Management Science

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

Do People Anticipate Loss Aversion?

Alex Imas, Sally Sadoff, Anya Samek

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–15
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

informs
http://dx.doi.org/10.1287/mnsc.2015.2402

Do People Anticipate Loss Aversion?

Alex Imas
Carnegie Mellon University, Pittsburgh, Pennsylvania 15213; and University of California, San Diego, La Jolla, California 92093, aimas@andrew.edu

Sally Sadoff
University of California, San Diego, La Jolla, California 92093, ssadoff@ucsd.edu

Anya Samek
University of Southern California, Los Angeles, California 90089, samek@usc.edu

There is growing interest in the use of loss contracts that offer performance incentives as up-front payments that employees can lose. Standard behavioral models predict a trade-off in the use of loss contracts: employees will work harder under loss contracts than under gain contracts, but, anticipating loss aversion, they will prefer gain contracts to loss contracts. In a series of experiments, we test these predictions by measuring performance and preferences for payoff-equivalent gain and loss contracts. We find that people indeed work harder under loss than gain contracts, as the theory predicts. Surprisingly, rather than a preference for the gain contract, we find that people actually prefer loss contracts. In exploring mechanisms for our results, we find suggestive evidence that people do anticipate loss aversion but select into loss contracts as a commitment device to improve performance, using one bias, loss aversion, to address another, dynamic inconsistency.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mnsc.2015.2402.

Keywords: economics; behavior and behavioral decision making; labor; utility preference; loss aversion; contracting; incentives; laboratory experiment
History: Received March 18, 2014; accepted June 25, 2015, by Teck-Hua Ho, behavioral economics. Published online in *Articles in Advance*.

1. Introduction

Attempts to take advantage of findings from behavioral economics have become increasingly popular in management and public policy (Madrian 2014). A prime example is the phenomenon of *loss aversion*, which predicts that gains and losses are evaluated relative to a reference point and that losses loom larger than gains.[1] An important behavioral prediction of loss aversion is that individuals first endowed with a payment will work harder to avoid losing it than to earn the same amount presented as a gain. Recent work has explored whether the design of incentive contracts can exploit this insight to increase effort and performance in the workplace (Hossain and List 2012, Fryer et al. 2012). These studies find that presenting incentives in the form of loss contracts (i.e., bonuses workers could potentially lose) increases productivity relative to payoff-equivalent gain contracts where the same bonuses are presented as gains. Yet loss-framed contracts tend to be rare in practice (Baker et al. 1988, Lazear 1991).

A natural criticism of the economic significance of loss contracts is that if people are averse to losses, they will have a preference for gain contracts. In turn, firms may have to pay a premium for workers to accept loss contracts, which could outweigh their productivity benefits and make loss contracts inefficient. As we discuss further in the next section, a standard behavioral model of reference-dependent preferences makes two central predictions: first, to avoid potential losses, individuals will exert more effort under a loss contract than a gain contract; and second, anticipating this loss aversion, individuals will have a strict preference for the gain contract.[2] The intuition is that because losses are painful, people will work harder to avoid losing a bonus than they would to receive

[1] Kahneman and Tversky's (1979) seminal work on prospect theory introduced and formalized loss aversion. Since then, loss aversion has been used to explain a variety of behavioral anomalies including the endowment effect (Kahneman et al. 1990) and status quo bias (Samuelson and Zeckhauser 1988). For more recent reviews of applications of reference-dependent preferences, see Camerer et al. (2004), DellaVigna (2009), Barberis (2013), and Ericson and Fuster (2014).

[2] Models using the status quo as the reference point (e.g., Thaler and Johnson 1990) predict that individuals will work harder under loss contracts conditional on the endowment being incorporated as the status quo. If the expectation is taken as the reference point (e.g., Kőszegi and Rabin 2006), then no difference between the contracts should be observed either in effort or in contract preference. As such, in our context, a *standard behavioral model* refers to models assuming the status quo as the reference point.

the same bonus offered as a gain. At the same time, working harder than is otherwise optimal in order to avoid losses lowers expected utility relative to working under a gain contract. If employees anticipate this difference in expected utility, they will prefer to work under gain contracts. As discussed above, several recent studies have provided evidence for the first prediction—that loss contracts increase workplace productivity (Brooks et al. 2012, Hossain and List 2012, Fryer et al. 2012).[3] However, less work has been done to directly investigate the second prediction, that individuals prefer gain contracts to loss contracts. Understanding individuals' preferences over contracts is critical in determining optimal contract design from the perspective of the firm or manager.

In this paper, we present results from a series of incentivized laboratory experiments that measure both the impact of gain and loss contracts on productivity and ex ante preferences for selecting into payoff-equivalent gain and loss contracts for the same task. We first test whether participants exert greater effort when payoff-equivalent incentives are presented as a loss contract rather than as a gain contract. We then conduct a second experiment to examine whether people anticipate the differential effect of the loss contract in line with the standard behavioral model—that is, whether they prefer gain rather than loss contracts. To do this, we use the same task and compare participants' willingness to pay (WTP) to work under the loss contract versus their WTP to work under the gain contract. Finally, we investigate the correlation of performance and selection into contracts with a separately elicited loss aversion parameter.[4]

In line with the standard model of prospect theory, we find that individuals assigned to the loss contract work harder than those assigned to the gain contract. However, we do not find support for the theoretical prediction that people prefer the gain contract to the loss contract. Surprisingly, people are willing to pay more to enter the loss contract than to enter the payoff-equivalent gain contract for the same task.

To shed light on the mechanisms behind our results, we examine the relationship between performance, contract preferences, and the separately elicited individual-level loss aversion parameter. As discussed above, in the standard behavioral model, losses are more painful for people who are more loss averse, which leads to two predictions: greater loss aversion results in greater effort when assigned to a loss contract, and this greater effort corresponds to lower expected welfare ex ante. In turn, if people anticipate their degree of loss aversion when selecting into a contract, more loss-averse people will have a greater ex ante preference against loss contracts.

We find suggestive evidence to support the first prediction: performance under loss contracts increases with the degree of loss aversion, whereas loss aversion has no relationship with performance under the gain contract. However, we find no evidence for the second prediction; rather, willingness to pay for loss contracts seems to increase rather than decrease with the degree of loss aversion. As discussed in Section 6, our findings potentially provide support for a behavioral model that incorporates dynamic inconsistency in preferences. In this framework, loss contracts can be viewed as a commitment device that sophisticated workers select so as to improve their performance and increase their expected earnings, using one bias, loss aversion, to address another, dynamic inconsistency.

The few other studies examining contract preferences in this context have found mixed evidence on selection into loss and gain contracts. Luft (1994) examines entry into gain and loss contracts (with an emphasis on calling them "bonus" and "penalty" contracts) and finds that participants are more likely to enter into contracts framed as gains. In work conducted concurrently to our own, De Quidt (2014) examines the effect of descriptive framing on entry into contracts offered on the crowd-sourcing platform Amazon Mechanical Turk. In line with our empirical results, De Quidt finds that entry rates are higher in the loss-framed contract than in the gain-framed contract. Together, our two studies demonstrate a preference for loss contracts in different settings using different methods to vary the contract type—De Quidt varies the descriptive framing of incentives as either bonuses (in gain contracts) or penalties (in loss contracts), whereas we vary the timing of receiving incentives, with potential rewards earned after completion of the task under gain contracts and endowed prior to the task under loss contracts.

To our knowledge, ours is the only paper studying preferences between contracts that separately examines the effect of contract type on productivity. This allows us to identify the role of loss aversion in both contract preference and productivity without issues of selection effects, which our results suggest could be a potentially serious confound. As far as we know, ours is also the first study to investigate how performance and contract preferences vary by individuals'

[3] Related field studies have compared the impact of incentives framed as gains with incentives framed as losses in other contexts, including incentives for student performance (Levitt et al. 2016) and child food choice (List and Samek 2014).

[4] Prior to conducting the experiments, we ran a series of pilots with a smaller sample, which are summarized in Appendix B. The results of the pilot experiments are consistent with the findings of the experiments reported here.

degree of loss aversion, allowing us to identify potential mechanisms driving the results.[5]

The remainder of the paper is organized as follows. In Section 2, we discuss four theoretical predictions of a simple behavioral model of reference-dependent preferences and loss aversion, which motivate our experimental design and analysis. Section 3 presents the design and results of our first experiment, which separately identifies the effect of contract type on productivity. Section 4 presents our second experiment investigating preferences over gain versus loss contracts. Section 5 describes our elicitation of an individual-level loss aversion parameter and subsequent tests. Section 6 discusses implications of the results, and Section 7 concludes.

