The Dark Side of Monetary Bonuses*

Victor Gonzalez-Jimenez[a], Patricio S. Dalton[b], Charles N. Noussair[c]

[a]*Erasmus School of Economics, Erasmus University Rotterdam*
[b]*Department of Economics, Tilburg University*
[c]*Department of Economics, University of Arizona*

Abstract

To incentivize workers, firms use bonuses tied to the achievement of production goals. These goals are often set by the workers themselves. We develop a theoretical framework that predicts that if workers have high loss aversion, monetary rewards for the achievement of self-chosen goals make workers set conservative goals and worsen performance. Conversely, if goal achievement is not rewarded monetarily, workers with high loss aversion set more ambitious goals, which in turn improves performance. Results from a laboratory experiment corroborate this prediction. The results highlight the limits of monetary bonuses as an effective incentive when workers set production goals and are loss averse.

JEL Codes: J41, D90, C91, D81
Keywords: contracts, goal setting, incentives, loss aversion, experiment

* This paper has benefited from insightful comments by Alexander Koch and Joaquin Gomez-Miñambres. We also thank Christine Eckert, Juha Tolvanen, Pol Campos-Mercade, participants at the 2020 Experimental Finance conference, the 2019 M-BEES conference, the 2019 Bogota Experimental Economics Conference, and the 2019 Young Economists Meeting for their comments. The experiments reported in this paper were funded by the University of Vienna and the Economic Science Laboratory of the University of Arizona.

1. Introduction

Offering monetary bonuses for the achievement of goals is a widespread practice used by firms to incentivize employees. According to Worldatwork (2018), close to 98% of publicly traded American companies use at least one compensation scheme that includes bonuses, and 73% of these companies report that these bonuses are awarded when an individual or organizational goal is reached. The main theoretical rationale behind including bonuses in compensation packages is that the additional monetary incentive that the bonus creates boosts performance (Gibbons and Roberts, 2013).

There is also ample evidence from psychology showing that a goal that is challenging but attainable can lead to greater effort exertion even if the goal is not monetarily incentivized (Heath et al.,1999; Wu et al., 2008). Moreover, instead of exogenously imposing goals on employees, firms are increasingly involving employees in the decision to set production goals (Gallo 2011, Bourne et al., 2013, Groen et al., 2015, de Morree, 2018).[1] Workers who set their own goals feel that they have more control over outcomes, and that they are more involved with the decisions of the firm (Groen et al., 2012, 2015).[2] Indeed, recent experimental evidence confirms that contracts with self-chosen goals are more cost-effective than contracts with exogenously set goals (Groen et al., 2015, Brookins et al., 2017).

In this paper, we ask whether and under what conditions the attainment of a self-chosen goal should be rewarded with a monetary bonus. Throughout, we refer to a bonus as a discrete jump in the agent's compensation triggered when a production goal is met (Oyer, 2000, Kim 1997, Park 1995). On one hand, offering monetary bonuses tied to the achievement of a goal can incentivize the setting of challenging goals, provided that more ambitious goals are accompanied by higher bonuses. On the other hand, a monetary bonus can be counterproductive if it crowds out the intrinsic motivation from setting a goal, as has been observed in other contexts (Gneezy and Rustichini, 2000, Ariely et al., 2009a, Ariely et al., 2009b, Gneezy et al., 2011). If the intrinsic motivation to achieve a goal is sufficiently strong, monetary bonuses for goal achievement may not provide any additional incentive and may even crowd out the motivation that the presence of a goal creates. If so, money spent on the bonuses would be a wasteful expenditure for an employer.

We develop a theoretical model, presented in Section 2, which shows that when workers have reference-dependent preferences and set their own goals, attaching a bonus to the goal can lower performance. This is because a loss averse worker will choose to set more conservative goals to increase her chance of obtaining the bonus. This behavior leads to lower performance, since the motivation from not falling short from the goal is not as strong as it would have been if the goal was higher. Consequently, workers with high loss aversion, that is, who are more sensitive to the losses associated with not attaining the bonus and

[1] Setting one's own goal is also common outside the workplace. Fitness apps allow the user to set her own target for minutes exercised or steps walked each day. Many banks offer a calculator in which the account holder sets a savings goal and the amount that has to be saved each month is calculated for them. Apps such as Goals-on-Track and Lifetick permit the user to set their goals in a variety of areas and to track their progress.

[2] There are at least two theoretical rationales for the effectiveness of goals. First, a goal can act as a costly self-commitment device (Koch and Nafziger, 2011, 2019, 2020; Hsiaw, 2013; Kaur, 2015) for workers with present-biased preferences. Second, a goal attains the status of a reference point, making the loss averse individual exert great effort to avoid experiencing the losses in utility that would result from falling short of the goal (Heath et al., 1999; Wu et al., 2008).

who are more motivated by goals, will set lower goals when goals are rewarded monetarily and, as a consequence, exhibit worse performance.

We conduct a laboratory experiment to test the predictions of our model. In the experiment, described in Section 3, participants must complete a task that requires effort. The monetary incentives to perform the task depend on the treatment they are randomly assigned to. The design of the experiment can be viewed from the perspective of a firm that is considering changes to a simple piece-rate contract. The first treatment, called LOPR, is a low-powered piece-rate contract. The treatment GOAL+BONUS is a contract in which LOPR is complemented with a self-chosen goal that yields a monetary bonus if the goal is achieved. Comparing GOAL+BONUS to LOPR provides a measure of whether there is sufficient improvement in performance from adding a self-chosen goal to more than offset the bonuses that are paid. The GOAL treatment adds to LOPR a self-chosen goal without any monetary bonus. Comparing GOAL+BONUS and GOAL is key to our paper and indicates how much more (or less) performance one can get from the bonus. The fourth and last treatment is a higher-powered piece-rate contract called HIPR. Comparing the HIPR and GOAL (or GOAL+BONUS) provides a comparison of performance under self-chosen goals vis-à-vis a substantial piece rate increase.

Our model predicts that if individuals exhibit high loss aversion, GOAL yields higher performance and higher goals than GOAL+BONUS. If this is the case, GOAL dominates GOAL + BONUS from the point of view of the employer, because it involves lower employer expenditure for higher output. Since our theoretical predictions depend on the extent of individuals' loss aversion, we elicit the participants' loss aversion. We implement the parameter-free elicitation method developed by Abdellaoui et al. (2008). This elicitation method has the advantage that it allows the independent measurement of the curvature of participants' utility functions and their degree of loss-aversion, while accounting for the possibility that participants might exhibit probability weighting (Tversky and Kahneman, 1992, Gonzalez and Wu, 1999, Abdellaoui, 2000).

The experimental data, reported in Section 4, confirm the main predictions of the model. On average, participants assigned to the GOAL treatment set more ambitious goals than those assigned to GOAL+BONUS, and exhibit higher performance than participants assigned to any of the other three contracts. Specifically, a non-paid goal contract leads to 11% more output, an increase in performance of 0.36 standard deviations, over a paid goal contract. Moreover, as predicted by the model, we find that the performance of loss averse participants is especially greater when goals are not rewarded monetarily. Finally, we observe that performance under goals with no payment is significantly higher than performance under the two piece-rate contracts. We conclude that a goal contract with no monetary bonus is the cheapest way to improve performance among the contracts we study.

This paper contributes to several strands of literature. It adds to the literature on incentives and contracting (e.g. Laffont and Tirole, 1993; Laffont and Martimort, 2002). We show that offering monetary bonuses can be counterproductive when they are linked to a production goal that is set by loss-averse workers. This result constitutes a proof of principle that offering additional monetary incentives could inhibit psychological motives that would otherwise stimulate effort. Our study also adds to the recent and expanding area of behavioral contract theory (see Köszegi (2014) for a review). In contrast to the results of De Meza and Webb (2007) and Herweg et al. (2010), we show that the principal may obtain lower performance when he offers an incentive scheme that includes a performance bonus to loss averse agents. The results of our work show that the principal should instead offer a contract that includes non-monetarily rewarded self-chosen goals to properly harness the worker's loss aversion.

Our paper also contributes to the emerging literature on goal setting in economics (Koch and Nafziger, 2011, 2019, 2020; Gómez-Miñambres, 2012; Corgnet et al., 2015, 2018; Kaur et al., 2015; Allen et al., 2017; Brookins et al., 2017; Markle et al., 2018). To our knowledge, this is the first paper studying theoretically and empirically how monetary bonuses interact with self-chosen goals and the role of loss aversion in moderating such interaction. We depart from the existing literature in two fundamental ways. We do not assume dynamic inconsistency as in Hsiaw (2013), Kaur et al. (2015), Hsiaw (2018) and Koch and Nafziger (2011, 2019). We also relax the assumption of deterministic performance, as assumed by Wu et al. (2008), Corgnet et al. (2015), Dalton et al. (2016a, 2016b), Brookins et al. (2017). Including uncertainty about reaching a self-chosen goal yields the novel prediction that adding monetary bonuses for the achievement of a goal can lower performance. This paper also contributes to this literature, as it is, to our knowledge, the first to quantify loss aversion and utility curvature to study their association with goal setting, monetary incentives, and performance. Eliciting these preference parameters also allows us to validate empirically the mechanism of the model.

Within the literature of goal setting, our study contributes to the literature on self-chosen goals. Kaur et al. (2015) show that self-chosen goals can act as a costly self-commitment device. In their experiment, workers were offered the option to choose a production goal. If they achieved the goal, they received a piece rate; if they fell short of it, they received only half of such piece rate for their output. We depart from Kaur et al (2015) in several ways. First, our focus is on loss aversion as a motivating force behind goal setting beyond time-inconsistent preferences. Moreover, in our goal contract, falling short of the self-chosen goal does not inflict any monetary penalty. On the contrary, setting a positive goal can only improve the expected payment from performance. In another related paper, Brookins et al. (2017) show evidence from a field experiment that self-chosen goals without a bonus improve performance compared to a cost-equivalent piece rate. In contrast to their work, our focus is on the effect of bonuses paid for goal achievement. Such monetary incentives are not considered in their study. We also differ from Brookins et al. (2017) in that we focus on the role of loss aversion and its interaction with bonuses and for that reason, we offer a model that predicts a crowding-out effect of motivation from bonuses, and we measure loss aversion at the individual level.

Finally, we add to the literature that examines how extrinsic incentives crowd-out intrinsic motivation (see Gneezy et al., 2011, and Bowles and Polanía-Reyes, 2012, for reviews and Benabou and Tirole, 2003, for a theoretical framework). In this literature, crowding-out effects appear when monetary incentives inhibit individuals from signaling to others, or to themselves, a favorable attribute such as intelligence (Ariely, et al., 2009b), pro-sociality (Ariely et al., 2009a, Mellstrom and Johannesson, 2016) or norm-conforming behavior (Frey and Oberholzer-Gee, 1997, Gneezy and Rustichini, 2000). Our main contribution to this literature is to provide a completely different microeconomic foundation, based on loss aversion, for the motivational crowding out effect. Specifically, the intrinsic incentives from setting an ambitious goal, which stem from loss aversion, are offset by monetary incentives that encourage more modest goal setting.

2. Theoretical Framework

Consider a worker hired by a principal to produce output $y \in [0, \overline{y}]$ on a task. The agent's action consists of exerting an effort level $e \in \{e_L, e_H\}$. Exerting effort implies incurring a cost $c(e)$, which is higher under high effort, e_H, than under low effort, e_L. For simplicity, we assume the following piece-wise function for $c(e)$:

Assumption 1. $c(e) = \begin{cases} c \; if \; e_H, \\ 0 \; if \; e_L, \end{cases}$

where $c > 0$. In addition to the agent's effort, other factors also affect production. We thus model y as a random variable conditional on effort. Both parties, principal and agent, know that y is distributed according to the cumulative distribution function $F(y|e)$, which has a probability density function $f(y|e)$. To keep the problem tractable, we assume that the mean output produced is positive and bounded for any effort level, $0 < \mathbb{E}(y|e) < \infty$. Finally, we assume that higher effort boosts production in a manner exhibiting the monotone likelihood ratio property:

Assumption 2. $\frac{\partial}{\partial y}\left(\frac{f(y|e_H)}{f(y|e_L)}\right) \geq 0, \forall \; y \in [0, \overline{y}]$.

An immediate consequence of Assumption 2 is the existence of a unique output level $\hat{y} \in (0, \overline{y})$ such that $f(\hat{y}|e_H) = f(\hat{y}|e_L)$. For $y \in [0, \hat{y})$ then $f(y|e_H) < f(y|e_L)$, likelihoods are higher under low effort, while for $y \in (\hat{y}, \overline{y}]$ the opposite is true. We take advantage of this partition and sometimes informally refer to $y \in [0, \hat{y})$ as low output levels and $y \in (\hat{y}, \overline{y}]$ as high output levels.

The principal is assumed to be risk neutral. Specifically, her objective function is:

$$\int_{\underline{y}}^{\overline{y}} \big(S(y) - w(y)\big) f(y|e) dy,$$

Where $S(y)$ captures benefit and exhibits the properties, $S'(y) > 0$ and $S''(y) < 0$, and $w(y)$ is the wage offered to the agent. From the above equation is evident that the objective of the principal is to elicit high effort as it extracts higher output (Assumption 2).

We focus on a setting in which the agent is incentivized with a goal contract, w_g. That contract consists of a piece rate $a > 0$, i.e., a monetary amount given in exchange of each additional unit of output produced, and a bonus $B(y, g)$. The bonus pays a monetary amount if the agent meets or exceeds a goal, $g \in [0, \overline{y}]$. Specifically, the payoff of the agent is:

$$w_g(y, B(y, g)) := ay + B(y, g), \qquad (1)$$

where

$$B(y, g) = \begin{cases} 0 \; if \; \mathrm{y} < g, \\ bg \; if \; \mathrm{y} \; \geq g, \end{cases} \qquad (2)$$

and $b > 0$. That is, the agent receives a larger bonus for achieving more ambitious goals, and the bonus is not awarded if the goal is not attained. According to Chung et al. (2014), this type of contract is classified as a combination of linear commission, i.e., the piece rate, and a bonus. See their Figure 1d.[3]

In this paper, we assume that the goal, g, is chosen by the agent. A key advantage of self-chosen goals, as opposed to exogenously set goals, is that the agent is more likely than the principal to know her own cost, c, and can set a goal that is tailored to this parameter. Self-chosen goals are also useful in settings where there is potential heterogeneity regarding workers' ability in the task, as reflected by the parameter c , but it is not possible for the employer to impose different contracts on different workers. In such a setting, the trade-off under the contract described above is clear: the agent will want to set a goal that is high enough so that she can earn a higher bonus, but not so high as to be unachievable.

The timing of the contract is as follows. First, the agent simultaneously decides g and e. Then, the level of output y is realized, and the agent receives the benefits corresponding to y as well as the bonus, if the goal is achieved. Since we assume that the agent has time-consistent preferences, the fact that goals and effort are simultaneously chosen is equivalent to considering a framework in which goals are chosen first and the choice of effort follows.[4]

Further, we assume that the worker has reference-dependent preferences. These preferences capture the notion that the agent does not only derive utility from the monetary incentives offered by the goal contract, but also that the presence of a goal induces a psychological (dis)utility from (not) achieving it. We assume that the goal acquires the status of a reference point (Heath et al., 1999, Wu et al., 2008, Markle et al., 2018). Hence, a production goal induces an intrinsic, non-monetary, psychological utility that satisfies the properties of Kahneman and Tversky's (1979) value function.

Assumption 3. *The value function is given by* $v(y,g) = \begin{cases} \mu(y-g) & if\ y \geq g \\ -\mu\lambda(g-y) & if\ \ y < g \end{cases}$, *with* $\mu \geq 0$ *and* $\lambda > 1$.

The parameter $\lambda > 1$ reflects the agent's loss aversion, i.e., the psychological loss from failing short of a goal by some amount looms larger than the gain from surpassing the goal by the same amount. The parameter $\mu \geq 0$ represents the weight of psychological utility on the overall utility. If $\mu = 0$, the agent's preferences are *standard* in the sense that no utility is derived from achieving or failing to achieve the goal, other than from the monetary payment that it yields.

The expected utility of an agent with reference-dependent preferences facing a self-chosen goal contract is:

---

[3] Note that this contract is different than that of Kaur et al. (2015). Let $a = \frac{b}{2}$. Our contract becomes $w_g = \begin{cases} \frac{b}{2}\ if\ \mathrm{y} < g, \\ \frac{b}{2} + bg\ if\ \mathrm{y}\ \geq g \end{cases}$.
Their contract is $w = \begin{cases} \frac{b}{2}\ \ if\ \mathrm{y} < g, \\ b\ if\ \mathrm{y}\ \geq g \end{cases}$. Setting a positive goal is, on expectation, dominated under Kaur's et al (2015) contract but not under ours.

[4] Formally, the agent can be thought as having exponential discounting with discount factor $\delta = 1$. See Eq. (3).

$$\mathbb{E}\left(U\left(w_g,e,g\right)\right)=\int_0^{\bar{y}} ay\, f(y|e)dy+\int_g^{\bar{y}} bg+\mu(y-g)\, f(y|e)dy-\int_0^{g} \lambda\mu(g-y)\, f(y|e)dy\ -c(e) \tag{3}$$

The first term in (3) is the expected monetary utility from the piece-rate payment. The second term is the expected utility from producing y above the goal. This term includes both monetary (if $b>0$) and psychological utility gains. The third term is the expected psychological disutility from producing y below the goal. This representation of preference with the goal as the reference point is widely adopted in the literature on goal setting (Koch and Nafziger 2011, Corgnet et al., 2015, 2018, Dalton et al., 2016a,2016b, Brookins et al. 2017).

After solving the integrals using integration by parts, equation (3) becomes:

$$\mathbb{E}\left(U\left(w_g,e,g\right)\right)=(a+\mu)\mathbb{E}(y|e)-c(e)+bg\big(1-F(g|e)\big)-\mu g-\mu(\lambda-1)\int_0^{g} F(y|e)dy. \tag{4}$$

When the agent sets goals herself, she chooses the goal that maximizes (4). The following lemma characterizes the optimal goal level as a function of effort, $g(e)$. This function can be thought as a best-response function of goals for any level of effort.

Lemma 1. *The optimal goal for a given effort,* $g(e)$*, satisfies the following condition:*

$$g(e)=\frac{b\big(1-F(g(e)|e)\big)-\mu\big(1+(\lambda-1)F(g(e)|e)\big)}{bf(g(e)|e)}.$$

This function is non-decreasing in the bonus, $\frac{dg(e)}{db}\geq 0$*, decreasing in loss-aversion,* $\frac{dg(e)}{d\lambda}<0$*, and independent of the piece-rate,* $\frac{dg(e)}{da}=0$*, and the cost of effort,* $\frac{dg(e)}{dc}=0$.

Lemma 1 shows that when the effort level is taken as given, higher bonuses motivate higher goals, and higher loss aversion decreases goals.

Consider now the agent's decision to exert effort for a given goal level. Using (4), we can establish that the agent chooses high effort if the following incentive compatibility constraint holds:

$$\begin{aligned} IC(g):\ \ &(a+\mu)\big(\mathbb{E}(y|e_H)-\mathbb{E}(y|e_L)\big)+bg\big(F(g|e_L)-F(g|e_H)\big)\\ &+\mu(\lambda-1)\int_0^{g} F(y|e_L)-F(y|e_H)dy\geq c. \end{aligned} \tag{5}$$

Note that the magnitude of c determines how affordable high effort is. From (5), it is evident that a higher c makes choosing high effort less likely. We think of this parameter as a measure of productivity of an agent.

The following lemma characterizes the decision to exert high effort as a function of a given goal. Again, this function can be thought as a best-response function of the agent for a given goal level.

Lemma 2. *For a given goal, the decision to exert high effort is non-decreasing in the bonus,* $\frac{dIC(g)}{db} \geq 0$, *loss-aversion,* $\frac{dIC(g)}{d\lambda} \geq 0$, *and the piece-rate,* $\frac{dIC(g)}{da} \geq 0$, *while it is decreasing with the cost of effort,* $\frac{dIC(g)}{dc} < 0$.

Lemma 2 shows that for a given goal, choosing high effort is more likely when an agent is more loss averse, or when the bonus or the piece-rate are higher.

Thus far, Lemma 1 and Lemma 2 point at bonuses as motivational devices: they boost effort (for a given goal) as well as goals (for a given effort). However, the way effort responds to goals is crucial to understanding why bonuses can backfire. This relationship depends on the bonus level in a non-trivial way. The following Lemma characterizes this dependence.

Lemma 3. *The decision to exert high effort depends on the goal level in the following way:*

i. *If* $b = 0$, *then* $\frac{dIC(g)}{dg} \geq 0$;

ii. *If* $b > 0$ *and* $g < \hat{y}$, *then* $\frac{dIC(g)}{dg} \geq 0$;

iii. *Let* $\hat{\lambda} := 1 + \frac{b}{\mu}\left(\frac{g\left(f(g|e_H) - f(g|e_L)\right)}{F(g|e_L) - F(g|e_H)} - 1\right)$. *If* $b > 0$, $\lambda > \hat{\lambda}$, *and* $g > \hat{y}$, *then* $\frac{dIC(g)}{dg} \geq 0$.

The key insight of Lemma 3 is that high effort increases with higher goals when there is no bonus associated with reaching goals. However, this is not necessarily true in the presence of a bonus. In that case, higher goals motivate high effort either if goals are not high enough, $g \in [0, \hat{y})$, or if goals are high, $g \in (\hat{y}, \overline{y}]$ *and* the agent's loss aversion is high enough (higher than the threshold $\hat{\lambda}$). Note that the threshold $\hat{\lambda}$ increases with the size of the bonus, so the probability that high goals motivate high effort decreases with higher bonuses.[5]

We are now in condition to define an optimal solution in this model. We assume that the principal anticipates $g(e)$, the optimal goal function of the agent for a given effort level (see Lemma 1) and offers a contract w_g that implements agent's high effort. We are interested in the properties of that contract. In the optimum, the agent chooses a goal consistent with high effort. The following proposition characterizes the solution of this model and shows that when the agent is sufficiently loss-averse, higher bonuses imply lower goals.

Proposition. *In the optimal solution, an agent with* $\mu > 0$ *and loss aversion level* $\lambda \geq \hat{\lambda}$ *sets a positive goal that is optimal under high effort,* $g^*(e_H) > 0$. *That optimal goal is decreasing in the bonus,* $\frac{\partial g^*(e_H)}{\partial b} < 0$. *Therefore, a higher bonus decreases the values c for which a solution with high effort can be implemented.*

[5] Lemma 4 in the Appendix show that a similar conclusion is achieved when optimal goals consistent with high effort, $g(e_H)$, are compared to optimal goals consistent with low effort, $g(e_L)$.

The Proposition shows that offering a monetary bonus for the achievement of a self-chosen goal will backfire if the agent is sufficiently loss averse. That is because the probability of losing the monetary bonus if the goal is not reached will induce a loss averse agent to set lower goals compared to a situation with no bonus. A lower goal leads to lower motivation to exert high effort. In such a case, the principal will achieve higher output at lower cost with a contract without a bonus than with one with bonus.

The intuition of the solution presented in the Proposition is as follows. When a bonus is offered, higher goals motivate high effort for an individual whose loss aversion is greater than $\hat{\lambda}$ (Lemma 3 (iii)). Since the loss aversion threshold $\hat{\lambda}$ depends positively on the bonus, higher levels of loss aversion are required for goals to be motivating under higher bonuses. However, higher levels of loss aversion also lead to lower goal setting (Lemma 1), and these lower goals in turn demotivate the choice of high effort. This demotivational effect counteracts the positive effect of higher loss aversion on effort (Lemma 2). Alternatively, the agent could set a low goal, $g^*(e_H) < \hat{y}$ (Lemma 3 (ii)), but doing so also leads to lower motivation than when no bonus is awarded.

We turn to the special case of $b = 0$, a goal with no monetary bonus. Our proposition shows that higher bonuses generate lower goals and thus lower motivation. Hence, the goal levels and the motivation to exert high effort must be greatest in the absence of a bonus. The following corollary confirms that intuition.

Corollary 1. *Consider an agent with $\mu > 0$ working under a self-chosen goal contract with $b = 0$. The agent sets the highest possible goal $g^*(e_H) = \bar{y}$ and exerts high effort for the largest possible range of costs c.*

When goals are not combined with a monetary bonus for the attainment of the goal, higher goals always incentivize high effort (Lemma 3 (i)). This is because, under $b = 0$, not achieving a goal is costless in the money dimension, while the motivation that the agent derives from not incurring psychological losses remains. Since higher goals lead to larger psychological losses, more challenging goals generate greater motivation.

This result seems to contradict Lemma 1, in which the optimal goal, conditional on effort approaches zero when $b = 0$. However, recall that the solution that is characterized implements high effort and, contrary to Lemma 1, does not consider goals for any effort level.[6]

To conclude this section, we investigate the effect of bonuses when the agent does not exhibit reference-dependent preferences. The following Corollary presents the result.

Corollary 2. *Consider an agent with $\mu = 0$ working under a self-chosen goal contract. As the bonus b increases, the agent exerts high effort for a greater range of costs, c.*

Under standard preferences ($\mu = 0$), higher bonuses motivate high effort and never backfire. When there is no psychological cost from not achieving the goal, the agent with standard preferences sets a goal and exerts high effort to maximize the chances of achieving it.

[6] One way to understand this result uses Lemma 4, in the Appendix. It states that $b = 0$, $\mu > 0$, and high loss aversion imply $g(e_H) > g(e_L)$. Thus, it cannot be that $g(e_H) = 0$. Let $g(e_L) = 0$ which is in turn consistent with Lemma 1. Then, according to Lemma 3 (i) it must be that $g(e_H)$ is set at the maximum.

2.2 Hypotheses

The model yields a set of predictions that we test with a laboratory experiment. When formulating these predictions, it is assumed, as it is typically found in laboratory experiments, that individuals are loss averse (Tversky and Kahneman 1992, Abdellaoui et al., 2007, 2008, 2016). Regarding performance, the main hypothesis is derived from our Proposition and Corollary 1.

Hypothesis 1*: Participants will exhibit higher performance under GOAL than under GOAL+BONUS and this difference will be larger for participants with relatively high levels of loss aversion.*

If we instead find that GOAL+BONUS leads to higher performance as compared to GOAL and subjects are not loss averse, the data will also support the model. In this case Corollary 2 will apply and successfully describe the data.

Two hypotheses predicting goal-setting behavior emerge from the Proposition and Corollary 1.

Hypothesis 2: *Participants will set higher goals in GOAL than in GOAL+BONUS.*

Hypothesis 3: *The difference in goal levels between GOAL and GOAL+BONUS will be larger for participants with high levels of loss aversion.*

We compare the performance of the goal contract with and without bonus with that generated by a pure piece rate contract. While a piece rate contract is optimal under rather stringent conditions (Holmström and Milgrom, 1987), it has a particularly simple structure and is frequently observed in organizations (Chiappori and Salanie, 2000). It is often discussed as a benchmark contract structure in the theory of incentives (Laffont and Martimort, 2002).

The contract offered in *LOPR* provides the same piece-rate as in GOAL but it does not require subjects to set an explicit goal. As a result, the extra motivation from not achieving the goal is not experienced by subjects. We hypothesize that such motivation improves performance.

Hypothesis 4*: Participants will exhibit higher performance in GOAL than in LOPR.*

3. Experimental Procedures

3.1 General Procedures

The experiment was conducted at the University of Arizona's Economic Science Laboratory in May 2018. Participants were all students at the university and were recruited using an electronic system. The dataset consists of 12 sessions with a total of 161 participants. On average, a session lasted approximately 70 minutes. Between 3 and 20 participants took part in each session. The currency used in the experiment was US Dollars. We used Otree (Chen, et al., 2016) to implement and run the experiment. Participants earned on average 20.7 US Dollars. The instructions of the experiment are provided in Appendix D.

The experiment consisted of two parts: A and B. Upon arrival, participants were informed that their earnings from either Part A *or* Part B would be their earnings for the session, and that this would be decided by chance at the end of the session. Whether participants faced Part A or Part B first was determined at random by the computer.

3.2. Treatment Structure: Comparison of the Contracts

In Part A, participants performed a task that required their effort and attention. The task consisted of counting the number of zeros in a table of 100 randomly distributed zeros and ones. This task has been widely used by other researchers (e.g. Abeler et al., 2011, Gneezy et al., 2017, and Koch and Nafzinger, 2019). Participants submitted their answers using the computer interface. Immediately after submission, a new table appeared on the computer screen and participants were invited again to count the number of zeros in the new table.

Participants had six rounds of five minutes each to complete as many correct tables as they could. To become acquainted with the task, participants also had a five-minute practice round where it was clear that their performance did not count toward their earnings. After each round ended, participants were given feedback about the number of tables they solved correctly and their earnings for that round. If applicable, they were reminded of their goal for that round and were told whether that goal was achieved. In total, aside from the practice round, participants had 30 minutes to work on the task, and were given small time intervals between rounds. Since the time between goal setting and performance was almost immediate, we rule out by design any self-control problems, which have been shown to affect goal-setting behavior.

There were four treatments, LOPR, HIPR, GOAL, and GOAL+BONUS. The treatments differed only with respect to the incentives offered to participants. Each participant was randomly assigned to one of the four treatments. We ensured randomization in our design by having participants in any given experimental session face the same chance of being assigned to any of the treatments. That is, within each session in the laboratory, different individuals were randomized into different treatments. The incentives in effect in each treatment were the following.

- *LOPR*: Participants were paid 0.20 dollars for each correctly solved table.
- *HIPR*: Participants were paid 0.50 dollars for each correctly solved table.
- *GOAL*: Participants were paid 0.20 dollars for each correctly solved table and were asked at the beginning of each round to set an individual goal regarding the number of tables that they aimed to solve in that round.
- *GOAL+BONUS:* Identical to the *GOAL* treatment, with the exception that participants were offered a monetary bonus for reaching their goals. The bonus in dollars was equivalent to the goal set by the participant, multiplied by a factor of 0.20.

LOPR can be viewed as a baseline condition to which different features that may improve performance are added. HIPR includes an increase in the piece rate, while GOAL adds a goal set by the worker. GOAL+BONUS adds the goal, as well as a monetary payment for reaching it, which is larger the more ambitious the goal.

Including the HIPR treatment has the purpose of evaluating how incentive schemes with self-chosen goals compare in terms of performance to a policy of considerably increasing the piece rate. The predicted boost in performance in GOAL from Corollary 1 is not only compared to a cost-equivalent piece rate LOPR but also to a higher piece-rate, HIPR.

3.3 Elicitation of Risk Attitudes

In Part B of the experiment, the task was to choose between two binary lotteries in multiple trials. This part of the experiment was designed to elicit participants' loss aversion and utility curvature. The lotteries yielded either only gains, or were mixed in the sense that either gains or losses were possible. We used the Abdellaoui et al. (2008) method, which has the advantage of eliciting risk and loss attitudes without making any assumptions about the decision model that participants use to evaluate outcomes or probabilities.

Our implementation of Abdellaoui et al.'s (2008) method consisted of 10 decision sets. Each decision set was designed to elicit indifference between two initial lotteries through bisection. The algorithm was programmed so that the participant's choice between two initial lotteries determined the next choice problem that the participant faced. Specifically, in the next choice trial, either the lottery chosen in the preceding trial was replaced by a less attractive alternative, or the one not chosen was replaced by a more attractive alternative, while the other choice remained the same. The participant was again invited to choose between the two available options. This process was repeated four times.

Decision sets 1 to 5 elicited participants' utility curvature. In each decision set, the algorithm elicited the certainty equivalent x_j of a lottery of the form $Lottery_j = (H_j, 0.5; L_j, 0.5)$, with $j = 1,2,3,4,5$, and $H_j \geq L_j \geq 0$. The values of H_j and L_j used in each decision set are shown in Table 1 below.

Table 1. High and Low Values Used in Lotteries to Measure Utility Curvature

Lottery	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
H_j	4	8	12	20	20
L_j	0	0	0	0	12

Panel A of Table 2 below details an example of the bisection algorithm used to find x_1, the certainty equivalent of L_1. Note that initially, option R is a degenerate lottery that pays the expected value of option L, which, in turn, is equal to L_1. The example shows that after having made a first choice, the participant faces a new problem whereby R, the option that was chosen before, becomes less attractive. In the remaining repetitions, the individual's preferred option is L, even though lottery R becomes more attractive. The certainty equivalent is eventually determined as x_2 =1.625, the midpoint between 1.75 and 1.5.

Decision sets 6 to 10 elicited participants' loss aversion. The program was designed to find the outcome $z_j < 0$ that made an individual indifferent between a sure outcome of zero and a mixed lottery of the form $(k_j, 0.5; z_j, 0.5)$, with $k_j > 0$ for $j = 1,2,3,4,5$. A loss averse participant would require low values of z_j to be indifferent, whereas a gain-seeking participant would require large values of z_j to be indifferent. The starting values of the program were set at the certainty equivalent of a decision set j, i.e. $k_j = x_i$, and its mirror image, that is $z_j = -x_j$.

Table 2. Example of the Elicitation Procedure for Certainty Equivalents and Loss Aversion

	Panel A			**Panel B**		
Repetition	Lottery L	Lottery R	Choice	Lottery L	Lottery R	Choice
Initial lottery	(4,0.5;0,0.5)	2	**R**	(1.62,0.5; -1.62,0.5)	0	**R**
1	(4,0.5;0,0.5)	1	**L**	(1.62,0.5; -0.81,0.5)	0	**L**
2	(4,0.5;0,0.5)	1.5	**L**	(1.62,0.5; -1.20,0.5)	0	**L**
3	(4,0.5;0,0.5)	1.75	**L**	(1.62,0.5; -1.40,0.5)	0	**R**
Final Elicitation		$\boldsymbol{x_1}$ **=1.625**		$\boldsymbol{z_1}$ **=-1.40**		

Note: This table presents an example of Abdellaoui et al.'s (2008) algorithm used to find certainty equivalents $\{x_1, x_2, x_3, x_4, x_5\}$ and the sequence of offsetting negative numbers $\{z_1 z_2, z_3, z_4, z_5\}$. The left panel presents how x_1 is elicited with the algorithm. The right panel shows how z_1 is elicited.

Panel B of Table 2 presents an example of the bisection algorithm used to find these negative outcomes. Note that in this example, the certainty equivalent elicited in Panel A, x_1 =1.625, is used as an outcome of the mixed lottery. Also note that the mirror image of x_1, -1.625, is used initially as the other outcome of the mixed lottery. The elicited value in this example, z_1 =-1.40, was the value that made the participant indifferent between the lottery $(1.625_j, 0.5; -1.40, 0.5)$ and zero.

Once participants finished both parts (A and B), they were reminded about their performance in each round of the real-effort task, as well as whether they achieved their goal in that round, if applicable. Also, participants were informed about the lottery that was chosen for potential compensation for Part B and its realization. They were also informed about whether Part A or B was chosen to become their final earnings. Finally, participants completed a questionnaire about their general willingness to take risks, as well as to take specific risks (health-, job-, and driving-related). The questions were taken from Dohmen et al. (2011). The questionnaire can be found in Appendix D.

4. Results

We first present results on risk attitudes. Measuring the degree of loss aversion and, by construction, reference dependence in our subject pool, allows us to later test our hypotheses regarding performance and goal setting.

4.1 Risk Attitudes

To measure subjects' attitudes towards risky monetary payments, we use the data from part B of the experiment, where we elicit the certainty equivalents of five lotteries that offer positive payments, $\{x_1, x_2, x_3, x_4, x_5\}$, and the negative outcomes, $\{z_1, z_2, z_3, z_4, z_5\}$, that make subjects indifferent between receiving zero and a mixed lottery $(x_j, 0.5; z_j, 0.5)$ for each $j = \{1,2,3,4,5\}$. We classify participants according to the curvature of their utility function, as well as their sensitivity toward losses.

Most participants in our sample have linear utility functions in the domain of gains. Specifically, 91 participants are classified as having a linear utility function (proportion test against 0.5, p=0.014), while 65 have concave, and only five have convex, utility. Details of this classification are included in Appendix B.[7] If a power utility function $u(x) = x^{\theta}$ is assumed, the pooled estimate overall participants is $\hat{\theta} = 0.945$, close to risk neutrality. In Appendix B, we show that similar conclusions are reached when other families of utility functions are assumed. These results confirm that our theoretical assumption that individuals have linear utility functions is reasonable.

In addition, the great majority of participants are loss averse. Specifically, 117 participants are classified as loss averse and 23 as gain-seeking (more sensitive to gains than losses of the same magnitude). Table B.1 in Appendix B presents further details of this classification.[8] Importantly, participants are balanced across treatments with respect to both the mean level of loss aversion and the proportion of loss averse participants.[9]

Table 3 presents descriptive statistics for the loss aversion coefficients, λ_j, obtained by computing $\lambda_j = x_j / z_j$, for $j = 1, \dots, 5$. Following Abdellaoui et al. (2008), we compute one loss aversion coefficient for each mixed lottery that we implement. We find that, on average, participants exhibit loss aversion for every mixed lottery. The null hypothesis that the loss aversion coefficient is equal to one is rejected for each lottery. Moreover, we cannot reject the null that the five loss aversion coefficients are equal to each other (F(4,805)=0.63, sphericity-corrected p-value=0.644), corroborating the result that sign-dependence, rather than the magnitude of the loss, determines loss aversion (Köbberling and Wakker, 2005).

Table 3. Loss Aversion Levels for Each of the Five Lotteries
$\lambda_j = z_j / x_j$

	$\boldsymbol{\lambda_1}$	$\boldsymbol{\lambda_2}$	$\boldsymbol{\lambda_3}$	$\boldsymbol{\lambda_4}$	$\boldsymbol{\lambda_5}$	$\boldsymbol{\lambda_{average}}$	$\boldsymbol{\lambda_{median}}$
Mean	3.459	3.676	3.712	3.548	3.945	3.668	3.417
Median	1.777	1.777	1.777	1.777	2.285	2.136	1.777
S.D.	4.798	4.716	4.699	4.166	4.566	3.752	4.152
25th perc.	0.516	1.066	1.066	1.230	1.066	1.208	1.230
75th perc,	3.2	3.2	3.2	5.333	5.333	4.48	3.2

Aggregating the coefficients across lotteries and participants, we observe that participants exhibit a median coefficient of 2.14. This implies that, for our participants, losses loomed 2.14 times larger than equally

[7] In short, to classify participants according to their curvature we constructed variables $\Delta_{ij} \equiv x_{ij} - EV_j$, where x_{ij} is the certainty equivalent of participant i for lottery j, EV_j stands for the expected value of the lottery, and $j = \{1, \dots, 5\}$ is an indicator of the lottery used. A participant was classified as having a linear utility function if for at least four Δ_{ij}s the null hypothesis that they are equal to zero was not rejected.

[8] A subject is loss averse when at least four of her variables λ_j, are greater than one, where $\lambda_{ij} \coloneqq x_{ij}/z_{ij}$. We adopt this classification from Abdellaoui et al. (2008).

[9] The proportion of loss averse participants is 0.717 in LOPR, 0.6923 in GOAL, 0.6923 in GOAL+BONUS and 0.804 in HIPR. We use a two-sample test of proportions and find that these proportions are not significantly different from each other (LOPR vs. GOAL (p=0.786), HIPR vs. GOAL (p= 0.230), GOAL+BONUS vs. GOAL (p= 0.985)). Mean loss aversion is on average 3.44 for GOAL, 3.94 for GOAL+BONUS, 3.55 for HIPR and 3.75 for LOPR. The distribution of loss aversion is not significantly different between pairs of treatments, according to Mann-Whitney tests (LOPR vs. GOAL (p= 0.461), HIPR vs. GOAL (p=0.639), GOAL+BONUS vs. GOAL (p= 0.533)).

sized gains for the median individual. Previous studies that used the same definition of loss aversion found a median loss aversion parameter of similar magnitude to our 2.14. For instance, Tversky and Kahneman's (1992) median estimate was 2.25, Abdellaoui et al. (2007) reported an estimate of 2.54, Abdellaoui et al. (2016) observed 1.88, and Abdellaoui et al. (2008), using the same method to elicit loss aversion as we have employed here, obtained a median loss aversion parameter equal to 2.61. As in all previous studies we observe considerable heterogeneity, reflected in the size of the interquartile range (the difference between the 25th and 75th percentile).

4.2 Performance Under the Different Contracts

Since most participants are loss averse (and therefore have reference-dependent preferences), we hypothesize that performance will be on average higher in the GOAL treatment than in the GOAL+BONUS treatment and this difference in average performance will be higher for participants with greater loss aversion (Hypothesis 1). Recall that we define performance in the experiment as the total number of tables an individual solves correctly.

Table 4 reports the descriptive statistics of performance by treatment and Figure 3 shows the Probability Density Functions (PDFs) of performance in GOAL and GOAL+BONUS. As predicted, paying a monetary bonus for achieving a goal backfires, resulting in lower output. On average, participants in GOAL solve more tables (47.58 tables), than participants in GOAL+BONUS (42.897 tables) ($t = 1.485$, $p = 0.07$).[10] The size of this effect is 11.1%, or 0.377 standard deviations.

Table 4. Performance by Treatment

Treatment	**N**	**Mean**	**Median**	**S.D.**	**25th**	**75th**	**Max**	**Min**	**Mean Cost (Dollars)**
GOAL +BONUS	39	42.897	43	13.480	34	48	80	17	9.74
GOAL	41	47.585	47	14.747	39	56	86	17	9.517
HIPR	41	42.073	43	15.408	31	48.5	72	5	21.036
LOPR	39	42.179	41	15.022	27	55	72	16	8.436
Total	**160**	**44.3**	**43.5**	**14.734**	**34**	**54**	**86**	**5**	**12.259**

In addition, participants in GOAL solve more tables than participants in LOPR ($t = 1.623$, $p = 0.054$) and even more than in HIPR ($t = 1.6548$, $p=0.051$).[11] These represent differences of 0.401 standard deviations and 0.406 standard deviations, respectively. These results validate Hypothesis 4. The fact that output is higher under GOAL than under HIPR suggests that a self-chosen goal contract without a bonus is highly cost-effective.

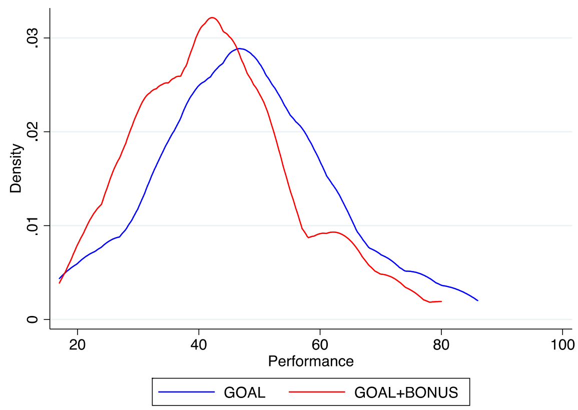
However, we do not find significant differences in performance between (a) GOAL+BONUS and LOPR ($t = 0.222$, $p = 0.824$), (b) GOAL+BONUS and HIPR ($t = 0.255$, $p = 0.799$) or (c) LOPR and HIPR ($t = 0.312$,

[10] A Wilcoxon-Mann-Whitney test generates the same conclusion (U= 1.584, p=0.056).

[11] Wilcoxon-Mann-Whitney tests of these differences yield U= 1.494 (p=0.074) and U= 1.512 (p=0.065), respectively.

p= 0.975).[12] Raising the piece rate or adding a self-chosen contract with a bonus did not improve average performance over LOPR. As it will be shown later, restricting the sample to individuals with linear utility, as it is assumed in the model, does lead to differences in performance across these treatments.

Figure 3. Probability Density Function of Performance in the Treatments with Goal Setting



We also perform regressions of individual performance on treatment dummies that confirm these results. We use Poisson count regressions to account for the count nature of the performance data. Table 5 shows that the coefficient associated with GOAL is positive and significant at the 5% level for all specifications, indicating that participants in GOAL exhibit higher average performance than participants in GOAL+BONUS, the benchmark category. Similarly, the coefficient of GOAL is significantly larger than the coefficient of LOPR ($\chi^2 = 13.28$, p= 0.001) and HIPR ($\chi^2 = 16.79$, p= 0.001).[13 14 15] These results are robust controlling for loss aversion (Column 2).

In column 3 of Table 5 we show regression estimates when the sample is restricted to participants with linear utility, in line of the model's assumption. All the results reported above are robust to this sample restriction, and we observe that performance in GOAL+BONUS is higher than performance in LOPR.

To test Hypothesis 3, we examine the role of loss aversion in explaining the performance differences between GOAL and GOAL+BONUS. We start by distinguishing participants who are loss averse from those who are not. We create a dummy variable labeled "Loss Averse" that equals one if the participant has a loss aversion parameter greater than one, and zero otherwise. We then extend the Poisson count regression model presented above by adding the interaction between this "Loss Averse" dummy and the GOAL dummy. We perform the analysis only with participants assigned to either GOAL or GOAL+BONUS.

[12] These conclusions are also confirmed by Wilcoxon-Mann-Whitney tests. The U-statistics of these comparisons and their respective p-values are U= 0.110 (p= 0.912), U= 0.048 (p= 0.961), and U= 0.034 (p= 0.973), respectively.

[13] The coefficients for the LOPR and HIPR treatments are not significantly different for the specification in columns (1) and (2) ($\chi^2 = 0.16$, p= 0.688).

[14] We use the estimates of column (2) in Table 5 for statistical inference.

[15] All results are robust to adding an ability variable in the regression, measured by the number of correct tables that participants completed in the 5-minute practice round.

Table 5. Performance as Function of Treatment and Preference Parameters

	(1)	(2)	(3)
	Performance (all participants)	**Performance (all participants)**	**Performance (linear utility only)**
GOAL	0.104^{***}	0.102^{***}	0.177^{***}
	(0.033)	(0.033)	(0.045)
HIPR	-0.019	-0.034	0.098^{**}
	(0.034)	(0.034)	(0.038)
LOPR	-0.017	-0.020	-0.134^{***}
	(0.035)	(0.035)	(0.048)
Loss Averse		0.135^{***}	0.169^{***}
		(0.028)	(0.038)
Constant	3.759^{***}	3.664^{***}	3.585^{***}
	(0.024)	(0.032)	(0.043)
Log-Likelihood	-846.727	-834.826	-442.010
N	160	160	91

Note: This table presents the estimates of Poisson count regressions of the statistical model $\text{Performance}_i = \beta_0 + \beta_1 \text{GOAL} + \beta_2 \text{HIPR} + \beta_3 \text{LOPR} + \beta_4 \text{Loss Averse} + \mathcal{E}_i$ with $\mathcal{E}_i \sim \text{Poisson}(\omega)$. "Performance" is the total number of tables a participant solves correctly over the six rounds of the real-effort task. Participants were assigned either to the GOAL, HIPR, LOPR, or GOAL+BONUS treatment. The GOAL+BONUS treatment is the benchmark category of the regression. ``Loss Averse'' is a dummy variable that indicates whether a participant is loss averse or not. A participant is classified as loss averse when at least four of her variables λ_j, where $\lambda_j \equiv x_j/z_j$, are greater than one. A participant is classified as having a linear utility function when for at least four Δ_{ij}s the null hypothesis that they are equal to zero was not rejected. Model (3) presents estimates of a regression including only those participants classified as having linear utility. Standard errors are in parenthesis. * indicates a p-value <0.1, ** indicates a p-value <0.05, ***indicates a p-value <0.01

Table 6, Column 1 presents the results of this regression. The coefficient of the interaction term is positive and significant, which implies that under the GOAL contact, loss averse participants perform better than non-loss averse participants. Moreover, loss averse participants perform better when assigned to GOAL than when assigned to GOAL+BONUS ($\chi^2(1)$ =7.76, p<0.01). Participants who are not loss averse perform equally under GOAL than under GOAL+BONUS. Figure 4(a) illustrates this result using the marginal effects of these estimates.

To dig further into the interaction between the levels of loss aversion and the monetary bonus, we classify participants as being either above (High Loss Averse) or below (Low Loss Averse) the median loss aversion level (1.77). We use the same Poisson count regression model with the interaction term as above. The results are presented in Column 2 of Table 6, and the marginal effects are presented in Figure 4(b). Again, we observe that High Loss Averse participants perform significantly better under GOAL than under GOAL+BONUS, while there is no significant difference in performance across treatments for Low Loss Averse participants.

To finalize this analysis, we also use a continuous measure of loss aversion. In that model, we allow for different slopes for loss averse relative to non-loss averse subjects. Our results are confirmed: for loss averse subjects the slope in GOAL is significant and positive whereas that in GOAL+BONUS is significant and negative (p=0.004). Hence, subjects with higher loss aversion exhibit a larger performance difference between GOAL and GOAL+BONUS. Results are presented in Table C.2 in Appendix C. Furthermore,

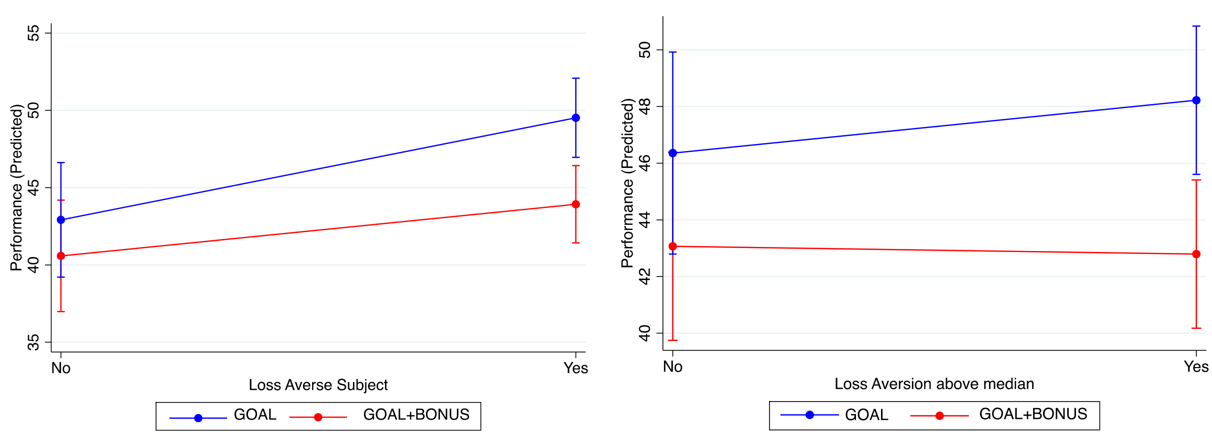
Table C.3 and Figure C.1. in Appendix C corroborate our model's main implication that subjects with high loss aversion *and* a linear utility for money are more affected by the monetary bonus, findings that are supportive of Hypothesis 1.

Table 6. Heterogeneity of Treatment Effects by Participant Loss Aversion Level

	(1)	**(2)**
	Performance	**Performance**
GOAL	0.055	0.073
	(0.063)	(0.055)
Loss Averse	0.079	
	(0.053)	
GOAL* Loss Averse	0.198^{***}	
	(0.052)	
High Loss Averse		-0.006
		(0.050)
GOAL* High Loss Averse		0.113^{***}
		(0.048)
Constant	3.703^{***}	3.762^{***}
	(0.045)	(0.039)
Log-Likelihood	-391.929	-396.641
N	80	80

Note: This table presents the estimates of the Poisson regression of the specification $\text{Performance}_i = \beta_0 + \beta_1 \text{GOAL} * \text{Loss Averse} + \beta_2 \text{GOAL} + \beta_3 \text{LOPR} + \beta_4 \text{LOPR} + \beta_5 \text{Loss Averse} + \mathcal{E}_i$ with $\mathcal{E}_i \sim \text{Poisson}(\omega)$ with $\mathcal{E}_i \sim \text{Poisson}(\omega)$. "Performance" is the total number of correctly solved tables by a participant over all rounds. GOAL+BONUS is the benchmark category. ``Loss Averse'' is a dummy variable that captures whether a participant is loss averse or not. A participant is loss averse when at least four of her variables $\lambda_j$, where $\lambda_j \equiv x_j/z_j$, are greater than one. In the second column ``Loss Averse'' is replaced by ``High Loss averse'' which equals 1 if her average λ is greater than that of the median participant in the sample and 0 otherwise. See Appendix B for a detailed explanation of these measurements. Standard errors are in parenthesis. * indicates a p-value <0.1, ** indicates a p-value <0.05, ***indicates a p-value <0.01.

Figure 4. Treatment differences in performance by loss aversion levels.



a) With $\lambda = 1$ as threshold

b) With median λ as threshold (median split)

We summarize the above analysis on performance as follows:

Result 1: *Performance is higher in GOAL than in any of the other treatments, including GOAL+BONUS. Performance in GOAL+BONUS is also higher than in LOPR for participants with linear utility.*

Result 2: *Performance differences between GOAL and GOAL+BONUS are more pronounced for participants with high loss aversion.*

4.3 Goal Setting

Regarding goals, the model predicts that participants, the majority of whom have reference-dependence preferences, will set higher average goals in GOAL than in GOAL+BONUS (Hypothesis 2). Also, the model predicts that the difference in goal levels between GOAL and GOAL+BONUS increases with loss aversion (Hypothesis 3).

Table 7 presents the descriptive statistics of goals by treatment and Figure 5 presents the Probability Density Functions (PDFs) of goals in the two treatments. As predicted by the model, participants in GOAL set significantly higher goals on average (48.82 tables) than participants in GOAL+BONUS (37.48 tables) (t=2.842, p=0.005).[16] This represents a difference of 30.2%, or 0.573 standard deviations.

In Table C.1 presented in Appendix C, we show that the difference in goals between GOAL and GOAL+BONUS increases in later rounds, indicating that not only do participants adjust their goals after being provided with feedback, but also that this adjustment induces a larger difference in goal setting between the two treatments. This rules out the alternative explanation that the observed difference in goals is an artifact of early periods and dissipates with experience.

In Table 8, we report estimates from regressions of the goals set by participants, controlling for loss aversion with a dummy variable in Columns (2) and (3). The results confirm that participants set higher goals when goal achievement is not rewarded monetarily. Moreover, more loss averse individuals set on average higher goals.

Table 7. Mean and Median Goals Set, by Treatment

Treatment	**N**	**Mean**	**Median**	**S.D.**	**25th perc.**	**75th perc.**	**Max.**	**Min.**
GOAL +BONUS	39	37.487	38	10.308	29	43	64	20
GOAL	41	48.829	45	25.706	34	54	180	18
Total	**80**	**44.300**	**40**	**22.758**	**31**	**48.5**	**180**	**18**

[16] A Wilcoxon-Mann-Whitney test yields the same conclusion (U=2.842, p=0.004).

Figure 5. PDFs of Goals Set, by Treatment

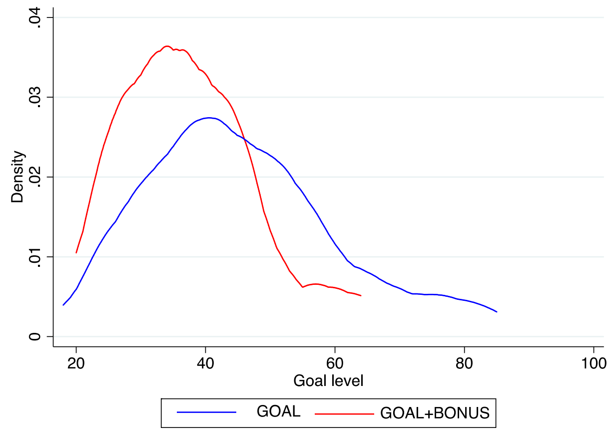


Table 8. Goals as Function of Treatment and Preference Parameters

	(1) **Goal Level (all participants)**	(2) **Goal Level (all participants)**	(3) **Goal Level (linear utility only)**
GOAL	0.264***	0.262***	0.378***
	(0.034)	(0.034)	(0.045)
Loss Averse		0.133***	0.241***
		(0.038)	(0.054)
Constant	3.624***	3.530***	3.426***
	(0.026)	(0.038)	(0.053)
Log-Likelihood	-475.870	-469.699	-280.819
N	80	80	44

Note: This table presents the estimates of Poisson count regressions of the statistical model $\text{Goal Level} = \beta_0 + \beta_1\text{GOAL} + \Gamma'\text{Controls} + \mathcal{E}_i$ with $\mathcal{E}_i \sim \text{Poisson}(\omega)$. "Goal Level" equals the sum of a participant's goals over all six rounds of the real-effort task. Participants were randomly assigned either to the "GOAL" or the "GOAL+BONUS" treatment. The latter is the benchmark condition of the regression. ``Loss Averse'' is a dummy variable that captures whether a participant is loss averse or not. A participant is classified as loss averse when at least four variables λ_j, where $\lambda_j \equiv x_j/z_j$, are larger than one. Model (3) presents estimates of a regression including only participants classified as having linear utility. Standard errors presented in parenthesis. * indicates a p-value <0.1, ** indicates a p-value <0.05, ***indicates a p-value <0.0

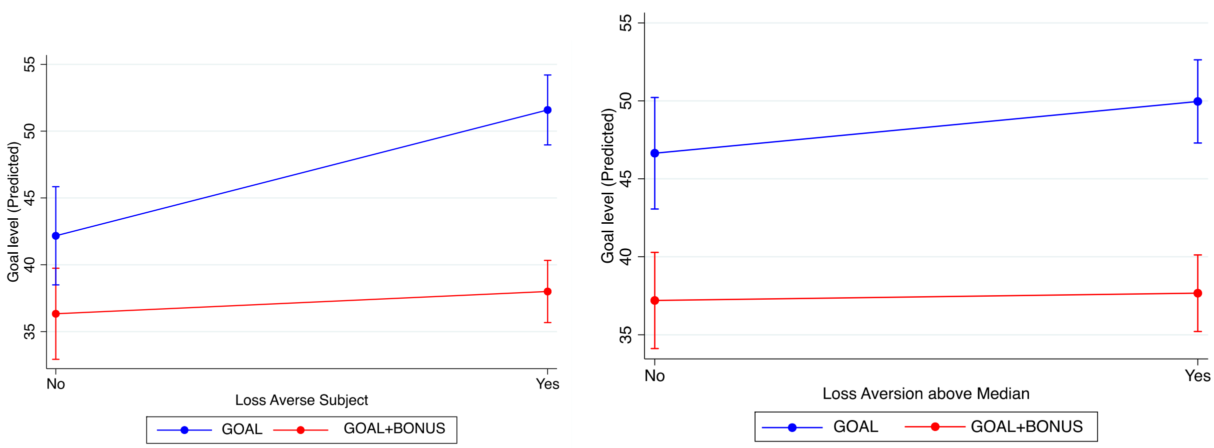
To test Hypothesis 3, we run models with interaction terms similar to those used to estimate treatment effects on performance and find similar results (Table 9). Participants in GOAL set higher goals than in GOAL+BONUS, and this difference is larger when participants are loss averse ($\chi^2(1)$ =15.37, p=0.001) (Column 1 and Figure 6(a)). In another specification (Col 2-Table 9), we investigate whether the difference in goal levels between GOAL and GOAL+BONUS is greater for participants with above median levels of loss aversion than for participants with below median levels of loss aversion. The results in Column 2 and Figure 6 (b) confirm this hypothesis.

Table 9. Heterogeneity of Treatment Effects by Participant Loss Aversion Level

	(1)	**(2)**
	Goal Level	**Goal Level**
GOAL	0.148**	0.159*
	(0.065)	(0.083)
GOAL * Loss Averse	0.350***	
	(0.054)	
Loss Averse	0.045	
	(0.057)	
GOAL * High Loss Averse		0.309***
		(0.072)
High Loss Averse		0.026
		(0.075)
Constant	3.593***	3.601**
	(0.048)	(0.067)
Log-Likelihood	-467.624	-469.951
N	80	80

Note: This table presents the estimates of the Poisson regression of the specification $\text{Goal level}_i = \beta_0 + \beta_1 \text{GOAL} * \text{Loss Averse} + \beta_2 \text{GOAL} + \beta_3 \text{LOPR} + \beta_4 \text{LOPR} + \beta_5 \text{Loss Averse} + I$ with $\mathcal{E}_i \sim \text{Poisson}(\omega)$. "Goal level" is the sum of the goals set by the participant over all rounds. GOAL+BONUS is the benchmark category for the regression. ``Loss Averse'' is a dummy variable that captures whether a participant is loss averse or not. A participant is loss averse when at least four of her variables $\lambda_j$, where $\lambda_j \equiv x_j/z_j$, are greater than one, ``Mild Loss averse' equals 1 if a participant is loss averse and her average λ is lower than that of the median participant in the sample, and equals 0 otherwise. ``High Loss averse' equals 1 if a participant is loss averse and her average λ is greater than that of the median participant in the sample and 0 otherwise. See Appendix B for a detailed explanation of these measurements. Standard errors are in parenthesis. * indicates a p-value <0.1, ** indicates a p-value <0.05, ***indicates a p-value <0.01.

Figure 6. Treatment effects in goal levels by loss aversion levels.



a) With $\lambda = 1$ as threshold

b) With median λ as threshold

Finally, we use a continuous measure of loss aversion, and estimate a model that considers different slopes for loss averse and non-loss averse subjects. Estimates given in Appendix C, Table C.2, confirm that participants with higher loss aversion display a larger goal setting difference between GOAL and GOAL+BONUS. That is because the slope of loss averse subjects under GOAL is significant and positive and that of loss averse subjects in GOAL+BONUS is significant and negative. Also, in Table C.3 and Figure C.1, we show that subjects with high loss aversion and linear utility exhibit a larger treatment effect. These results are consistent with Hypotheses 2 and 3.

We summarize the above analysis on goal setting as follows:

Result 3: *Goals are higher under GOAL than under the GOAL+BONUS.*

Result 4: *The difference in goal levels between GOAL and GOAL+BONUS is greater for participants with high loss aversion.*

5. Conclusion

This paper shows that offering monetary bonuses for the achievement of self-chosen goals can backfire on the employer, and result in worse performance than would be achieved without the bonus. In a setting in which workers exhibit a sufficient degree of loss aversion, a monetary bonus for meeting a self-chosen goal can crowd out the motivational effect of the goal itself. Loss averse workers will set lower (though positive) goals to increase the likelihood of reaching the monetary bonus, which will in turn lead to lower performance. Thus, a self-chosen goal contract with no monetary payment can lead to better performance than one where achieving the goal yields a monetary bonus. The results from our experiment confirm this prediction.

Insofar as our empirical findings generalize to less controlled, non-laboratory environments, this paper suggests that including monetary bonuses in a worker's compensation scheme does not necessarily guarantee better performance. Our finding in this regard is that it depends on worker preferences. We have shown this in a setting in which increasing the piece rate did not improve performance, perhaps because the income effect from a greater piece rate reduces effort that may offset the substitution effect that increases effort. A scheme in which individuals set their own goals, when achieving the goal carries no monetary bonus, is the most effective incentive scheme that we have studied. From an employer's point of view, given that there is already a piece rate in place, a non-monetarily rewarded self-chosen goal scheme dominates the others we have studied, though obviously further research would be required to enable stronger claims about the generality of our results.

References

Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science*, *46*(11), 1497–1512.

Abdellaoui, M., Bleichrodt, H., and L'Haridon, O. (2008). A Tractable Method to Measure Utility and Loss Aversion Under Prospect Theory. *Journal of Risk and Uncertainty*, *36*(245), 1659-1674.

Abdellaoui, M., Bleichrodt, H., L'Haridon, O., and van Dolder, D. (2016). Measuring Loss Aversion under Ambiguity : A Method to Make Prospect Theory Completely Observable. *Journal of Risk and*

Uncertainty, *52*, 1–20.
Abdellaoui, M., Bleichrodt, H., and Paraschiv, C. (2007). Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. *Management Science*, *53*(10), 1659–1674.
Abeler, J., Falk, A., Goette, L., and Huffman, D. (2011). Reference Points and Effort Provision. *American Economic Review*, *101*(2), 470–492.
Allen, E., Dechow P., Pope D. and Wu G. (2017). Reference-Dependent Preferences: Evidence from Marathon Runners", *Management Science*, 63, 1657-1672.
Ariely, B. D., Bracha, A., and Meier, S. (2009). Doing Good or Doing Well ? Image Motivation and Monetary Incentives in Behaving Prosocially. *The American Economic Review*, *99*(1), 544–555. (Ariely et al., 2009a).
Ariely, D., Gneezy, U., Loewenstein, G., and Mazar, N. (2009). Large Stakes and Big Mistakes. *The Review of Economic Studies*, *76*, 451–469. (Ariely et al., 2009b)
Benabou, R., and Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies*, *70*(1), 489–520.
Bourne, M., Neely, A., Mills, J. and Platts, K. (2003). Implementing Performance Measurement Systems: a Literature review. International Journal of Business Performance Management, 5(1), 1-24.
Bowles, S., and Polanía-Reyes, S. (2012). Economic Incentives and Social Preferences: Substitutes or Complements? *Journal of Economic Literature*, *50*(2), 368–425.
Brookins, P., Goerg, S. J., and Kube, S. (2017). Self-chosen Goals, Incentives and effort. *Working Paper*.
Chan, W., Proschan, F., & Sethuraman, J. (1990). Convex-ordering among functions, with applications to reliability and mathematical statistics. *Lecture Notes-Monograph Series*, 121-134.
Chen, D. L., Schonger, M., and Wickens, C. (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.
Chung, D. J., Steenburgh, T., and Sudhir, K. (2013). Do Bonuses enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-based Compensation Plans. *Marketing Science*, *33*(2), 165-187.
Chiappori, P. A., & Salanié, B. (2002). Testing contract theory: A survey of some recent work. *Available at SSRN 318780*.
Corgnet, B., Gómez-Miñambres, J., and Hernán-gonzalez, R. (2015). Goal Setting and Monetary Incentives : When Large Stakes Are Not Enough. *Management Science*, *61*(12), 2926–2944.
Corgnet, B., Gómez-Miñambres, J., and Hernán-Gonzalez, R. (2018). Goal Setting in the Principal-Agent Model: Weak Incentives for Strong Performance. *Games and Economic Behavior*, *109*, 311–326.
Dalton, P. S., Gonzalez-Jimenez, V., and Noussair, C. N. (2016). Self-Chosen Goals: Incentives and Gender Differences, *CentER Discussion Paper Series* No. 2016-036. (Dalton et al., 2016a)
Dalton, P. S., Ghosal S., and Mani A. (2016). Poverty and Aspirations Failures, *Economic Journal*, *126(590)*, 165-188. (Dalton et al., 2016b)
De Meza, D., and Webb, D. C. (2007). Incentive design under loss aversion. *Journal of the European Economic Association*, *5*(1), 66-92.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522-550.
Frey, B. S., and Oberholzer-gee, F. (1997). The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out. *The American Economic Review*, *87*(4), 746–755.
Gallo (2011, February 7) Making sure your employees succeed. Retrieved from https://hbr.org/2011/02/making-sure-your-employees-suc#comment-section
Gibbons, R., and Roberts, J. (2013). *The Handbook of Organizational Economics*. Princeton: Princeton University Press.
Gneezy, U., Goette, L., Sprenger, C., and Zimmermann, F. (2017). The limits of expectations-based reference dependence. *Journal of the European Economic Association*, *15*(4), 861–876.
Gneezy, U., Meier, S., and Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, *25*(4), 191–210.
Gneezy, U., and Rustichini, A. (2000). Pay Enough or Don't Pay at All. *Quarterly Journal of Economics*,

115(3), 791–810.
Gómez-Miñambres, J. (2012). Motivation through goal setting. *Journal of Economic Psychology*, *33*(6), 1223–1239.
Gonzalez, R., and Wu, G. (1999). On the Shape of the Probability Weighting Function. *Cognitive Psychology*, *166*, 129–166.
Groen, B. A., Wouters, M. J., and Wilderom, C. P. (2012). Why do employees take more initiatives to improve their performance after co-developing performance measures? A field study. *Management Accounting Research*, *23*(2), 120-141.
Groen, B.A., and Wilderom, C.P. and Wouters M.J. (2015). High Job Performance through co-developing performance with employees. *Human resource management*, *56*(1), 111-132.
Heath, C., Larrick, R. P., and Wu, G. (1999). Goals as reference points. *Cognitive Psychology*, *38*(1), 79–109.
Herweg, F., Müller, D., and Weinschenk, P. (2010). Binary payment schemes: Moral hazard and loss aversion. *American Economic Review*, *100*(5), 2451-77.
Holmström, B., & Milgrom, P. (1987). Linearity and Aggregation in the Provision of Intertemporal Incentives. *Econometrica*, *55*(2), 303-28.
Hsiaw, A. (2013). Goal-setting and self-control. *Journal of Economic Theory*, *148*(2), 601–626.
Hsiaw, A. (2018). Goal bracketing and self-control. *Games and Economic Behavior*, *111*(December), 100–121.
Kaur, S., Kremer, M., and Mullainathan, S. (2015). Self-control at work. *Journal of Political Economy*, *123*(6), 1227-1277.
Kim, S. K. (1997). Limited liability and bonus contracts. *Journal of Economics & Management Strategy*, *6*(4), 899-913.
Köbberling, V., & Wakker, P. P. (2005). An index of loss aversion. *Journal of Economic Theory*, *122*(1), 119-131.
Koch, A. K., and Nafziger, J. (2011). Self-regulation through Goal Setting. *Scandinavian Journal of Economics*, *113*(1), 212–227.
Koch, A. K., and Nafziger, J. (2016). Goals and bracketing under mental accounting. *Journal of Economic Theory*, *162*, 305–351.
Koch, A. K., and Nafziger, J. (2019). Motivational goal bracketing. *Journal of Economic Theory* (Forthcoming).
Koszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature*, *52*(4), 1075-1118.
Laffont, J. J., and Martimort, D. (2002). *The theory of incentives: The Principal-Agent Model*. Princeton University Press.
Laffont, J. J., and Tirole, J. (1993). *A Theory of Incentives in Procurement and Regulation*. Cambridge(Mass.): MIT-Press.
Larkin, I. (2014). The cost of high-powered incentives: Employee gaming in enterprise software sales. *Journal of Labor Economics*, *32*(2), 199-227.
Markle, A., Wu, G., White, R., and Sackett, A. (2018). Goals as Reference Points in Marathon Running: A Novel Test of Reference Dependence. *Journal of Risk and Uncertainty*, *56*(1), 19-50.
Mellstrom, C., and Johannesson, M. (2016). Crowding out in Blood Donation: Was Titmuss Right? *Journal of the European Economic Association*, *6*(4), 845–863
Morree de, Pim (2018) The Science of Target Setting (and how most companies get it wrong). Retrieved from https://corporate-rebels.com/target-setting/
Park, E. S. (1995). Incentive contracting under limited liability. *Journal of Economics & Management Strategy*, *4*(3), 477-490.
Oyer, P. (2000). A theory of sales quotas with limited liability and rent sharing. *Journal of labor Economics*, *18*(3), 405-426.
Tversky, A., and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.
Worldatwork. (2018). *2018 Incentive Pay Practices for Publicly-Traded Companies*.

Wu, G., Heath, C., and Larrick, R. (2008). A prospect theory model of goal behavior. *Mimeo*. Retrieved from http://faculty.chicagobooth.edu/george.wu/research/papers/wu heath larrick (prospect theory model of goal behavior).pdf

Appendix A. Proofs

Lemma 1.

Proof. The optimal goal for given effort level, $g(e)$, satisfies the following first-order condition:

$$b\big(1-F(g(e)|e)\big)-bg(e)f(g(e)|e)-\mu-\mu(\lambda-1)F(g(e)|e)=0. \quad (A1)$$

Eq. $(A1)$ can be rewritten as

$$g(e)=\frac{b\big(1-F(g(e)|e)\big)-\mu\big(1+(\lambda-1)F(g(e)|e)\big)}{bf(g(e)|e)} \qquad (A2)$$

For $g(e)$ to be a maximum, it is sufficient that the second-order condition evaluated at $g(e)$ satisfies:

$$-\big(2b+\mu(\lambda-1)\big)f(g^*(e)|e)-bg^*(e)f'(g^*(e)|e)<0. \qquad (A3)$$

Note that if $b=0$ then $(A3)$ is always satisfied. If $b>0$, a sufficient condition for Eq. $(A3)$ to hold is $f'(g^*(e)|e)>0$. Instead, if Eq. $(A3)$ does not hold, the agent chooses $g(e)=0$ since $\mathbb{E}\left(U\left(w_g,e,0\right)\right)>\mathbb{E}\left(U\left(w_g,e,\overline{y}\right)\right)$. Thus, to compute comparative statics, we focus on the case in which Eq. (3) is satisfied and $g(e)$ is interior.

From Eq. $(A1)$ is evident that $\frac{dg(e)}{da}=0$ and $\frac{dg(e)}{dc}=0$. To investigate the effect of bonuses, we differentiate implicitly $(A1)$ with respect to b to obtain:

$$\frac{dg(e)}{db}=\frac{\big(1-F(g(e)|e)\big)-g(e)f(g(e)|e)}{\big(2b+\mu(\lambda-1)\big)f(g(e)|e)+bg(e)f'(g(e)|e)}. \qquad (A4)$$

According to the condition in Eq. (3) it must be that the denominator of Eq. $(A4)$ is positive. Moreover, from Eq. (2) it is evident that $\max\{g(e)\}=\frac{1-F\left(g(e)|e\right)}{f(g|e)}$, so $g(e)\leq\frac{1-F\left(g(e)|e\right)}{f(g|e)}$ and the numerator of Eq. $(A4)$ is non-negative. Hence, $\frac{dg(e)}{db}\geq 0$.

Next, implicitly differentiate $(A1)$ with respect to λ to obtain:

$$\frac{dg(e)}{d\lambda}=\frac{-\mu F(g(e)|e)}{\big(2b+\mu(\lambda-1)\big)f(g(e)|e)+bg(e)f'(g(e)|e)}. \qquad (A5)$$

Due to Eq. $(A3)$ it must be that at the optimum the denominator of Eq. $(A5)$ is positive. Therefore, $\frac{dg(e)}{d\lambda}<0$. ■

Lemma 2.

Proof. Fix g. Differentiation of Eq. (6) with respect to a gives:

$$\frac{dIC(g)}{da}=\mathbb{E}(y|e_H)-\mathbb{E}(y|e_L). \qquad (A6)$$

Due to Assumption 2, the monotone likelihood ratio property, then $\mathbb{E}(y|e_H) - \mathbb{E}(y|e_L) \geq 0$. So $\frac{dIC(g)}{da} \geq 0$. To investigate the effect of bonuses on the decision to exert high effort, we differentiate Eq. (6) with respect to b to obtain:

$$\frac{dIC(g)}{db} = g(F(g|e_L) - F(g|e_H)). \qquad (A7)$$

Due to Assumption 2, $F(g|e_L) - F(g|e_H) \geq 0$ so $\frac{dIC(g)}{db} \geq 0$.

Moreover, the derivative of (6) with respect to λ gives:

$$\frac{dIC(g)}{d\lambda} = \mu \int_0^g \big(F(y|e_L) - F(y|e_H)\big)dy > 0. \qquad (A8)$$

The above equation shows that $\frac{dIC(g)}{d\lambda} \geq 0$. Finally, differentiating (6) with respect to c is equal to $\frac{\partial IC(g)}{\partial c} = -1 < 0$. ■

Lemma 3

Proof. Fix g. To investigate the effect of higher goals on the decision to exert high effort, we derive (6) with respect to g to obtain:

$$\frac{\partial IC(g)}{\partial g} = \big(b + \mu(\lambda - 1)\big)\big(F(g|e_L) - F(g|e_H)\big) + bg\big(f(g|e_L) - f(g|e_H)\big). \qquad (A9)$$

Assumption 2 implies $F(g|e_L) - F(g|e_H) \geq 0$, so the first expression of the above equation is non-negative. The sign of the second expression in $(A9)$ depends on the location of g. Let $g \leq \hat{y}$. In that case $f(g|e_L) \geq f(g|e_H)$ and that second expression non-negative, which implies $\frac{dIC}{dg} \geq 0$. Instead, if $g > \hat{y}$ then $f(g|e_L) < f(g|e_H)$ and $\frac{dIC}{dg} \geq 0$ is not guaranteed. A condition for $\frac{\partial IC}{\partial g} \geq 0$ under $g > \hat{y}$ is that λ is sufficiently large so that the first expression in $(A9)$ outweighs the second expression in that equation. Specifically, $\frac{dIC}{dg} \geq 0$ as long as $\lambda \geq \hat{\lambda} := \frac{b}{\mu}\left(\frac{g\big(f(g|e_H) - f(g|e_L)\big)}{F(g|e_L) - F(g|e_H)} - 1\right) + 1$.

Finally, let $b = 0$. The derivative of (6) with respect to g becomes:

$$\frac{\partial IC(g)}{\partial g} = \mu(\lambda - 1)\big(F(g|e_L) - F(g|e_H)\big) \qquad (A10)$$

then $\frac{dIC}{dg} \geq 0$ regardless of whether $g \leq \hat{y}$ or $g > \hat{y}$. ■

The following Lemma is not included in the main text of the paper as it yields a similar conclusion as Lemma 3. Specifically, it shows sufficient conditions on b and μ such that optimal goal increases in effort. This result is however achieved comparing optimal goals at different effort levels.

Lemma 4

i. *If* $b > \mu$*, then* $g(e_H) > g(e_L)$*;*
ii. *If* $\mu > b$ *and* $g(e_L) \geq \hat{y}$*, then* $g(e_H) > g(e_L)$*;*
iii. *If* $\mu > b$*,* $g(e_L) < \hat{y}$*, and large enough loss aversion, then* $g(e_H) > g(e_L)$*.*

Proof. Suppose that $g(e_H) \geq g(e_L)$. We use the expression in Eq. $(A2)$ to obtain:

$$g(e_H) \geq g(e_L) \Leftrightarrow \left(1-\frac{\mu}{b}\right)\left(\frac{1-F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{1-F(g(e_L)|e_L)}{f(g(e_L)|e_L)}\right) - \frac{\mu\lambda}{b}\left(\frac{F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{F(g(e_L)|e_L)}{f(g(e_L)|e_L)}\right) \geq 0. \qquad (A11)$$

Assumption 2 implies $\frac{1-F(g(e_H)|e_H)}{f(g(e_H)|e_H)} \geq \frac{1-F(g(e_L)|e_L)}{f(g(e_L)|e_L)}$, hazard rate dominance. That assumption also implies $\frac{F(g(e_H)|e_H)}{f(g(e_H)|e_H)} \leq \frac{F(g(e_L)|e_L)}{f(g(e_L)|e_L)}$, inverse hazard rate dominance. Hence, according to Eq. $(A11)$ a sufficient condition for $g(e_H) \geq g(e_L)$ is $b > \mu$.

Let instead $\mu > b$. Rewrite $(A11)$ as

$$\frac{1}{\lambda}\left(\frac{b}{\mu}-1\right) > \frac{\frac{F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{F(g(e_L)|e_L)}{f(g(e_L)|e_L)}}{\left(\frac{1-F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{1-F(g(e_L)|e_L)}{f(g(e_L)|e_L)}\right)}. \qquad (A12)$$

Since $0 \geq \frac{1}{\lambda}\left(\frac{b}{\mu}-1\right) \geq -1$, a sufficient condition for $(A12)$ to hold is $-1 > \frac{\frac{F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{F(g(e_L)|e_L)}{f(g(e_L)|e_L)}}{\left(\frac{1-F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{1-F(g(e_L)|e_L)}{f(g(e_L)|e_L)}\right)}$, which after some manipulations gives $1 > \frac{f(g(e_L)|e_L)}{f(g(e_H)|e_H)}$. Hence, under $\mu > b$ the fact that $g(e_L) > \hat{y}$ guarantees the inequality in $(A12)$.

Finally, we establish the conditions for $g(e_H) < g(e_L)$. Using similar steps as above, we obtain:

$$g(e_H) < g(e_L) \Leftrightarrow \frac{1}{\lambda}\left(\frac{b-\mu}{\mu}\right) < \frac{\frac{F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{F(g(e_L)|e_L)}{f(g(e_L)|e_L)}}{\left(\frac{1-F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{1-F(g(e_L)|e_L)}{f(g(e_L)|e_L)}\right)}. \qquad (A13)$$

Since $0 > \frac{1}{\lambda}\left(\frac{b}{\mu}-1\right) > -1$, a necessary condition for (13) to hold is

$$-1 < \frac{\frac{F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{F(g(e_L)|e_L)}{f(g(e_L)|e_L)}}{\left(\frac{1-F(g(e_H)|e_H)}{f(g(e_H)|e_H)}-\frac{1-F(g(e_L)|e_L)}{f(g(e_L)|e_L)}\right)} < 0 \Leftrightarrow 1 < \frac{f(g(e_L)|e_L)}{f(g(e_H)|e_H)} < \frac{F(g^*(e_L)|e_L)}{F(g^*(e_H)|e_H)} \quad (A14)$$

Hence, $g(e_L) < \hat{y}$ is necessary for $g(e_H) < g(e_L)$. From $(A13)$ it is evident that $g(e_H) < g(e_L)$ is more likely as $b \to 0^+$. Instead $(A13)$ when $b \to \mu^-$ then $g(e_H) > g(e_L)$ is more likely. ■

The following lemma is invoked in the Proposition and is key to Proof that at the optimum goals decrease in bonuses. This result is originally due to Chan et al. (1990).

Lemma 5. Assumption 2 implies that $F(y|e_H)$ is more convex than $F(y|e_L)$.

Proof. The distribution $F(y|e_H)$ is more convex than $F(y|e_L)$ if $F(y|e_H)F(y|e_L)^{-1}$ is a convex function in [0,1]. The function $F(y|e_H)F(y|e_L)^{-1}$ is convex if and only if $\frac{F(y|e_L)^{-1}f(y|e_L)}{F(y|e_L)^{-1}f(y|e_H)}$ is increasing in y. To see how, compute the derivative of $F(y|e_H)F(y|e_L)^{-1}$ with respect to y to obtain

$$\frac{F(y|e_L)^{-1}f(y|e_H)}{F(y|e_L)^{-1}f(y|e_L)} + F(y|e_H)F(y|e_L)^{-1}. \qquad (A15)$$

Thus, that the second derivative of $F(y|e_H)F(y|e_L)^{-1}$ is positive requires that $\frac{F(y|e_H)^{-1}f(y|e_L)}{F(y|e_H)^{-1}f(y|e_H)}$ increases in y. Since, $F(y|e_H)^{-1}$ is an increasing function, then it is sufficient and necessary that the ratio $\frac{f(y|e_H)}{f(y|e_L)}$ increases in y. This is equivalent to the monotone likelihood ratio property (Assumption 2). ■

Proposition.

Proof. According to Eq. $(A2)$, the self-chosen goal compatible with exerting high effort is equal to:

$$b\big(1 - F(g(e_H)|e_H)\big) - bg(e_H)f(g(e_H)|e_H) - \mu - \mu(\lambda - 1)F(g(e_H)|e_H) = 0. \qquad (A16)$$

Similarly, the self-chosen goal compatible with exerting low effort is equal to:

$$b\big(1 - F(g(e_L)|e_L)\big) - bg(e_L)f(g(e_L)|e_L) - \mu - \mu(\lambda - 1)F(g(e_L)|e_L) = 0. \qquad (A17)$$

We find an expression of $g(e_H)$ as a function of $g(e_L)$ using Eq. $(A16)$ and Eq. $(A17)$. We do so by obtaining an expression for $\mu - b$ for each equation, and equating those expressions to get:

$$\big(b + \mu(\lambda - 1)\big)\big(F(g(e_L)|e_L) - F(g(e_H)|e_H)\big) - b\big(g(e_H)f(g(e_H)|e_H) - g(e_L)f(g(e_L)|e_L)\big) = 0. \quad (A18)$$

Equation $(A18)$ is rewritten to obtain:

$$g(e_H) = \left(1 + \frac{\mu(\lambda - 1)}{b}\right)\left(\frac{F(g(e_L)|e_L) - F(g(e_H)|e_H)}{f(g(e_H)|e_H)}\right) + \frac{g(e_L)f(g(e_L)|e_L)}{f(g(e_H)|e_H)}. \qquad (A19)$$

Denote by $g^*(e_H)$ and $g^*(e_L)$ the goal levels that satisfy $(A19)$. Moreover note that Eq. $(A19)$ shows that by $g^*(e_H) \geq g^*(e_L)$ since $F(g^*(e_L)\,|e_L) - F(g^*(e_H)|e_H) \geq 0$ (Assumption 2), $b > 0, \lambda > 1$, and $\mu > 0$. Lemma 4 shows that a necessary condition for $g^*(e_H) \geq g^*(e_L)$ is $g^*(e_L) > \hat{y}$. Hence, at the optimum it must be that $g^*(e_L) \geq \hat{y}$ and, therefore, that $g^*(e_H) > 0$.

To evaluate the impact of bonuses on goals we implicitly differentiate $(A18)$ and obtain

$$\frac{dg^*(e_H)}{db} = \frac{F(g^*(e_L)|e_L) - F(g^*(e_H)|e_H) - \big(g^*(e_H)f(g^*(e_H)|e_H) - g^*(e_L)f(g^*(e_L)|e_L)\big)}{\big(2b + \mu(\lambda - 1)\big)\big(f(g^*(e_H)|e_H) - f(g^*(e_L)|e_L)\big) + b\big(g^*(e_H)f'(g^*(e_H)|e_H) - g^*(e_L)f'(g^*(e_L)|e_L)\big)}. \quad (A20)$$

Note that under $g^*(e_L) > \hat{y}$ then $f(g^*(e_H)|e_H) > f(g^*(e_L)|e_L)$. Moreover, Lemma 5 shows that Assumption 2 implies $f'(g^*(e_H)|e_H) > f'(g^*(e_L)|e_L)$ for all y. Hence, the denominator of Eq. $(A20)$ is positive. Moreover, the numerator of Eq. $(A20)$ is weakly positive if:

$$g^*(e_H) \leq \frac{F(g^*(e_L)|e_L) - F(g^*(e_H)|e_H)}{f(g^*(e_H)|e_H)} + \frac{g^*(e_L)f(g^*(e_L)|e_L)}{f(g^*(e_H)|e_H)}. \quad (A21)$$

Eq. $(A19)$ shows that the inequality in $(A21)$ cannot hold at the optimum if $\mu > 0$ and $b > 0$. Thus, $\frac{dg^*(e_H)}{db} < 0$.

Finally, Lemma 3 states that $\frac{\partial IC(g^*(e))}{\partial g^*(e)} > 0$ if $g^*(e_L) > \hat{y}$ and if $\lambda > \hat{\lambda}$, i.e. a sufficiently large level of loss aversion. Since $\frac{dg^*(e_H)}{db} < 0$ for any λ, then higher bonuses yield a lower affordability of high effort if $\lambda > \hat{\lambda}$ through their adverse effect on goals.

Corollary 1

Proof. Let $b = 0$. Equation $(A19)$ implies that $g^*(e_H) = \bar{y}$. In the absence of bonuses, a goal consistent with high effort is the maximum possible goal. That equation also demonstrates that $g^*(e_H) > g^*(e_L)$. From Lemma 4 we know that $g^*(e_H) \geq g^*(e_L)$ if $g^*(e_L) > \hat{y}$. Hence, at the optimum it must be that $g^*(e_L) \in (\hat{y}, \bar{y}]$.

Under $b = 0$, (6) becomes

$$IC\colon\ (a + \mu)\big(\mathbb{E}(y|e_H) - \mathbb{E}(y|e_L)\big) + \mu(\lambda - 1)\int_0^g F(y|e_L) - F(y|e_H)dy > c. \quad (A22)$$

Assumption 2 implies $F(y|e_L) \geq F(y|e_H)$ and $\mathbb{E}(y|e_H) > \mathbb{E}(y|e_L)$. Therefore $\frac{dIC}{dg} > 0$. Higher goals always lead to higher effort in the absence of bonuses. Note that Eq. $(A22)$ is largest whenever $g(e_H) = \bar{y}$ since $F(y|e_L) - F(y|e_H) \geq 0$ and $g \geq 0$ so $\int_0^g F(y|e_L) - F(y|e_H)dy$ attains its maximum value when g is largest. As a result, high effort can be implemented for larger values of c in the absence of bonuses.

Corollary 2.

Proof . Let $\mu = 0$. Equation $(A18)$ becomes,

$$F(g(e_L)|e_L) - F(g(e_H)|e_H) - \big(g(e_H)f(g(e_H)|e_H) - g(e_L)f(g(e_L)|e_L)\big) = 0. \quad (A23)$$

So, Eq. $(A23)$ can be rewritten as:

$$g(e_H) = \frac{F(g(e_L)|e_L) - F(g(e_H)|e_H)}{f(g(e_H)|e_H)} + \frac{g(e_L)f(g(e_L)|e_L)}{f(g(e_H)|e_H)}. \qquad (A24)$$

The optimal goal $g^*(e_H)$ satisfying Eq. $(A24)$ is non-negative since $F(g^*(e_L)|e_L) - F(g^*(e_H)|e_H) \geq 0$ (Assumption 2) and $g(e_L) \in [0, \overline{y}]$. Furthermore, Eq. $(A24)$ demonstrates that $\frac{dg^*(e_H)}{db} = 0$.

Under $\mu = 0$, Eq. (6) evaluated at $g^*(e_H)$ becomes

$$IC\colon\ a\big(\mathbb{E}(y|e_H) - \mathbb{E}(y|e_L)\big) + bg^*(e_H)\big(1 - F(g^*(e_H)|e_H)\big) - bg^*(e_L)\big(1 - F(g^*(e_L)|e_L)\big) \geq c \quad (A25)$$

Differentiating $(A25)$ with respect to b gives

$$\frac{dIC(g^*(e_H))}{db} = \ g^*(e_H)\big(1 - F(g^*(e_H)|e_H)\big) - g^*(e_L)\big(1 - F(g^*(e_L)|e_L)\big) \qquad (A26)$$

Assumption 2 implies $1 - F(g^*(e_H)|e_H) \geq 1 - F(g^*(e_L)|e_L)$, and Eq. $(A24)$ demonstrates that $g^*(e_H) \geq g^*(e_L)$ even if $\mu = 0$. Therefore $\frac{dIC(g^*(e_H))}{db} > 0$. As a conclusion, while higher bonuses do not influence optimal goals when $\mu = 0$, $\frac{dg^*(e_H)}{db} = 0$, they do increase the likelihood of attaining high effort $\frac{dIC(g^*(e_H))}{db} \geq 0$.

Appendix B. Details of the Risk Attitude Classifications

In this appendix, we report some additional analysis of the data from part B of the experiment, in which utility curvature and loss-aversion are measured. These elicited data consist of a vector of positive certainty equivalents $\{x_1, x_2, x_3, x_4, x_5\}$ for five different lotteries in the domain of gains, and the vector of offsetting loss outcomes $\{z_1, z_2, z_3, z_4, z_5\}$ for each participant. These values can be analyzed to understand (1) the risk attitudes of participants when outcomes are restricted to the gain domain, (2) whether risk attitudes have a sign-dependent component as proposed in prospect theory, and (3) how loss averse a participant is.

B.1. Risk attitudes in the domain of gains

We begin by studying the elicited sequence $\{x_1, x_2, x_3, x_4, x_5\}$ which informs us about the risk attitudes of participants in the domain of gains. We classify each participant according to their risk attitude. To that end we compute the difference $\Delta_{ij} \equiv x_{ij} - EV_j$, where the index j indicates the lottery number, and EV_j stands for the expected value of that lottery, that is, $EV_j = 0.5H_j + 0.5L_j$. The sign of Δ_{ij} is a non-parametric measure of the risk attitude of participant i with respect to lottery j. If the participant exhibits $\Delta_{ij} > 0$, the lowest price at which she is willing to sell the lottery is larger than its expected value, denoting a risk-seeking attitude. If $\Delta_{ij} < 0$, the participant has a risk averse attitude toward that lottery. Also, whenever $\Delta_{ij} = 0$, the participant is risk neutral.

We perform a classification of participants based on the statistical significance of their elicited Δ_{ij}. We compute confidence intervals around zero to determine whether a Δ_{ij} is statistically relevant given the overall variation in the data. Specifically, we calculate the standard deviation of $\sum_i \Delta_{ij}$ for each lottery j, and multiply it by the factors 0.64 and -0.64. A significantly positive Δ_{ij} indicates that participant i is risk seeking with respect to lottery j, while a significantly negative Δ_{ij} indicates risk aversion. Under the assumption that the data follow a normal distribution, approximately 50% of the data must lie within this confidence interval. Furthermore, to account for response error, we classify a participant to have a risk averse attitude when at least four of her Δ_{ij}s are negative. A participant is risk seeking when at least four of her Δ_{ij}s are positive, and a participant has linear utility when at least four Δ_{ij}s are not different from zero. This is also the approach followed by Abdellaoui (2000) and Abdellaoui et al. (2008).

Table C.1 shows that, by this criterion, the majority of participants, 57%, are classified as having linear utility. A proportion test rejects the hypothesis that the fraction is 0.5, p=0.049, indicating that a significant majority has linear utility, 40% of participants are classified as having concave utility, and only five individuals (3%) as having convex utility.

The second analysis we conduct on these data assumes that the utility function of participants follows a particular functional form. We fit the certainty equivalents to these functionals to examine the participants' risk attitudes. Specifically, we assume that the utility of participants is of the power utility form, $x_{ij} = {EV_j}^{\alpha}$, which belongs to the CRRA family, or of the exponential utility form, $x_{ij} = 1 - \exp(-\rho EV_j)$, which belongs to the CARA family. We estimate the parameters of these functionals, using the non-linear least squares method, for the pooled data of all participants.

Table B.1 Classification of individuals' risk attitudes towards lotteries with gains

Lottery/Classification	Concave Utility	Convex Utility	Linear Utility
Lottery 1	60	24	77
Lottery 2	80	15	66
Lottery 3	98	16	47
Lottery 4	77	9	75
Lottery 5	100	11	50
Total num. participants	**65**	**5**	**91**

Note: This table presents the classification of individuals according to their risk attitudes in the domain of gains. Each row presents the number of participants classified as having concave, convex or linear utility, which is equivalent to saying that they are risk averse, risk seeking, or risk neutral, respectively. A participant i is classified as having concave utility for lottery j whenever the difference $\Delta_{ij} \equiv x_{ij} - EV_j < 0$. A participant i is classified as having convex utility for lottery j whenever the difference $\Delta_{ij} > 0$. A participant i is classified as having linear utility for lottery j whenever the difference Δ_{ij} is not significantly different from zero. The last row presents the number of participants classified as having concave, convex, and linear utility over all lotteries. A participant is classified to have concave utility if she displays risk averse attitudes toward at least four lotteries. A participant is classified to have convex utility if she displays risk seeking attitudes toward at least four lotteries. Having risk neutral attitudes for at least four lotteries classify a participant as having linear utility.

Table B.2 presents the estimates of the parameters. We find that the estimate $\hat{\alpha}$ is significantly less than one, though only modestly so. This implies a utility function for a representative agent that has slightly risk averse attitudes (F(1,160)=127.651, p<0.001). Similarly, we find that the estimate $\hat{\rho}$ is significantly greater than zero, though still fairly small in magnitude. This also reveals a small degree of risk aversion on the part of the representative individual (F(1,160)= 3.4e+05, p<0.001). Thus, we find that participants in aggregate display moderate risk aversion when functional forms with CARA and CRRA are estimated. This is in line with the statistical analysis of the data at the individual level, which found that the majority of participants were risk-neutral, but that there was also a share of participants more appropriately classified as risk averse.

Table B.2 Estimates of parameters of the utility function of the representative participant

Parametric form	**Coefficient**	**St. Error**	R^2
$x_{ij} = EV_j^{\alpha}$	$\alpha =.9480115$	.0048	0.946
$x_{ij} = \dfrac{1 - \exp(-\rho EV_j)}{\rho}$	$\rho =. 01854$	.0016	0.942

B.2 Loss aversion

Next, we analyze the sequence of negative outcomes $\{z_1, z_2, z_3, z_4, z_5\}$ that made the participants indifferent between receiving zero for sure and a lottery consisting of $(z_j, 0.5; x_j, 0.5)$, where x_j was an elicited certainty equivalent of one of the lotteries containing only positive outcomes. The analysis of this data reveals the degree of a participant's loss aversion. The measure of loss aversion is the coefficient $\lambda_j \equiv x_j/z_j$. When the λ_i coefficient takes the value of one, the participant is indifferent between accepting and declining a lottery that consists of a gain and a loss of equal magnitude, each occurring with probability 0.5, and thus the participant exhibits no loss aversion or gain seeking. If the coefficient λ_i takes on a value larger than one, it indicates the presence of loss aversion.

We first analyze the loss aversion coefficients at the individual level. A participant is classified as loss averse when at least four (out of five) of her loss aversion coefficients satisfy $\lambda_i > 1$. A participant is classified as gain-seeking when at least four (out of five) of her loss aversion coefficients have $\lambda_i < 1$. Finally, a participant is classified as having mixed attitudes toward losses if she cannot be classified as either loss averse or gain seeking. Table B.3 shows that the large majority of participants is loss averse. Specifically, 72% of participants are classified as loss averse and 14% as gain-seeking.

Table B.3 Individuals classified as loss averse for each x_j

Classification	**Loss Averse**	**Gain-Seeking**	**Mixed**
Lottery 6	106	55	-
Lottery 7	125	36	-
Lottery 8	128	33	-
Lottery 9	128	33	-
Lottery 10	124	37	-
Total	**117**	**23**	**21**
Average	**134**	**27**	**0**

We also perform a second analysis of these data featuring the average of the loss aversion coefficient that an individual exhibits over all five lotteries. The last row in Table B.3 shows that according to this analysis, 137 participants, or 85% of participants, are loss averse, and the remaining 27 participants are gain-seeking. Thus, both analyses conclude that the great majority of the participants is loss averse.

Appendix C. Additional Tables and Analyses

Table C.1 Descriptive statistics of goals set by round in the GOAL+BONUS and GOAL treatments

	round 1	**round 2**	**round 3**	**round 4**	**round 5**	**round 6**
GOAL+BONUS						
mean	6.076	5.948	6.179	6.205	6.461	6.615
median	6	5	6	6	6.5	6
S.D.	2.056	1.986	2.186	2.142	1.889	2.034
GOAL						
mean	8.658	8.024	7.902	8	7.829	8.414
median	7	7	7	8	7	8
S.D.	5.803	4.470	4.409	4.494	4.510	4.460

Table C.2 Effect of treatment and continuous loss aversion level

	(1)	(2)
	Performance	**Goal Level**
GOAL* Loss Aversion	0.018***	0.025***
	(0.007)	(0.006)
GOAL* No Loss aversion	-0.243	-0.350***
	(0.182)	(0.191)
Loss Aversion	-0.012***	-0.013**
	(0.005)	(0.005)
No Loss Aversion	0.207	0.349**
	(0.1548)	(0.163)
Constant	3.825***	3.780***
	(0.021)	(0.022)
Log-Likelihood	-396.611	-496.029
N	80	80

Note: This table presents the estimates of the Poisson regression of the specification $\text{Performance}_i = \beta_0 + \beta_1 \text{GOAL} * \text{Loss Aversion} + \beta_2 \text{GOAL} * \text{No Loss Aversion} + \beta_3 \text{Loss Aversion} + \beta_4 \text{ No Loss Aversion} + \varepsilon_i$ with $\varepsilon_i \sim \text{Poisson}(\omega)$, as well as the Poisson regression of the statistical model $\text{Goal level}_i = \beta_0 + \beta_1 \text{GOAL} * \text{Loss Aversion} + \beta_2 \text{GOAL} * \text{No Loss Aversion} + \beta_3 \text{Loss Aversion} + \beta_4 \text{ No Loss Aversion} + \varepsilon_i$ with $\varepsilon_i \sim \text{Poisson}(\omega)$. "Performance" is the number of tables a participant correctly solves over all rounds and "Goal setting" the sum of the goals set by the participant over all rounds. Participants were assigned either to the GOAL or GOAL+BONUS treatment, where the latter is the benchmark category for the regression. ``Loss Aversion'' is a variable constructed as $\max(\lambda_i - 1, 0)$ and ``No Loss Aversion'' is a variable constructed as defined as $\min(\lambda_i - 1, 0)$ Standard errors are in parenthesis. * indicates a p-value <0.1, ** indicates a p-value <0.05, ***indicates a p-value <0.01.

Table C.3 Effect of treatment and loss aversion on performance and goals for participants with linear utility only

	(1)	(2)	(3)	(4)
	Performance	**Performance**	**Goal Level**	**Goal Level**
GOAL* Loss Averse	0.412***		0.559***	
	(0.073)		(0.073)	
GOAL* High Loss Averse		0.213***		0.476***
		(0.061)		(0.062)
Loss Averse	0.260***		0.134*	
	(0.075)		(0.077)	
High Loss Averse		0.067		0.094
		(0.065)		(0.069)
GOAL	0.247*	0.241***	0.218***	0.373***
	(0.094)	(0.071)	(0.095)	(0.073)
Constant	3.517***	3.670***	3.505***	3.550***
	(0.065)	(0.048)	(0.065)	(0.051)
Log-Likelihood	-203.44	-437.376	-278.999	-288.899
N	44	44	44	44

Note: This table presents the estimates of the Poisson regression of the specification $\text{Performance}_i = \beta_0 + \beta_1 \text{GOAL} * \text{Loss Averse} + \beta_2 \text{GOAL} + \beta_3 \text{LOPR} + \beta_4 \text{LOPR} + \beta_5 \text{Loss Averse} + \varepsilon_i$ with $\varepsilon_i \sim \text{Poisson}(\omega)$, as well as the Poisson regression of the statistical model $\text{Goal level}_i = \beta_0 + \beta_1 \text{GOAL} * \text{Loss Averse} + \beta_2 \text{GOAL} + \beta_3 \text{LOPR} + \beta_4 \text{LOPR} + \beta_5 \text{Loss Averse} + \varepsilon_i$ with $\varepsilon_i \sim \text{Poisson}(\omega)$. "Performance" is the total number of tables a participant solves correctly over all rounds and "Goal setting" is the sum of the goals set by the participant over all rounds. Participants were assigned either to the GOAL, HIPR, LOPR, or GOAL+BONUS treatment. GOAL+BONUS is the benchmark category for the regression. ``Loss Averse'' is a dummy variable that captures whether a participant is loss averse or not. A participant is loss averse when at least four variables $\lambda_j$, where $\lambda_j \equiv x_j/z_j$, are greater than one, ``Loss averse Mild' equals 1 if a participant is loss averse and her average λ is lower than the median participant in the sample, and 0 otherwise. ``Loss averse High' equals 1 if a participant is loss averse and her average λ is lower than the median participant in the sample and 0 otherwise. All models include only participants classified as having linear utility. Standard errors presented in parenthesis. * indicates a p-value <0.1, ** indicates a p-value <0.05, ***indicates a p-value <0.01.

Figure C.1 Effect of treatment and loss aversion on performance and goals for participants with linear utility only

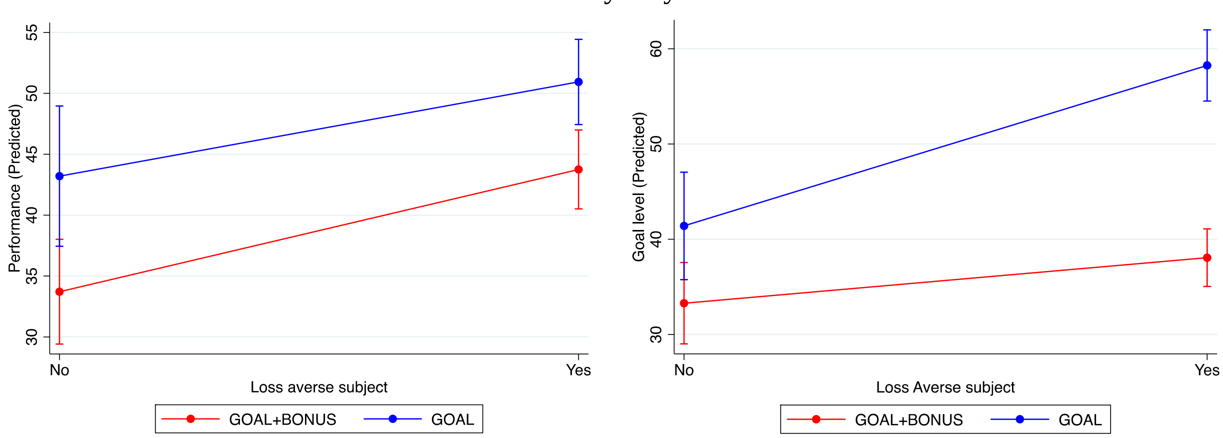
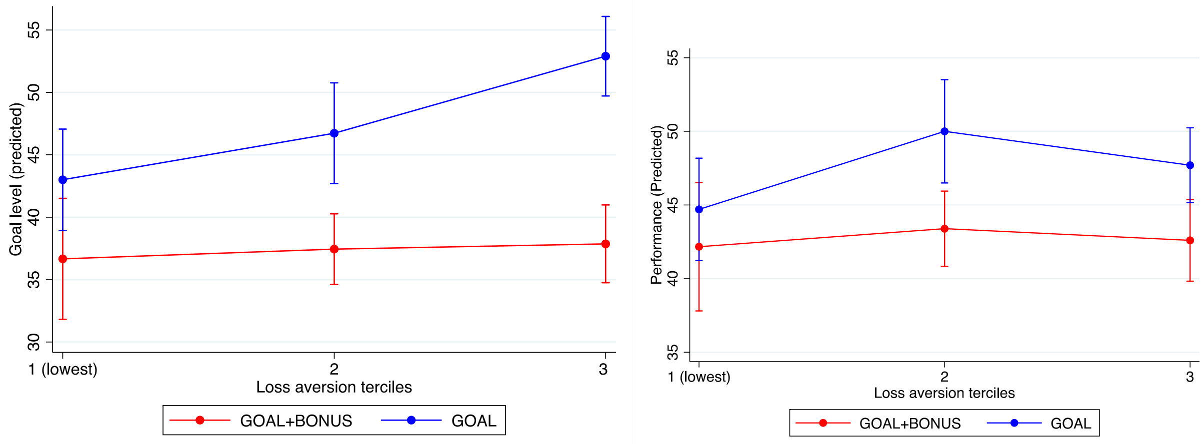


Table C.4 Effect of treatment and loss aversion terciles on performance and goals

	(1) **Performance**	(2) **Goal Level**
GOAL* Loss Aversion Highest tercile	0.123**	0.366***
	(0.070)	(0.074)
GOAL* Loss Aversion Middle tercile	0.170**	0.242***
	(0.075)	(0.061)
GOAL* Loss Aversion Lowest tercile	0.058	0.159*
	(0.078)	(0.082)
Loss Aversion Highest tercile	0.010	0.032
	(0.074)	(0.032)
Loss Aversion Middle tercile	0.028	0.020
	(0.072)	(0.077)
Constant	3.741***	3.601***
	(0.062)	(0.067)
Log-Likelihood	-395.325	-468.352
N	80	80

Note: This table presents the estimates of the Poisson regression of the specification $\text{Performance}_i = \beta_0 + \beta_1 \text{GOAL} * \text{Concave Utility} + \beta_2 \text{GOAL} + \beta_5 \text{Loss Averse} + \mathcal{E}_i$ with $\mathcal{E}_i \sim \text{Poisson}(\omega)$, as well as the Poisson regression of the statistical model $\text{Goal level}_i = \beta_0 + \beta_1 \text{GOAL} * \text{Concave Utility} + \beta_2 \text{GOAL} + \beta_5 \text{Loss Averse} + \mathcal{E}_i$ with $\mathcal{E}_i \sim \text{Poisson}(\omega)$. "Performance" is the number of tables a participant correctly solves over all rounds and "Goal setting" the sum of the goals set by the participant over all rounds. