Perspectives*, 5(2): 89–110.

List, John A. 2002. "Testing neoclassical competitive market theory in the field." *Proceedings of the National Academy of Sciences*, 99(24): 15827–15830.

List, John A. 2003. "Does market experience eliminate market anomalies?" *The Quarterly Journal of Economics*, 118(1): 41–71.

List, John A. 2004*a*. "Neoclassical theory versus prospect theory: Evidence from the marketplace." *Econometrica*, 72(2): 615–625.

List, John A. 2004*b*. "Testing neoclassical competitive theory in multilateral decentralized markets." *Journal of Political Economy*, 112(5): 1131–1156.

Ljungqvist, Alexander, and William J Wilhelm Jr. 2005. "Does prospect theory explain IPO market behavior?" *The Journal of Finance*, 60(4): 1759–1790.

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): 307–326.

Lu, Bo, Robert Greevy, Xinyi Xu, and Cole Beck. 2011. "Optimal nonbipartite matching and its statistical applications." *The American Statistician*, 65(1): 21–30.

Luft, Joan. 1994. "Bonus and penalty incentives contract choice by employees." *Journal of Accounting and Economics*, 18(2): 181–206.

Malmendier, Ulrike. 2018. "Behavioral corporate finance." In *Handbook of Behavioral Economics*. Vol. 1, , ed. Douglas Bernheim, Stefano Dellavigna and David Laibson. Elsevier.

Manski, Charles F. 1986. "Analog estimation of econometric models." In *Handbook of Econometrics*. Vol. 4, , ed. R. F. Engle and D. McFadden, Chapter 43, 2559–2582. 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): 211–257.

Murry, Charles. 2018. "The effect of retail competition on relationship-specific investments: evidence from new car advertising." *International Journal of Industrial Organization*, 59: 253–281.

Murry, Charles, and Henry S Schneider. 2016. "The economics of retail markets for new and used cars." *Handbook on the Economics of Retailing and Distribution*, 343.

O'Donoghue, Ted, and Charles Sprenger. 2018. "Reference-dependent preferences." In *Handbook of Behavioral Economics - Foundations and Applications 1*. Vol. 1, , ed. B. Douglas Bernheim, Stefano DellaVigna and David Laibson, 1 – 77. North-Holland.

Olivares, Marcelo, and Gérard P Cachon. 2009. "Competing retailers and inventory: An empirical investigation of General Motors' dealerships in isolated US markets." *Management Science*, 55(9): 1586–1604.

Oyer, Paul. 1998. "Fiscal year ends and nonlinear incentive contracts: The effect on business seasonality." *The Quarterly Journal of Economics*, 113(1): 149–185.

Pierce, Lamar. 2012. "Organizational structure and the limits of knowledge sharing: Incentive conflict and agency in car leasing." *Management Science*, 58(6): 1106–1121.

Rees-Jones, Alex. 2018. "Quantifying loss-averse tax manipulation." *The Review of Economic Studies*, 85(2): 1251–1278.

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): 109–126.

Roberts, Adrienne. 2018. "Car dealerships face conundrum: Get big or get out." *Wall Street Journal*.

Rubin, Jared, Anya Samek, and Roman M Sheremeta. 2018. "Loss aversion and the quantity–quality tradeoff." *Experimental Economics*, 21(2): 292–315.

Steenburgh, Thomas J. 2008. "Effort or timing: The effect of lump-sum bonuses." *Quantitative Marketing and Economics*, 6(3): 235.

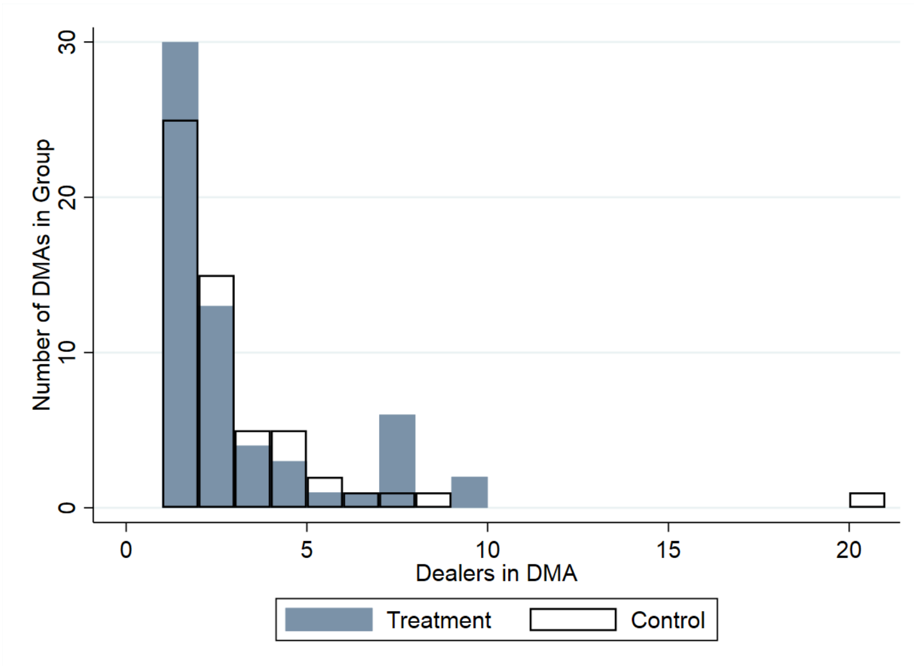
Thakral, Niel, and Lihn Tô. 2019. "Daily labor supply and adaptive reference points." *American Economic Review*, Forthcoming.

Appendix

1. Group Assignment

Figure A.1 shows the distribution of the number of participants in the DMAs assigned to each group. The outlier is the same control group DMA referenced in the discussion of Figure 3, containing the unusually large dealers.

Figure A.1. : Distribution of DMA Dealer Count By Condition



2. Additional Time Trends

Figures A.2 and A.3 present variants of Figure 6 constructed by model group. We plot average dealer unit sales by month for both the control and treatment group as well as each group's sales average across both the pre- and post-treatment periods. Figures A.4–A.6 compare trends for participants and non-participants.

Figure A.2. : Participant Time Trends for Model Group A

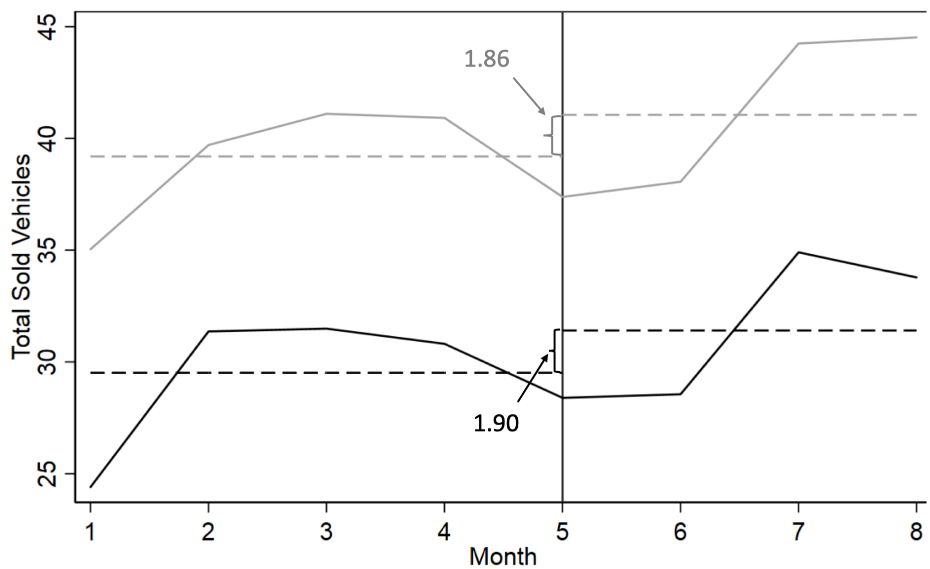


Figure A.3. : Participant Time Trends for Model Group B

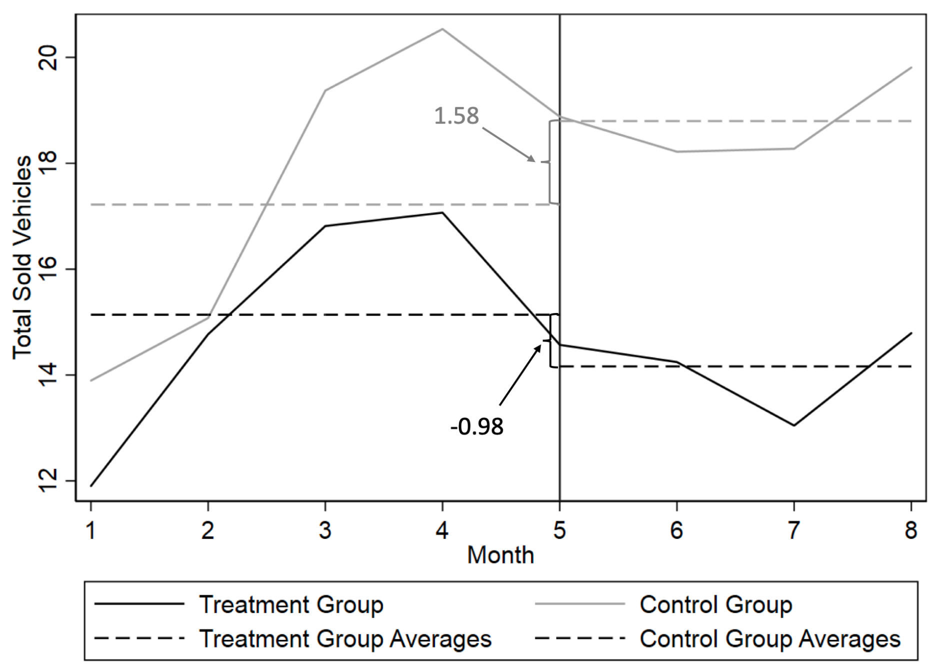


Figure A.4. : Non-Participant Sales Time Trends for Total Sales

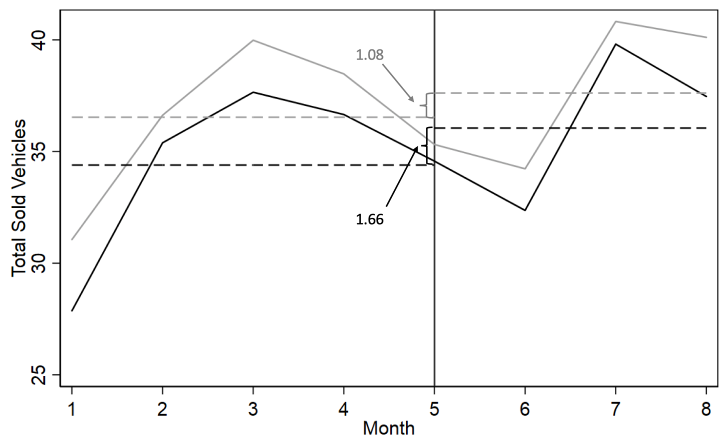


Figure A.5. : Non-Participant Time Trends for Model Group A

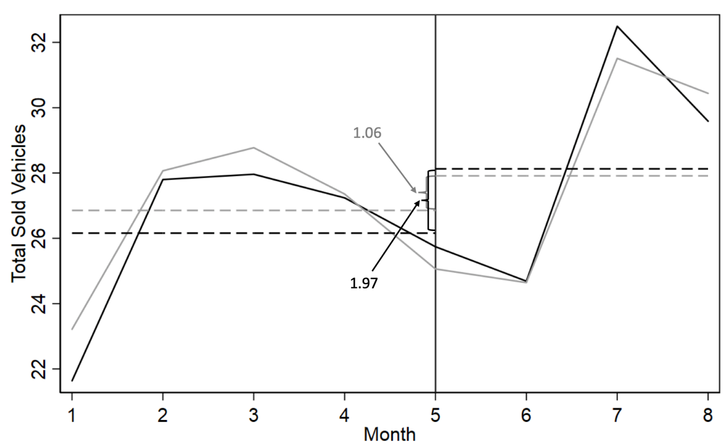
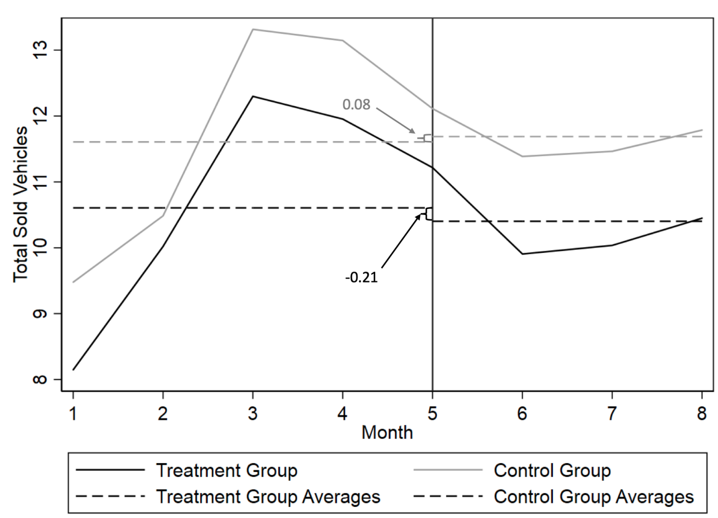


Figure A.6. : Non-Participant Time Trends for Model Group B



3. Interview and Survey Evidence

Following the completion of the experiment, we gathered information from the participating dealers on both their experiences and their actions resulting from the loss-framed incentive structure. We first conducted interviews with a subset of dealers, then launched an online survey. There were three goals for the interviews and surveys. First, we wanted to understand how dealers' self-assessed their reaction to the intervention—e.g., whether they felt that it induced loss aversion, or whether they viewed it as an interest-free loan. Second, we sought to understand the extent to which dealers might change effort and risk mitigation in response to monthly fluctuations in the probability of achieving the discrete sales target, and whether the loss-framing changed that behavior. Finally, we sought to understand if dealers adjusted their internal pay structure to managers and salespeople when receiving the monthly prepayments.

Interview Evidence

The interviews were conducted over the phone in May, 2018, with either the general manager or principal (owner) of 14 participating dealers. Dealer interviews were arranged by CarCo with a mix of targeted regions and experimental conditions. Participation was not random nor representative of the dealers as a whole. The interviews, which lasted 20 to 30 minutes, were relatively unstructured, but focused on five key questions: 1. "How do the [program] targets influence formal policies or managerial attention in a given month for both you or other dealers?" 2. "How do you think about the two separate brands and respective targets, both before and after achieving the targets?" 3. "When you received the prepayments, did this affect any of your policies or attention, and how was the cash used?" 4. "How do you imagine other dealers might deal with these prepayments, and do you see them as generating potential benefits or problems for these other dealers?" 5. "What do you see as the immediate strengths and weaknesses of the current setup, where 110% has such a large immediate payoff?"

The interviews revealed relatively consistent views on the role of the incentive program in dealer strategy and policy. Respondents universally emphasized that the incentive program represents the majority of their sales profits. One dealer noted that "you gotta get there. If you don't hit the 110 number, your operating profit is next to nothing." Another stated that "it's critical to reach those plateaus for achieving profitability." They also consistently noted how the program dominates their sales strategy. "It's an integral part of how we shape and process just about everything." Another dealer explained that "(the incentive plan) is running our business. It's the most important thing. Financially, it's crucial—55-65% of operating profit." Several dealers also noted that hitting the objectives early changed their behavior, saying "it allows us to start focusing on margin," while also noting that it makes them "relaxed." Dealers also explained that the incentive program is tightly tied to the incentives of the sales managers and salespeople.

Many indicated that they carefully track progression toward the targets throughout the month, adjusting resources and pricing based on the likelihood of hitting the target. One dealer reported that "if we don't hit a fast start, then we back off. If we're close, we use incentives to get there." This resource and attention allocation also applied across brands. A dealer noted that "if we're cruising on (Model Group A), we'll put attention to (Model Group B). If we get to the 20th of the month, we'll focus on the one closest to the 110% target." One dealer noted that "we think about (the brands) entirely separately, until we're close. Then I'll move people." Another explained that "mid-month, when it looks likely to hit one more than the other, we allocate resources accordingly." Another dealer noted the difference between the brands. "We make a lot more money off (Model Group A) so we focus there. We're not going to take a (Model Group A) buyer and put them in (Model Group B)."

Although dealers were reluctant to discuss attempts to move customers across months and brands, several indicated that they had some ability to do so, and many more explained that they "know there are dealers who play that game," or opaquely explained "if we're having a slow month, we think about next month." Some dealers also indicated the importance of having smooth earnings across months. "You can't have peaks and valleys in your sales year. You need to be consistent."

The interviewed dealers indicated a mix of responses to the prepayment condition. One mentioned that "the reason I signed up was for an interest-free loan," with some other dealers expressing similar sentiments. Another mentioned that "this is simply cash up front, and we put it in a separate account," explaining that "my sisters are very accounting oriented." Others, however noted the motivational effect it had on their business. One dealer noted that "nobody wanted to give it back." Another explained that "it's like giving my son $100—it's going to be spent. Your back's in the corner so you've got to produce." Another also mentioned that "nobody wanted to give it back," referring to it as "a little extra spice in the stew," and explaining that he "didn't want to have that conversation with the owner." One successful dealer explained that the prepayment did not impact their behavior during the experiment because "I knew I would hit my goals. I would be hesitant to do it this year, because I don't want it yanked back from me." Another stated that it "always felt like there was a lot more pressure. It changed the intensity." Two dealers also noted that the prepayment did help with cash flow.

Post-Experiment Survey

The post-experiment survey was conducted in July, 2018, and was distributed to the primary contacts at all 294 participating dealers. The survey asked about the use of the advanced funds as well as the respondent's perception about how the advanced funds changed behavior throughout the month. Follow-up emails were used to attempt to increase participation, resulting in sixty-two dealers (21%) completing the survey.

Only 13 dealers reported distributing advanced funds to employees, with all but two focusing this loss-framing on the general manager or sales manager. One respondent freely responded to this question by answering "ARE YOU CRAZY?" The majority of respondents (44) reported setting the funds aside until the end of the month. Twenty-two dealers reported that they agreed that the advanced funds increased pressure and motivation to meet program goals. Eighteen dealers also reported that they increased their emphasis on meeting program targets under the prepayment condition, while 16 indicated that the prepayment condition made them more closely track progress toward the monthly targets.

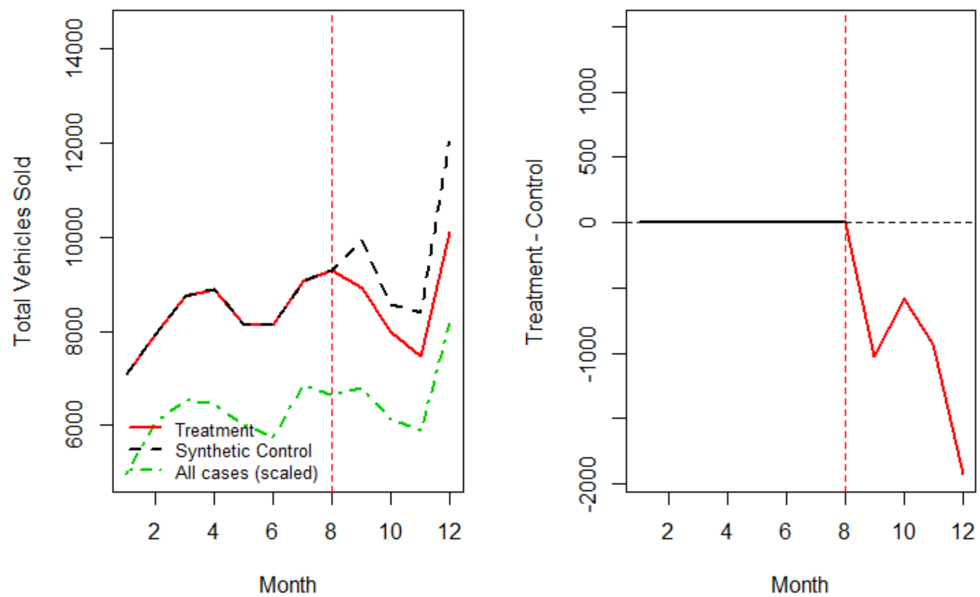
Fourteen dealers reported that the prepayment condition changed their approach to the two brands, with 11 reporting that it increased their focus on whichever brand was closest to the target.

Sixteen dealers believed that the prepayment system increased their sales, while one believed it slightly decreased them. Finally, 26 dealers would prefer advanced funds moving forward, while 24 would prefer end-of-the-month and 10 were indifferent.

Collectively, the interviews and survey indicate that for a select set of dealers, the prepayment condition changed behavior and mind-frame in many dealers while having little effect in others. This is consistent with there being a variety of management practices across car dealers resulting from both franchise law protections and other sources of heterogeneity. We draw several important conclusions from these self-reported data. First, the prepayment condition was salient enough to successfully treat some dealers with loss-framing. Second, this loss-framing was not universal, which deflates any average treatment effects estimated from the study. Third, the behavioral responses of the treated dealers are consistent with our model of increased effort and risk mitigation across months and brands.

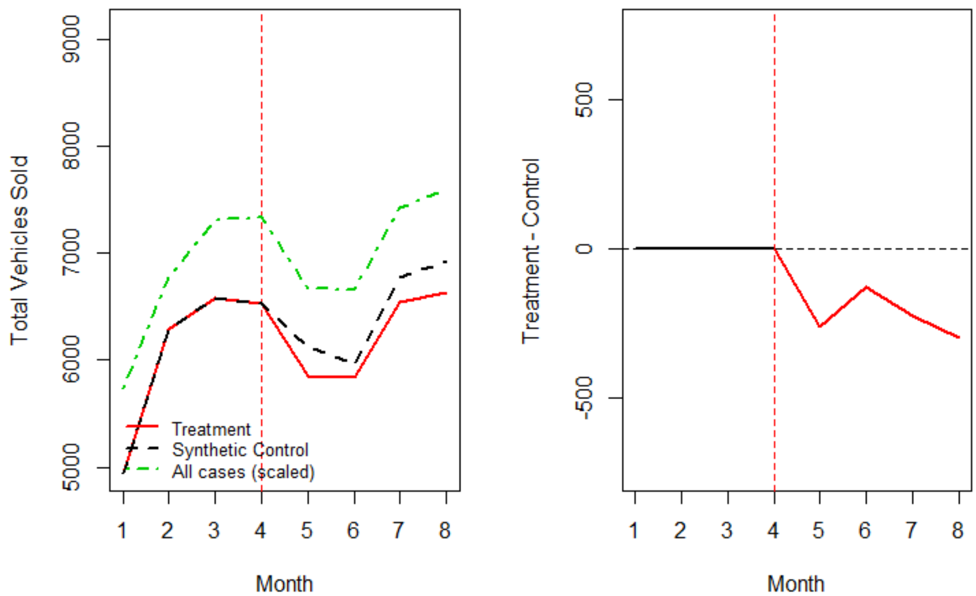
4. Synthetic Control Models

Figure A.7. : Synthetic Control Estimates of Treatment Effect in Second Treatment Window



Notes: This figure shows results from a synthetic control model, estimating the impact of the control group recieving the loss-framing treatment in months 9–12. The vertical axes represent the total monthly sales across the treatment and synthetic control groups. The synthetic control group is built from non-participants in the treatment group and matches on dealer size (in units sold), dual-dealer status, and pre-treatment sales trends. The model estimates a treatment of -12%.

Figure A.8. : Synthetic Control Treatment Estimates of Treatment Effect in First Treatment Window



Notes: This figure shows results from a synthetic control model of the treatment effect on the treatment group in months 5–8. The vertical axes represent the total monthly sales across the treatment and synthetic control groups. The synthetic control group is built from participants in the control group and matches on dealer size (in units sold), dual-dealer status, and pre-treatment sales trends. The model estimates a treatment of -4%.