2. Theoretical Predictions

We consider a standard behavioral model in which people derive additively separable utility from consumption net of costs, *net consumption utility*, as in the standard framework; they also derive *gain-loss utility* relative to a reference point, as in the prospect theory model (Kahneman and Tversky 1979). We assume the reference point is determined by the status quo (e.g., Thaler and Johnson 1990). In our context, a performance incentive can be offered as a potential gain, in which the status quo is *not having* the incentive, or as a potential loss, in which the status quo is *having* the incentive. In gain contracts, people work to increase the probability they will receive the incentive. In loss contracts, people work to increase the probability they will avoid losing the incentive. There are two critical assumptions of this model. First, people experience utility relative to a reference point, deriving positive gain-loss utility from gains and negative gain-loss utility from losses (the utility from remaining at the reference point is normalized to zero).

Second, losses loom larger than gains such that the negative gain-loss utility from a loss of x relative to the reference point is larger in absolute value than the positive gain-loss utility from a gain of x. This greater sensitivity to losses, *loss aversion*, leads to our first two predictions for relative performance under gain versus loss contracts. See Appendix A for proofs of all results.

[5] De Quidt (2014) includes an unincentivized survey measure of risk preferences using hypothetical lotteries. As discussed in Section 5, unlike our incentivized measure, the survey measure does not separately identify loss aversion from utility curvature in the gain and loss domains, and hence it cannot be used as a clean test of the theory. De Quidt also tests for selection on observables (including the survey measure) and does not find significant differences across contracts; the experimental design precludes ruling out selection on unobservables.

Prediction 1. *If people are loss averse, performance will be higher under a loss contract than under a gain contract.*

Under both a gain and a loss contract, individuals choose optimal effort to maximize expected utility—i.e., the effort level at which the marginal benefits from increasing the likelihood of earning the incentive equal the marginal costs. Both contract types have identical marginal costs of effort and identical marginal benefits to consumption utility and thus identical marginal benefits to net consumption utility. Where they differ is in their marginal benefits to gain-loss utility, which in the gain contract is the increased likelihood of experiencing a pleasant gain and in the loss contract is the increased likelihood of avoiding an unpleasant loss. If people are loss averse, the gain-loss utility from avoiding the loss is greater than the gain-loss utility from obtaining the gain. In turn, the marginal benefit of effort under loss contracts is greater than under gain contracts. People will therefore work harder under the loss contract than under the gain contract.

Prediction 2. *Among people who are loss averse, performance differences between contracts are increasing in individuals' degree of loss aversion.*

As discussed above, a greater sensitivity to losses than gains leads to performance differences between gain and loss contracts. Larger differences in sensitivity will lead to larger differences in performance. If a person is not at all loss averse, she is equally sensitive to gains and losses, and her performance will not differ between contracts.[6]

Our next two behavioral predictions concern preferences over gain and loss contracts. We consider the decision between selecting into a contract, and potentially earning the incentive, or accepting a fixed payment. The highest amount someone is willing to forgo in order to participate in the contract is her WTP. A person's WTP is determined by her maximum expected utility from working under the contract, i.e., exerting the optimal level of effort as discussed in Prediction 1.

Prediction 3. *If people have dynamically consistent preferences and rational expectations, the willingness to pay for the gain contract will be higher than the willingness to pay for the loss contract.*

Under the behavioral model, expected utility is the sum of expected net consumption utility plus expected gain-loss utility. We first compare expected gain-loss utility across contracts. Gain-loss utility is

[6] If a person is less sensitive to losses than she is to gains, performance will be higher under gain contracts, and the gain-loss gap will increase as loss sensitivity decreases.

positive for gains and negative for losses. As such, gain-loss utility is always (weakly) greater under gain contracts than under loss contracts. This is because the worst a person can do under a gain contract is not receive the incentive; if she does not receive the incentive, the agent will remain at her reference point and derive zero gain-loss utility. Any positive probability of earning the incentive increases her expected gain-loss utility above zero. By contrast, under loss contracts the best a person can do is to keep the incentive—she will remain at her reference point and derive zero gain-loss utility. Any positive probability that she does not earn the incentive decreases her expected gain-loss utility below zero.

We now compare expected net consumption utility, the expected consumption utility minus effort costs as in the standard expected utility model, across contracts:

$$e[u(b)] - c(e), \qquad (1)$$

where $e \in (0, 1)$ is the probability of earning the incentive $b > 0$, $u(\cdot)$ is consumption utility from the incentive, and $c(\cdot)$ is the cost of effort. We assume u is increasing and concave and c is increasing and convex, and we normalize consumption utility from not receiving the incentive to zero. Let e_S^* maximize (1). As discussed in Prediction 1, optimal effort under the loss contract e_L^* will be greater than optimal effort under the gain contract e_G^*. By the same logic, e_G^* will be greater than e_S^* (because effort increases the marginal benefits to gain-loss utility). Because effort under the gain contract e_G^* does not equal e_S^*, it cannot maximize (1)—in particular, it is too high. Effort under the loss contract e_L^* is even farther from maximizing (1), as it is higher than e_G^*. Thus, $e_S^* u(b) - c(e_S^*) > e_G^* u(b) - c(e_G^*) > e_L^* u(b) - c(e_L^*)$. Expected net consumption utility is higher under gain contracts than under loss contracts.

Note that because effort is higher under loss contracts than under gain contracts, expected earnings will also be higher. However, this increase in effort represents a distortion from what is otherwise optimal from the perspective of maximizing consumption utility net of effort costs. It occurs because people are trying to compensate for the negative expected gain-loss utility imposed by facing the threat of potential losses.

Both expected gain-loss utility and expected net consumption utility are lower in loss contracts than in gain contracts. If people have rational expectations, they will anticipate this difference and will have a higher willingness to pay for a gain contract than for a loss contract. If people do not have rational expectations regarding their degree of loss aversion and, in turn, the differential effect of the loss contract on behavior, they will expect their reference point and optimal effort under the loss contract to be the same as it is under the gain contract. In this case, the willingness to pay for the loss contract will be equal to the willingness to pay for the gain contract.

Prediction 4. *If people are dynamically consistent and have rational expectations, differences in willingness to pay will be larger among people who are more loss averse.*

As discussed above, a greater sensitivity to losses than gains decreases WTP for loss contracts through both a decrease in gain-loss utility and a distortion in effort (relative to gain-loss utility and effort under gain contracts). Larger differences in sensitivity lead to larger gaps in WTP for a gain contract compared with a loss contract.

Note that this model assumes that people are dynamically consistent. That is, the preferences of a person when choosing the contract are consistent with her preferences when working under a contract. As we discuss further in Section 6, individuals who have dynamically inconsistent preferences—where the relative weight placed on the cost of effort is disproportionately greater in the period they have to work than in the preceding periods—may actually prefer loss contracts as commitment devices to induce their future selves to work harder.

3. Experiment 1: Effort Under Gain and Loss Contracts

3.1. Experimental Design

To test whether people anticipate loss aversion, we first need to establish that individual effort is indeed differentially affected by the two contract types, which is what we set out to do in our first experiment. Experiment 1 was implemented with 83 subjects at Carnegie Mellon University (CMU), with 4–8 subjects in each session. Subjects were randomized at the session level to either a GAIN or LOSS treatment, which corresponded to working under a gain contract or a loss contract, respectively. Then, subjects participated in a real-effort task where we offered them an incentive based on their performance. Performance in the real-effort task is the primary outcome measure in Experiment 1.

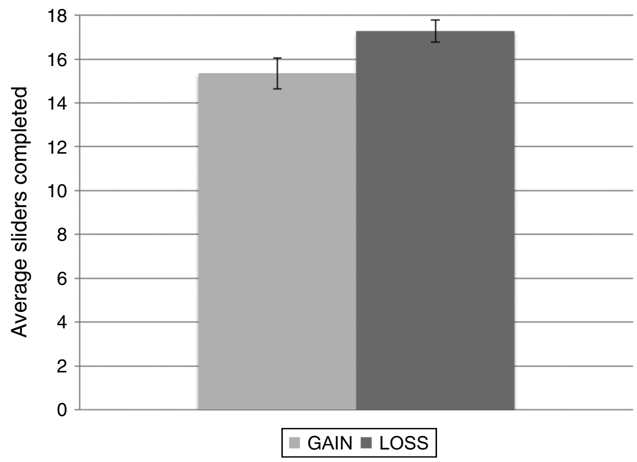
Upon arriving at the lab, subjects were assigned to a private computer station and given the instructions for the real-effort task, which were also read out loud. We used a modified one-shot version of the slider task developed by Gill and Prowse (2012). In this task, subjects complete a series of sliders by moving them sequentially on their computer screen to a randomly assigned point along a bar using their computer mouse. Subjects were incentivized to complete as many sliders as possible (max 30) in 1.5 minutes; see the online appendix (available as supplemental material at http://dx.doi.org/10.1287/mnsc.2015.2402) for the instructions and a screenshot of the task.

All subjects were offered a nonmonetary incentive for completing as many or more sliders than a previously determined exogenous threshold. In both treatments, the threshold was the average performance of subjects who previously completed the same task for a piece rate.[7] In both treatments, the incentive was a custom made T-shirt with an unknown outside value and a subjective personal value (its actual cost was about $8). The exogenous threshold, which was the same for both the GAIN and LOSS treatments, ensures that expectations about the level of effort required to end up with the incentive does not vary by treatment and does not depend on beliefs about the performance of other participants in the experiment. The value of the threshold was also unknown to subjects in advance, and hence beliefs were such that any increase in performance should increase the probability of ending up with the incentive.

In the GAIN treatment, the experimenter held up the T-shirt at the front of the room and told subjects that they would receive it if their performance on the slider task was equal to or above the threshold; otherwise, they would receive nothing. In the LOSS treatment, participants were first given the T-shirt, which remained at their station throughout the session. The experimenter told subjects that they would keep the T-shirt if their performance was equal to or above the threshold; otherwise, they would have to return it. This design created two contracts that were equivalent in payoffs—requiring the same level of effort to end up with the incentive. However, in the GAIN treatment the contract incentivized participants to work to gain the payoff, whereas the LOSS treatment incentivized participants to work to avoid losing the same payoff. After learning about the incentive scheme, subjects performed the slider task.[8]

At the end of the real-effort task, but before participants learned whether they earned the T-shirt, we endowed all participants with an additional $10 and

Figure 1 Average Performance in Gain and Loss Contracts



Notes. Average performance and standard error bars are shown for each treatment. The difference in average performance between a GAIN and LOSS is significant at the $p < 0.05$ level.

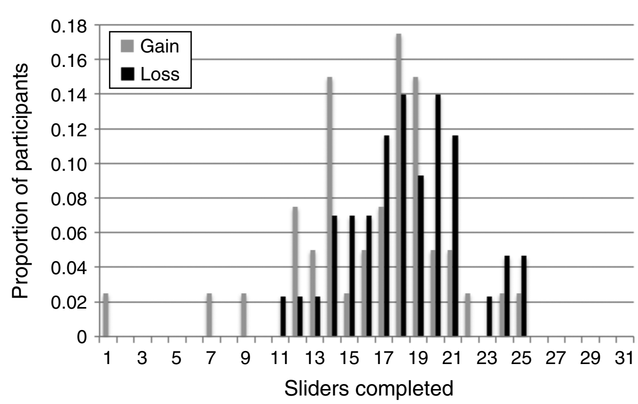
elicited individual loss aversion parameters using a multiple price list (we discuss details and results in Section 5). After completing the experiment, subjects filled out a short survey and received payment from the task including a preannounced show-up fee of $5. Participation in Experiment 1 took about 25 minutes.

3.2. Results

The behavioral model presented in Section 2 predicts that if people are loss averse, the number of sliders completed (performance) will be higher under a loss contract than under a gain contract (Prediction 1). The results of Experiment 1 support this prediction. As illustrated in Figure 1, subjects in the GAIN treatment complete an average of 15.35 sliders ($N = 40$, $SD = 4.50$) and subjects in the LOSS treatment complete an average of 17.28 sliders ($N = 43$, $SD = 3.33$). The 0.4-standard-deviation difference in performance is statistically significant (Student's t-test, two-tailed, $p = 0.03$; Wilcoxon-Mann-Whitney test of distribution, $p = 0.05$).

Figure 2 provides a histogram of performance in each treatment. To ensure that our results are robust

Figure 2 Distribution of Performance in Gain and Loss Contracts



[7] Subjects were told that the threshold was determined by the average performance of a group of CMU participants who worked on the same task but who were paid cash based on their performance. The previous group had worked on the task in the previous year, earning $0.50 per completed slider. Average earnings in the previous task were $9.50, which is similar to the cost of the nonfinancial incentive we offered participants in our experiment.

[8] De Araujo et al. (2015) survey prior work using the Gill and Prowse (2012) slider task and find a weak relationship between effort provision and incentives. We note that our task differs from the original in that we used nonmonetary incentives; the slider had to be moved to a value that was randomly assigned for each iteration rather than the middle value for every slider; an unknown, exogenous threshold was used to determine payment rather than a piece rate; and the task was one-shot with no practice round.

to outliers, we conduct nonparametric permutation tests on the performance distributions under the two contracts. We construct test statistics using permutation methods based on Schmid and Trede (1995) and run one-sided tests for stochastic dominance and separatedness of the distributions (see also Anderson et al. 2011, DiTraglia 2006, Imas 2014). The test statistics identify the degree to which one distribution lies to the right of the other, and take into account both the consistency of the differences between the distributions (i.e., how often they cross) and the size of the differences (i.e., the magnitudes). We compute p-values by Monte Carlo methods with 100,000 repetitions. The results reveal a significant difference between the performance distributions under the gain and loss contracts ($p = 0.02$), implying that the performance distribution in the LOSS treatment is shifted to the right of the performance distribution in the GAIN treatment.

4. Experiment 2: Anticipation of Loss Aversion and Choice Between Contracts

4.1. Experimental Design

Next, we examine whether people anticipate loss aversion in line with the standard behavioral model—that is, whether they prefer to work under a gain rather than under a loss contract. In Experiment 2, we elicited subjects' WTP to participate in one of the two incentive schemes used in the first experiment. As discussed in Section 2, if people anticipate the differential effect of the loss contract demonstrated by Experiment 1, then the WTP to work under the gain contract will be higher than the WTP to work under the loss contract (Prediction 3). The elicited WTP is the primary outcome measure for Experiment 2.

Experiment 2 was implemented using 85 subjects at CMU, with 4–8 subjects in each session. Using a between-subjects design, we randomized subjects to one of the two treatments described in Experiment 1: GAIN or LOSS. As in Experiment 1, subjects were randomized to treatment at the session level. Unlike in Experiment 1, rather than simply participating in the real-effort task, subjects were asked to indicate their WTP to participate.

Upon arriving at the lab, subjects were assigned to a private computer station and given the instructions, which were also read out loud. The experiment proceeded in two parts. In the first part, we explained the slider task (using the same instructions as in Experiment 1) and then elicited WTP to participate in the slider task with the T-shirt as the incentive. In both the GAIN and LOSS treatments, the experimenter held up the T-shirt at the front of the room and read the instructions describing either the gain or the loss contract from Experiment 1.

To elicit WTP, we asked subjects to make a series of trade-offs between either working under the respective contract (gain or loss) or receiving a sum of money.[9] We used a multiple price list, which has been employed as an incentive-compatible method to elicit attitudes for risk (Holt and Laury 2002, Sprenger 2015, Charness et al. 2013) and time preferences (Andersen et al. 2008). In our paradigm, participants made a series of decisions between either participating in the task or not participating and receiving an "additional payment" at the end of the experiment. The additional payment was \$0 for the first decision and increased to \$5 by the last decision in increments of \$0.50 (see the online appendix for the instructions).[10]

We used a die roll to randomly choose a single decision from the list to be implemented. The additional payment offered in the implemented decision determined the opportunity cost of working under the contract. If a subject had indicated that she was willing to forego the payment and participate, then she participated and received no additional compensation. If a subject indicated she preferred to receive the payment, then she waited at the computer terminal during the 1.5 minutes of the task instead of participating (and received the additional payment at the end of the experiment).

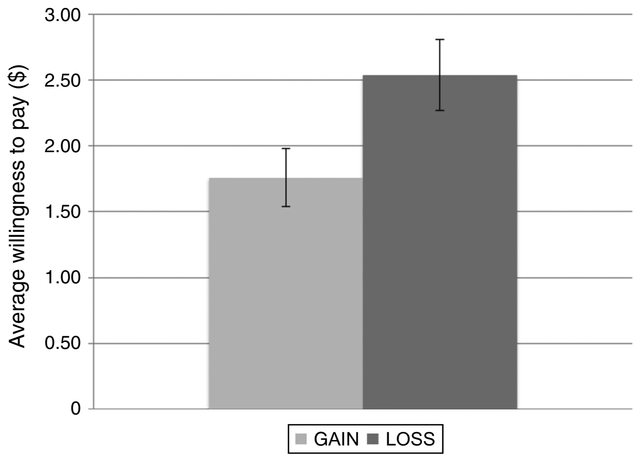
In the second part of the experiment, those who were willing to pay the randomly selected cost participated in the contract described in Experiment 1. In the GAIN treatment, those who elected to participate completed the slider task and received the T-shirt if their performance was equal to or above the performance threshold. Those who opted to participate in the LOSS treatment were first given the T-shirt to keep at their desk, performed the slider task, and got to keep the T-shirt if their performance was equal to or above the threshold, or they had to return it if their performance was below the threshold. Those who opted not to participate waited at their computer terminals until the slider task was complete.

Finally, at the end of the experiment (before participants learned whether they earned the T-shirt),

[9] In this design, subjects are paying to participate with the foregone payoff from not participating. This is more natural for subjects who are used to earning money, rather than spending money, to participate in experiments. In addition, it models the opportunity costs employees are willing to forgo in order to enter a contract and matches the theoretical model discussed in Section 2.

[10] Note that rather than presenting participants with a list of decisions, each decision was presented one at a time on a separate page. This was done to make the decision problem more natural, reducing confusion that typically results in inconsistent answers such as multiple switch points when price list measures are used (e.g., Charness et al. 2013).

Figure 3 **Average WTP to Participate in Gain and Loss Contracts**



Notes. Average WTP and standard error bars are shown for each treatment. The difference in average WTP between GAIN and LOSS is significant at the $p < 0.05$ level.

we elicited individual loss aversion parameters. As in Experiment 1, we endowed all participants with an additional $10 and used multiple price lists, which we describe in detail in Section 5. At the end of the session, all participants filled out a short survey and received their preannounced $5 show-up fee plus additional payments earned in the experiment. The session lasted approximately 45 minutes.

4.2. Results

The behavioral model discussed in Section 2 predicts that WTP to work under a gain contract will be higher than WTP to work under a loss contract (Prediction 3). To test this, we compare participants' maximum WTP to participate in the gain contract to maximum WTP to participate in the loss contract. We measure maximum WTP as the lowest additional payment an individual chooses to accept rather than participate in the slider task and potentially earn the T-shirt.

The results from Experiment 2 do not support the prediction of the standard behavioral model. As shown in Figure 3, the maximum average WTP is higher for the loss contract ($2.54, $N = 41$, SD = $1.73) than for the gain contract ($1.76, $N = 44$, SD = $1.48). This represents a statistically significant difference between LOSS and GAIN, which goes in the opposite direction predicted by the theory (Student's t-test, two-tailed, $p = 0.03$; Wilcoxon-Mann-Whitney test of distribution, $p = 0.04$).[11] Because WTP is censored at $0 and $5, we also use a Tobit regression to confirm our results. Regressing a treatment dummy (LOSS = 1, GAIN = 0) on WTP reveals a similar result: the coefficient on the dummy is 0.93 and is statistically significant ($p = 0.03$).

[11] Of 85 participants, 92% have consistent choices; i.e., once they choose to accept the additional payment (rather than participate), they continue to do so for all higher values of the payment. Dropping participants who are not consistent (five subjects in LOSS and two in GAIN) does not affect the results ($p = 0.04$ for the Student's t-test and the Wilcoxon-Mann-Whitney test when comparing LOSS and GAIN).

Figure 4 **Distribution of WTP to Participate in Gain and Loss Contracts**

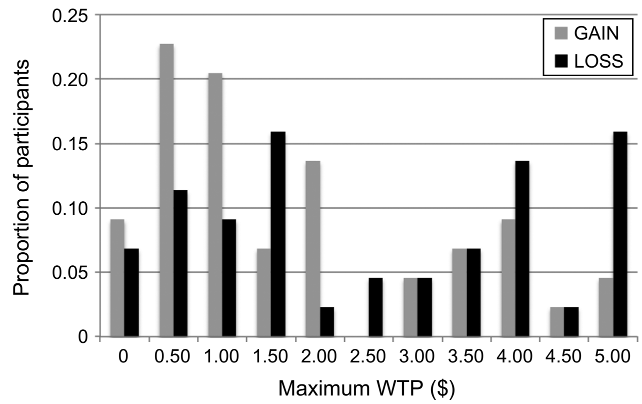


Figure 4 presents a histogram of maximum WTP by treatment. The majority of participants in the GAIN treatment prefer to receive $1.00 (or less) rather than work under a gain contract, and less than 5% are willing to pay the maximum allowed amount of $5. By contrast, nearly a third of participants in the LOSS treatment are willing to pay at least $4 to work under a loss contract, with half of those willing to forgo $5 in order to participate. To test for differences between the distributions, we run the same nonparametric distribution test as in Experiment 1. The results show that the distribution of WTP in the LOSS treatment is significantly to the right of the distribution in the GAIN treatment ($p = 0.03$).

5. The Effect of Loss Aversion on Performance and Contract Preferences

Anticipation of loss aversion in our standard behavioral model predicts that WTP to enter loss contracts should be lower than WTP to enter gain contracts, a hypothesis that is rejected by our data. An alternative possibility, that people fail to anticipate loss aversion, predicts no difference in WTP between the two contracts. This alternative is also rejected by our data. Instead, we find that people do anticipate loss aversion but react to it in the opposite direction predicted by the standard behavioral model—people actually prefer the loss contract to the gain contract. What mechanisms drive these results? To shed light on this question, we examine the relationship between participants' behavior and a separately elicited individual loss aversion parameter.

5.1. Construction of the Loss Aversion Parameter

At the end of Experiments 1 and 2, we elicited individuals' preferences over a series of gambles using multiple price lists. Subjects received an additional \$10 and made a series of 30 binary decisions, choosing between sure payoffs or risky payoffs with outcomes to be determined by a coin flip (see the online appendix for the instructions). Only one decision was randomly determined to be paid out at the end of the experiment. Similar to Abdellaoui et al. (2008), the set of 30 decisions allows us to separately estimate the three parameters of a prospect theory value function, α, β, and λ, for each individual:

$$v(x) = \begin{cases} (x)^{\alpha} & \text{if } x \geq 0, \\ -\lambda(-x)^{\beta} & \text{if } x < 0, \end{cases}$$

where α is the risk aversion parameter in the gain domain, β is the risk aversion parameter in the loss domain, and λ is the loss aversion parameter. We aimed to separately elicit each of the three parameters because, as noted in Section 2, the theory predicts an explicit relationship between behavior and the loss aversion parameter λ rather than risk preferences in general. As such, the three-part identification outlined below is required to provide identified evidence for or against the theoretical predictions.

To identify the parameter α, we asked participants to make a series of choices over gambles where all of the outcomes were positive. The multiple price list offered subjects a series of 10 decisions between a lottery and a sure amount, where the lottery was constant for each decision and the sure amount gradually increased. The risky option was \$0 with 50% probability and \$5 with 50% probability; the sure option started at \$0.50 in the 1st decision and increased in \$0.50 increments to \$5 in the 10th decision. The choice pattern generally observed is that subjects start out choosing the lottery and then switch to the sure outcome when the sure outcome becomes large enough. The choice at which a subject switches is taken as the indifference point between the lottery and the sure outcome. Since all outcomes are positive, the formulation of the decision problem does not involve β or λ, and it is straightforward to estimate the parameter α.

The parameter β is estimated in a similar manner using a multiple price list with only negative outcomes. In the second set of 10 decisions, subjects made the choice between either a risky option of −\$5 with 50% probability and \$0 with 50% probability or a sure option of −\$0.50 through −\$5.00 in increments of \$0.50. Here, the decision problem does not involve α or λ, so β is separately identified.

The final multiple price list offered mixed gambles. In the last set of 10 decisions, subjects made the choice between either a risky option of \$5 with 50% probability and −\$1 through −\$10 with 50% probability in increments of \$1 or a sure option of \$0. We use the multiple price list with mixed gambles to estimate λ by setting up the decision problem at the indifference point and using the α and β parameters we estimated from the other two multiple price lists.

People who exhibited extreme risk attitudes by never switching were excluded from the analysis, leaving 134 subjects (for similar exclusion criteria, see Andreoni and Sprenger 2012, Sprenger 2015). The median λ in the sample is 1.59 (median $\alpha = 0.87$, $\beta = 0.87$), which corresponds to significant loss aversion as reported in prior work (Tversky and Kahneman 1992, Abdellaoui et al. 2008).

5.2. Loss Aversion and Performance

Standard behavioral theory predicts that in Experiment 1, we should observe that the difference in performance between the loss and gain contracts should be increasing with the degree of loss aversion (Prediction 2). The more loss averse the individual, the harder she is willing to work to avoid experiencing a loss under the loss contract. Under a gain contract, since the individual does not face the possibility of a loss, we do not expect a significant relationship between loss aversion and performance.[12]

In Table 1, we examine the relationship between an individual's estimated loss aversion parameter λ and performance in gain and loss contracts. In all regressions, the outcome variable is the number of sliders completed. The first column includes participants in the GAIN treatment only, and the second column includes participants in the LOSS treatment only. The third column includes all participants in Experiment 1 and adds a dummy variable for the treatment (0 = GAIN, 1 = LOSS) and the interaction of the treatment with the loss aversion parameter λ. Our estimates offer suggestive evidence for Prediction 2. Loss aversion has no significant relationship with performance in the GAIN treatment ($p = 0.44$). In the LOSS treatment, the coefficient on λ is large, positive, and marginally significant ($p = 0.06$). The interaction term is also positive but not significant at conventional levels ($p = 0.13$). These results offer suggestive support for the prediction of a differential effect of loss aversion on performance between the two contract types.

5.3. Loss Aversion and Contract Preferences

The standard behavioral theory predicts that in Experiment 2, if individuals anticipate loss aversion, we

[12] Since loss aversion is measured as sensitivity to losses relative to gains, finding $\lambda > 1$ could indicate gain loving rather than loss aversion. In our analysis, we follow Abdellaoui et al. (2008) and Sprenger (2015) in interpreting $\lambda > 1$ as indicative of an aversion to losses. Note, however, that the between-treatment analysis to test Prediction 2, as well as Prediction 4 below, holds under both interpretations.

Table 1 Effect of Loss Aversion on Performance

	GAIN	LOSS	All treatments
λ	0.14	1.38*	0.14
	(0.18)	(0.71)	(0.16)
LOSS (= 1)			−1.20
			(1.60)
$\lambda \times$ LOSS			1.24
			(0.82)

Notes. The table presents ordinary least squares estimates with standard errors in parentheses. The dependent variable is performance measured by the number of sliders completed; λ is the estimated loss aversion parameter. The "GAIN" column estimates the effect of λ on performance in the GAIN treatment; the "LOSS" column estimates the effect in the LOSS treatment. The "All treatments" column estimates the effect of λ, the LOSS treatment, and their interaction.

*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Table 2 Effect of Loss Aversion on WTP

	GAIN	LOSS	All treatments
λ	−0.02	0.18*	−0.21
	(0.06)	(0.09)	(0.16)
LOSS (= 1)			0.50
			(0.82)
$\lambda \times$ LOSS			0.20*
			(0.11)

Notes. The table presents ordinary least squares estimates with standard errors in parentheses. The dependent variable is WTP measured by the lowest additional payment participants choose rather than working under the loss or gain contract; λ is the estimated loss aversion parameter. The "GAIN" column estimates the effect of λ on WTP in the GAIN treatment; the "LOSS" column estimates the effect in the LOSS treatment. The "All treatments" column estimates the effect of λ, the LOSS treatment, and their interaction.

*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

should observe larger treatment effects on WTP among more loss-averse individuals (Prediction 4). Individuals who are more loss averse should be willing to accept a smaller additional payment instead of working under the loss contract since the loss frame will hurt them most. That is, the WTP to enter the loss contract should be significantly decreasing in λ, whereas we do not expect a significant relationship between λ and WTP to enter the gain contract.

Examining the relationship between loss aversion and preferences between contracts in Experiment 2 does not support the prediction of the standard behavioral theory; in fact, our results suggest the opposite pattern. We find that those who exhibit greater loss aversion are willing to pay *more* to participate in the loss contract than the gain contract. Table 2, which has the same structure as Table 1 except that the dependent variable is WTP in Experiment 2, summarizes these results. The first column includes participants in the GAIN treatment only. The second column includes participants in the LOSS treatment only. The third column includes all participants in Experiment 2 and adds a dummy variable for the treatment (0 = GAIN, 1 = LOSS) and the interaction of the treatment with the loss aversion parameter λ. The standard behavioral model predicts a more negative λ coefficient in the LOSS treatment than the GAIN treatment, and thus the interaction term should be negative. Our estimates do not support these predictions. The coefficient on λ is small and not significant for participants in the GAIN treatment. However, it is *positive* and marginally significant for participants in the LOSS treatment ($p = 0.07$). The interaction term is also positive and marginally significant ($p = 0.07$). Rather than avoiding loss contracts, our results suggest that more loss averse individuals are *more* likely to select into them.

6. Interpretation

Our results demonstrate that, as predicted, individuals do exert higher effort under loss contracts; however, in contrast to the theory, they prefer loss contracts to gain contracts. Further, we find evidence that those who are most sensitive to losses are also the ones who are most likely to select into loss contracts. That is, the most loss-averse individuals both work harder *and* have the highest WTP for loss contracts. As discussed in Section 2, the greater the upward distortion in effort under loss contracts (relative to gain contracts), the larger the utility costs experienced under loss contracts. Why then would people want to enter a contract that induces them to work "too hard" by imposing the pain of potential losses?

We suggest that one possible mechanism driving our results is that individuals select into loss contracts as a commitment device because of a dynamic inconsistency in their preferences. In the workplace, dynamic inconsistency can take the following form: individuals may have an ex ante preference for working hard and maximizing their chance of earning a performance bonus, but when it comes time to exert the actual effort, their preferences reverse—they shirk and fail to earn the incentive (Kaur et al. 2010, Cadena et al. 2011).[13] If individuals are sophisticated about their dynamic inconsistency, they may anticipate this preference reversal and value commitment devices that impose additional costs on shirking (Laibson 1997, O'Donoghue and Rabin 1999).

[13] Although models of dynamic inconsistency are generally applied to choices occurring days, weeks, or months apart, previous work has demonstrated that dynamic inconsistency can also occur in choices made over the span of several minutes (Solnick et al. 1980, McClure et al. 2004, Brown et al. 2009). This suggests that the time frame of our study may be sufficient for dynamic inconsistency to emerge.

In particular, a potential mechanism for favoring a loss contract is that individuals with dynamically inconsistent preferences may recognize that they will work harder under the threat of potential losses than they would for potential gains and select into loss contracts in order to commit their future selves to improved performance and higher expected earnings. As such, in line with our empirical results but in contrast to the predictions of the standard behavioral model, individuals who are loss averse should both work harder under loss contracts *and* be more willing to enter into them.[14]

Although our results are consistent with people's anticipation of loss aversion as a commitment device, we cannot rule out alternative mechanisms. For example, the anticipation of obtaining the reward sooner under the loss contract than the gain contract may cause some to perceive the status quo as owning the shirt. As such, individuals offered the loss contract would be more likely to select the contract to retain the possibility of keeping the reward. De Quidt (2014) also explores potential explanations for the observed preference for loss contracts. He compares loss-framed contracts that offer a high base pay with a potential penalty to gain-framed contracts that offer a low base pay with a potential bonus. Although the expected pay—i.e., base pay net of the potential bonus or penalty—is equivalent in both contracts, De Quidt concludes that the higher stated base pay in the loss-framed contracts is more salient to workers than the expected pay, which they must calculate themselves. Thus, in De Quidt's setting, higher entry rates into loss-framed contracts compared with gain-framed contracts may be driven by differences in salience across contracts. However, this is less relevant in our context, where we offer a single fixed reward and vary only the timing of when workers will receive it.

Our study informs a growing literature demonstrating that people are willing to take up commitment contracts in contexts where dynamic inconsistencies in preferences exist (Ashraf et al. 2006; Augenblick et al. 2015; Hsiaw 2013, 2015; Sadoff et al. 2015).[15] These contracts often include a loss component such as penalties for failing to meet performance targets (Kaur et al. 2015, Royer et al. 2015, Giné et al. 2012, Schwartz et al. 2014, John et al. 2011). However, to our knowledge, little research has explicitly examined the role of loss aversion in demand for commitment.[16] More work is needed to understand how loss aversion affects preferences between contracts as well as its possible role in the take up of commitment devices.

7. Conclusion

Understanding the extent to which there are trade-offs between employee productivity and employee preferences is critical for managers and organizations considering the use of loss contracts. Standard behavioral models predict that such a trade-off exists: employees will work harder under loss contracts than they will under equivalent gain contracts, but, anticipating loss aversion, employees will select into gain contracts rather than loss contracts. Despite growing interest in the use of loss contracts, little is known about the extent to which these trade-offs exist in practice.

This study is among the first to examine both performance and preferences for gain versus loss contracts. We find that although individuals work harder under a loss contract than they do under a gain contract (as predicted), they prefer the former to the latter (in contrast to the standard prediction). We also find heterogeneity in susceptibility to loss contracts. More loss-averse individuals exert higher effort *and* have a greater preference for loss contracts. This suggests that firms may not need to pay a premium to persuade potential employees to work under loss contracts and that offering such contracts could be beneficial for all parties.

Our results also inform theory. Whether people anticipate loss aversion and how they react to it has important implications for modeling the decision making of individuals with reference-dependent preferences. Our study is among the first to explore the general question of whether people anticipate loss aversion. We find evidence that people do anticipate loss aversion, but rather than deterring them as predicted under standard models of reference dependence, greater loss aversion may make the preference for loss contracts stronger. Related work also finds evidence that people do not anticipate loss aversion as predicted by standard behavioral models with rational expectations. In the context of eliciting willingness to pay and willingness to accept values for a mug,

[14] It should be noted that the benefits to the individual of choosing a loss contract as a commitment device are from the perspective of a long-run self, since the "self" who is actually exerting the effort under a loss contract is worse off in expectation. As discussed in Section 2, barring a wedge between the preferences of the long-run and short-run selves—dynamic inconsistency—even if people anticipate the increase in expected earnings under the loss contract, they should still prefer the gain contract.

[15] In a related line of research, Arlen and Tontrup (2015) argue that people may take up contracts that delegate choice as a commitment device to counter endowment effects (in their case, biases against trading).

[16] De Quidt (2014) also explores the role of commitment in explaining the preference for penalty contracts. He finds that a significant gap in entry rates persists when loss contracts have less commitment value (because the incentivized task requires less costly effort), suggesting that in his context the salience of the penalty contract may outweigh other factors driving take up.

Loewenstein and Adler (1995) and Van Boven et al. (2000) find evidence that prior to being endowed, subjects underestimate their willingness to accept.

Examining preferences and selection effects is crucial for applying behavioral insights in management and policy more broadly. For example, several studies find that people are reluctant to realize losses on assets (Barberis 2013). If this is the case, whether and how people anticipate such behavior is critical for understanding their trading decisions. The anticipation of future preferences has been explored in other areas, such as models of rational addiction (Becker and Murphy 1988), projection bias (Loewenstein et al. 2003), and time preferences (Laibson 1997, O'Donoghue and Rabin 1999). These models allow us to evaluate the extent to which we can view individuals' decision making as rational and the extent to which people may be making optimization mistakes. Further studies in the lab and field can help shed light on this important yet underexplored question in the context of reference-dependent preferences.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2015.2402.

Acknowledgments

The authors thank Uri Gneezy, George Loewenstein, Yuval Rottenstreich, David Schkade, Richard Thaler, three anonymous contributors, and the assistant editor, as well as numerous seminar and conference participants, for valuable feedback. Christa Gibbs, Stephanie Schwartz, and Michael Rosenberg provided truly outstanding research assistance. This research has been conducted with institutional review board approval. Please direct correspondence to the second author.

Appendix A. Proofs of Theoretical Predictions

To formalize the intuition discussed above, consider a representative agent whose utility V is given by

$$V = V(e, b, r) = e[u(b) + v(b \mid r)] + [1-e]v(0 \mid r) - c(e),$$

where an individual receives a payoff $b > 0$ with probability equal to effort $e \in (0, 1)$ and receives 0 with probability $1-e$; $u(\cdot)$ corresponds to standard consumption utility over the payoff, $v(\cdot \mid r)$ is the gain-loss prospect theory value function, and $c(\cdot)$ is the cost of effort e that an individual exerts to obtain the bonus.[17] Let u be an increasing and concave function of b, and let c be an increasing and convex function of e. Normalize $u(0) = 0$. We define the utility derived in relation to a reference point r as follows:

$$v(x \mid r) = \begin{cases} (x-r)^{\alpha} & \text{if } x \geq r, \\ \lambda(x-r)^{\beta} & \text{if } x < r, \end{cases}$$

where $\lambda > 1$ is the loss aversion parameter, and we assume $\alpha = \beta$.[18] An individual chooses optimal effort e^* to maximize overall utility V:

$$\max_e V(e, b, r) = \max_e \{e[u(b) + v(b \mid r)] + [1-e]v(0 \mid r) - c(e)\}.$$

Taking the status quo model of prospect theory, when working under the gain contract ($r = 0$), optimal effort in the gain contract e_G^* satisfies the following first-order condition:

$$c'(e_G^*) = u(b) + b^{\alpha}. \qquad \text{(A1)}$$

Under the loss contract ($r = b$), optimal effort in the loss contract e_L^* satisfies the following first-order condition:

$$c'(e_L^*) = u(b) + \lambda b^{\beta}. \qquad \text{(A2)}$$

The first-order conditions lead to our first prediction.

Proof of Prediction 1. If people are loss averse, performance will be higher under a loss contract than under a gain contract.

Proof. Under the assumptions that costs c are convex, the left-hand sides of Equations (A1) and (A2) are increasing in effort e. If $\lambda > 1$, then the right-hand side of Equation (A2) is greater than the right-hand side of Equation (A1), $u(b) + \lambda b^{\beta} > u(b) + b^{\alpha}$ (under the assumption that $\alpha = \beta$).[19] Thus, optimal effort in the loss contract will be greater than optimal effort in the gain contract, $e_L^* > e_G^*$. Under the assumption that performance is an increasing function of effort, performance will be higher under the loss contract than the gain contract. If $\lambda = 1$, optimal effort and performance will be the same in loss and gain contracts; if $\lambda > 1$, effort and performance will be higher in loss contracts. □

Note that by the same logic as Prediction 1, if $\lambda < 1$, effort and performance will be lower in loss contracts than in gain contracts.

Proof of Prediction 2. Among people who are loss averse, performance differences between contracts are increasing in individuals' degree of loss aversion.

Proof. As discussed above, among people who are loss averse, greater optimal effort under loss contracts versus gain contracts is due to greater sensitivity to losses than gains $\lambda > 1$. The difference in optimal effort $e_L^* - e_G^*$ is increasing in the difference in gain-loss utility λ (under the assumption that effort costs are convex). That is, the more loss averse someone is, the harder she will work to avoid losses relative to working for gains. Given the assumption that performance is an increasing function of effort, performance differences between loss and gain contracts will be higher among more loss-averse people compared with less loss-averse people. □

[17] In practice, we measure effort through performance, which we assume is increasing in effort e.

[18] Tversky and Kahneman (1992) estimate $\lambda = 2.25$ and median $\alpha = \beta = 0.88$. As discussed in Section 5, we estimate λ, α, and β separately in our data and find similar support for our assumptions.

[19] This result does not require that $\alpha = \beta$. It is sufficient that $\lambda > b^{\alpha}/b^{\beta}$.

Note that among people for whom $\lambda < 1$, effort and performance are higher under gain contracts, and this difference increases as λ decreases. That is, if a person is less sensitive to losses than she is to gains, performance will be higher under gain contracts, and the gain-loss gap will increase as loss sensitivity decreases.

We now consider the participation constraint in which individuals are offered a choice between a certain amount $w \geq 0$ or the chance to participate in the task and earn the uncertain payoff b. An individual will participate if $V(e^*, b, r) \geq u(w) + w^\alpha$, where $V(e^*, b, r)$ is the agent's utility under optimal effort e^*. Note that, as argued and demonstrated by Kahneman et al. (1990), individuals choosing between goods without being endowed with either behave as if their reference point was the status quo. We follow this assumption when outlining the decision problem between a certain payoff and the contract.

The greatest amount the agent would be willing to forgo in order to participate (i.e., maximum willingness to pay) in the gain contract w_G solves the following:

$$u(w_G) + w_G^\alpha = V(e_G^*, b, 0), \tag{A3}$$

where e_G^* is optimal effort under the gain contract as defined in Equation (A1).

Assuming that the agent has rational expectations over her preferences under the loss contract, her maximum WTP to participate in the loss contract w_L solves the following:

$$u(w_L) + w_L^\alpha = V(e_L^*, b, b), \tag{A4}$$

where e_L^* is optimal effort under the loss contract as defined by Equation (A2).

Proof of Prediction 3. If people have dynamically consistent preferences and rational expectations, willingness to pay for the gain contract will be higher than willingness to pay for the loss contract, $w_G > w_L$.

Proof. Subtracting Equation (A4) from Equation (A3) gives

$$[u(w_G)+w_G^\alpha]-[u(w_L)+w_L^\alpha]=V(e_G^*,b,0)-V(e_L^*,b,b). \tag{A5}$$

We will show that the right-hand side is positive, which implies that $w_G > w_L$, under the assumption that u is increasing. Expanding terms,

$$\begin{aligned} &V(e_G^*, b, 0) - V(e_L^*, b, b)\\ &= (e_G^*[u(b) + b^\alpha] - c(e_G^*)) - (e_L^* u(b) + [1 - e_L^*]\lambda(-b^\beta) - c(e_L^*))\\ &= ([e_G^* u(b) - c(e_G^*)] - [e_L^* u(b) - c(e_L^*)])\\ &\quad + (e_G^* b^\alpha + [1 - e_L^*]\lambda b^\beta). \end{aligned} \tag{A6}$$

We first consider the term $e_G^* b^\alpha + [1 - e_L^*]\lambda b^\beta$ from (A6), which is the difference in expected gain-loss utility under gain and loss contracts. From the assumptions that $e \in (0, 1)$, $b > 0$, and $\lambda > 1$, the term is positive:[20]

$$e_G^* b^\alpha + [1 - e_L^*]\lambda b^\beta > 0. \tag{A7}$$

[20] Note that this result requires only reference-dependent preferences and does not depend on the degree of loss aversion as long as $\lambda > 0$.

We next consider the term $[e_G^* u(b) - c(e_G^*)] - [e_L^* u(b) - c(e_L^*)]$, which is the difference between gain and loss contracts in expected consumption utility net of costs. This is the expected utility from the standard framework where an agent chooses effort to maximize the following objective function:

$$\max_e \{e[u(b)] - c(e)\}.$$

Optimal effort under the standard framework e_S^* satisfies the following first-order condition:

$$c'(e_S^*) = u(b). \tag{A8}$$

The right-hand side of Equation (A8) is less than the right-hand side of Equation (A1) (under the assumption that $b > 0$). Thus, $e_G^* > e_S^*$ (under the assumption that effort costs are convex). From Prediction 1, if people are loss averse, then $e_L^* > e_G^*$. Since e_S^* optimizes Equation (A8), e_G^* and e_L^* cannot be the optimal effort—they are too high. Because $e_L^* > e_G^*$, e_L^* is further from the optimal e_S^* than e_G^*. Thus, $e_S^* u(b) - c(e_S^*) > e_G^* u(b) - c(e_G^*) > e_L^* u(b) - c(e_L^*)$, and

$$[e_G^* u(b) - c(e_G^*)] - [e_L^* u(b) - c(e_L^*)] > 0. \tag{A9}$$

By Equations (A5)–(A7) and (A9), $w_G > w_L$. The maximum WTP for the gain contract is higher than the maximum WTP for the loss contract. □

Note that if people do not have rational expectations regarding their degree of loss aversion and, in turn, the differential effect of the loss contract on behavior, they will expect their reference point and optimal effort under the loss contract to be the same as it is under the gain contract. In this case, the maximum WTP for the loss contract will be equal to maximum WTP for the gain contract, given by (A3), $w_L = w_G = V(e_G^*, b, 0)$.

Our last prediction follows.

Proof of Prediction 4. If people are dynamically consistent and have rational expectations, differences in willingness to pay will be larger among people who are more loss averse.

Proof. We will show that the right-hand side of Equation (A5) is increasing in λ. This implies that the difference in WTP for gain and loss contracts $w_G - w_L$ is increasing in loss aversion (under the assumptions that u is increasing and concave). Differentiating with respect to λ gives

$$\begin{aligned} &\frac{\partial}{\partial\lambda}(V(e_G^*, b, 0) - V(e_L^*, b, b))\\ &\quad = -\frac{\partial e_L^*}{\partial\lambda} u(b) - \frac{\partial e_L^*}{\partial\lambda}\lambda b^\beta + [1 - e_L^*]b^\beta + \frac{\partial c}{\partial e_L^*}\frac{\partial e_L^*}{\partial\lambda}\\ &\quad = [1 - e_L^*]b^\beta + \frac{\partial e_L^*}{\partial\lambda}\left(\frac{\partial c}{\partial e_L^*} - [(u)b + \lambda b^\beta]\right)\\ &\quad = [1 - e_L^*]b^\beta, \end{aligned}$$

where the final equality follows from Equation (A2), which shows that $c'(e_L^*) - [(u)b + \lambda b^\beta] = 0$ evaluated at e_L^*. The right-hand side is positive $[1 - e_L^*]b^\beta > 0$ under the assumptions that $e \in (0, 1)$ and $b > 0$. Thus $V(e_L^*, b, 0) - V(e_L^*, b, b)$ is increasing in λ, which implies that the difference in the willingness to pay for gain and loss contracts $w_G - w_L$ is increasing in individuals' degree of loss aversion. □

Appendix B. Description of Pilot Experiments

Below we describe the design and results of two pilot experiments conducted prior to the experiments discussed in the main text. Pilot Experiment 1 is analogous to Experiment 1 (see Section 3). Pilot Experiment 2 is analogous to Experiment 2 (see Section 4).

Pilot Experiment 1: Effort Under Gain and Loss Contracts

Experimental Design. Pilot Experiment 1 was implemented among 62 participants at the University of California, San Diego. Subjects were randomized at the session level to either a GAIN or LOSS treatment and then participated in a one-shot task (sessions included six people on average and lasted about 15 minutes).

Upon arriving in the lab, subjects were assigned to a computer station and given the instructions, which were read aloud. In both treatments, we first explained the task students would perform and then offered a performance-based incentive. For the real-effort task, we used the slider task discussed in Section 3. Subjects had two minutes to move up to 48 "sliders."

All subjects were offered an incentive for correctly completing more sliders than a previously determined threshold. The threshold was set within each treatment such that half of the participants in each group were expected to receive the incentive.[21] In the GAIN treatment, subjects received the incentive if their performance on the slider task was equal to or above the threshold. In the LOSS treatment, participants were endowed with the incentive before performing the slider task and were told they would keep the incentive if their performance was equal to or above the threshold. If their performance was below the threshold, participants in the LOSS treatment had to return the incentive. This design created two payoff-equivalent contracts: one framed as a gain and the other framed as a loss (the intratreatment threshold ensures that earnings do not differ across treatments even if average effort does).

In both treatments, the incentive was a custom made T-shirt with an unknown outside value and a subjective personal value (its actual cost was about $8). In the GAIN treatment, the experimenter held up the T-shirt at the front of the room and told subjects they would receive it if their performance on the slider task was equal to or above the threshold; otherwise, they would receive nothing. In the LOSS treatment, participants were given a T-shirt, which remained at their station throughout the session. The experimenter told subjects that they would keep the T-shirt if their performance was equal to or above the threshold; otherwise, they would have to return it. Subjects then performed the slider task for two minutes. After completing the task, subjects filled out a short survey and received payment, including a show-up fee of $5.

[21] The threshold was determined by the average performance from a randomly chosen previous session of the same treatment. Participants were informed of what constituted the threshold, but not its value, prior to performing the effort task. In the first session of each treatment, we used an average from a previous pilot study.

Results. Similar to the results of Experiment 1, the results in Pilot Experiment 1 support the prediction of the standard behavioral model that performance will be higher in loss contracts than in gain contracts. Subjects in the GAIN treatment completed an average of 11.88 sliders ($N = 32$, $SD = 5.55$) compared with an average of 15.27 sliders ($N = 30$, $SD = 4.44$) in the LOSS treatment. The 0.6-standard-deviation difference in performance is statistically significant at the $p < 0.01$ level.

Pilot Experiment 2: Anticipation of Loss Aversion and Choice Between Contracts

Experimental Design. In Pilot Experiment 2, we examine whether people anticipate loss aversion—that is, whether they are more likely to select into a gain rather than a loss contract. To do this, we elicited participants' willingness to pay to participate in each of the two incentive schemes used in Pilot Experiment 1.

Pilot Experiment 2 was implemented among 60 participants at the University of Wisconsin–Madison BRITE (Behavioral Research Insights through Experiments) Laboratory. Using a between-subjects design, we elicited willingness to pay to participate in one of the two treatments described in Pilot Experiment 1: GAIN or LOSS. As in Pilot Experiment 1, we randomized at the session level (sessions included 10 people on average and lasted about 40 minutes).

Upon arriving in the lab, subjects were assigned to a computer station and given the instructions, which were read aloud. The experiment proceeded in two parts. In the first part, subjects were given two minutes to participate in the slider task for no pay. In the second part, we elicited WTP to participate in an incentivized version of the task. In the GAIN treatment, the experimenter held up the T-shirt at the front of the room and read the instructions describing the gain contract from Pilot Experiment 1. The LOSS treatment was identical except that the experimenter read the instructions describing the loss contract from Pilot Experiment 1.

Subjects were then asked to indicate their maximum WTP out of their $10 show-up fee to work under the offered contract. We elicited WTP using a multiple price list. In our paradigm, participants made a series of decisions between paying a price and participating or paying nothing and not participating. The decision to not participate was constant (i.e., $0), whereas the price to participate increased from $0 to $10 from the first decision to the last. We then used a die roll to randomly choose a single decision from the list to be implemented. If a subject indicated she was willing to pay the chosen cost, she participated and the cost was deducted from her show-up fee. If she indicated she was not willing to pay the chosen cost, she did not participate and nothing was deducted from her show-up fee.

In the GAIN treatment, those who paid to participate completed the slider task and received the T-shirt if their performance was above average. Participating subjects in the LOSS treatment were first given the T-shirt, then performed the slider task, and either got to keep the T-shirt or had to return it, again depending on their performance. At the end of the session, all participants filled out a short survey and received payment.

Results. Similar to Experiment 2, the results from Pilot Experiment 2 do not support the prediction of the standard behavioral model that WTP will be higher for the gain contract than the loss contract. As in Experiment 2, the average WTP in Pilot Experiment 2 was higher for the loss contract ($2.58, $N = 30$, SD = $1.97) than for the gain contract ($2.17, $N = 30$, SD = $2.14).[22] Overall, we find no evidence that people prefer GAIN to LOSS.

References

Abdellaoui M, Bleichrodt H, L'Haridon O (2008) A tractable method to measure utility and loss aversion under prospect theory. *J. Risk Uncertainty* 36(3):245–266.

Andersen S, Harrison GW, Lau MI, Rutstrom EE (2008) Eliciting time and risk preferences. *Econometrica* 76(3):583–618.

Anderson LR, DiTraglia FJ, Gerlach JR (2011) Measuring altruism in a public goods experiment: A comparison of U.S. and Czech subjects. *Experiment. Econom.* 14(3):426–437.

Andreoni J, Sprenger C (2012) Risk preferences are not time preferences. *Amer. Econom. Rev.* 102(7):3357–3376.

Arlen J, Tontrup S (2015) Does the endowment effect justify legal intervention? The debiasing effect of institutions. *J. Legal Stud.* 44(1):143–182.

Ashraf N, Karlan D, Yin W (2006) Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *Quart. J. Econom.* 121(1):635–672.

Augenblick N, Niederle M, Sprenger C (2015) Working over time: Dynamic inconsistency in real effort tasks. *Quart. J. Econom.* 130(3):1067–1115.

Baker G, Jensen M, Murphy K (1988) Compensation and incentives: Practice vs. theory. *J. Finance* 43(3):593–616.

Barberis NC (2013) Thirty years of prospect theory in economics: A review and assessment. *J. Econom. Perspect.* 27(1):173–196.

Becker GS, Murphy KM (1988) A theory of rational addiction. *J. Political Econom.* 96(4):675–700.

Brooks RRW, Stremitzer A, Tontrup S (2012) Framing contracts: Why loss framing increases effort. *J. Institutional Theoret. Econom.* 168(1):62–82.

Brown AL, Chua ZE, Camerer CF (2009) Learning and visceral temptation in dynamic saving experiments. *Quart. J. Econom.* 124(1):197–231.

Cadena X, Cristea SA, Delgado-Medrano HM (2011) Fighting procrastination in the workplace: An experiment. NBER Working Paper 16944, National Bureau of Economic Research, Cambridge, MA.

Camerer CF, Loewenstein G, Rabin M, eds. (2004) *Advances in Behavioral Economics* (Princeton University Press, Princeton, NJ).

Charness G, Gneezy U, Imas A (2013) Experimental methods: Eliciting risk preferences. *J. Econom. Behav. Organ.* 87(March):43–51.

de Araujo FA, Carbone E, Conell-Price L, Dunietz MW, Jaroszewicz A, Landsman R, Lamé D, Vesterlund L, Wang S, Wilson AJ (2015) The effect of incentives on real effort: Evidence from the slider task. CESifo Working Paper 5372, CESifo Group Munich, Munich, Germany.

DellaVigna S (2009) Psychology and economics: Evidence from the field. *J. Econom. Literature* 47(2):315–372.

De Quidt J (2014) Your loss is my gain: A recruitment experiment with framed incentives. Working paper, Stockholm University, Stockholm. http://dx.doi.org/10.2139/ssrn.2418218.

DiTraglia FJ (2006) Experimental public goods in Prague and Williamsburg: An international comparison. Working paper, College of William and Mary, Williamsburg, VA.

Ericson KMM, Fuster A (2014) The endowment effect. *Ann. Rev. Econom.* 6(1):555–579.

Fryer RG Jr, Levitt SD, List JA, Sadoff S (2012) Enhancing the efficacy of teacher incentives through loss aversion: A field experiment. NBER Working Paper 18237, National Bureau of Economic Research, Cambridge, MA.

Gill D, Prowse V (2012) A structural analysis of disappointment aversion in a real effort competition. *Amer. Econom. Rev.* 102(1):469–503.

Giné X, Goldberg J, Silverman D, Yang D (2012) Revising commitments: Field evidence on the adjustment of prior choices. NBER Working Paper 18065, National Bureau of Economic Research, Cambridge, MA.

Holt CA, Laury SK (2002) Risk aversion and incentive effects. *Amer. Econom. Rev.* 92(5):1644–1655.

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

Hsiaw A (2013) Goal-setting and self-control. *J. Econom. Theory* 148(2):601–626.

Hsiaw A (2015) Goal-bracketing and self-control. Working paper, Brandeis University, Waltham, MA.

Imas A (2014) Working for the "warm glow": On the benefits and limits of prosocial incentives. *J. Public Econom.* 114(June):14–18.

John LK, Loewenstein G, Troxel AB, Norton L, Fassbender JE, Volpp KG (2011) Financial incentives for extended weight loss: A randomized, controlled trial. *J. General Internal Medicine* 26(6):621–626.

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–292.

Kahneman D, Knetsch JL, Thaler RH (1990) Experimental tests of the endowment effect and the Coase theorem. *J. Political Econom.* 98(6):1325–1348.

Kaur S, Kremer M, Mullainathan S (2010) Self-control and the development of work arrangements. *Amer. Econom. Rev. Papers Proc.* 100(2):624–628.

Kaur S, Kremer M, Mullainathan S (2015) Self-control at work. *J. Political Econom.* 123(6):1227–1277.

Kőszegi B, Rabin M (2006) A model of reference-dependent preferences. *Quart. J. Econom.* 121(4):1133–1165.

Laibson D (1997) Golden eggs and hyperbolic discounting. *Quart. J. Econom.* 112(2):443–478.

Lazear EP (1991) Labor economics and the psychology of organizations. *J. Econom. Perspect.* 5(2):89–110.

Levitt SD, List JA, Neckermann S, Sadoff S (2016) The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *Amer. Econom. J.: Econom. Policy*. Forthcoming.

List JA, Samek AS (2014) The behavioralist as nutritionist: Leveraging behavioral economics to improve child food choice and consumption. *J. Health Econom.* 39(January):133–146.

Loewenstein G, Adler D (1995) A bias in the prediction of tastes. *Econom. J.* 105(431):929–937.

Loewenstein G, O'Donoghue T, Rabin M (2003) Projection bias in predicting future utility. *Quart. J. Econom.* 118(4):1209–1248.

Luft J (1994) Bonus and penalty incentives contract choice by employees. *J. Accounting Econom.* 18(2):181–206.

Madrian BC (2014) Applying insights from behavioral economics to policy design. *Ann. Rev. Econom.* 6(1):663–688.

McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306(5695):503–507.

O'Donoghue T, Rabin M (1999) Doing it now or later. *Amer. Econom. Rev.* 89(1):103–124.

Royer H, Stehr MF, Sydnor JR (2015) Incentives, commitments and habit formation in exercise: Evidence from a field experiment with workers at a Fortune-500 company. *Amer. Econom. J.: Appl. Econom.* 7(3):51–84.

[22] One participant in the gain treatment reported inconsistent WTP across the multiple price list. The results reported above use the subject's first switching point (i.e., lowest WTP). Dropping the participant from the analysis decreases the average WTP in the gain treatment to $2.14.

Sadoff S, Samek AS, Sprenger C (2015) Dynamic inconsistency in food choice: Experimental evidence from a food desert. Becker Friedman Institute for Research in Economics Working Paper 2572821, University of Chicago, Chicago.

Samuelson W, Zeckhauser R (1988) Status quo bias in decision making. *J. Risk Uncertainty* 1(1):7–59.

Schmid F, Trede M (1995) A distribution free test for the two sample problem for general alternatives. *Comput. Statist. Data Anal.* 20(4):409–419.

Schwartz J, Mochon D, Wyper L, Maroba J, Patel D, Ariely D (2014) Healthier by precommitment. *Psych. Sci.* 25(2):538–546.

Solnick JV, Kannenberg CH, Eckerman DE, Waller MB (1980) An experimental analysis of impulsivity and impulse control in humans. *Learn. Motivation* 11(1):61–77.

Sprenger C (2015) An endowment effect for risk: Experimental tests of stochastic reference points. *J. Political Econom.* 123(6): 1456–1499.

Thaler RH, Johnson EJ (1990) Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Sci.* 36(6):643–660.

Tversky A, Kahneman D (1992) Advances in prospect theory: Cumulative representation of uncertainty. *J. Risk Uncertainty* 5(4):297–323.

Van Boven L, Dunning D, Loewenstein G (2000) Egocentric empathy gaps between owners and buyers: Misperceptions of the endowment effect. *J. Personality Soc. Psych.* 79(1):66–76.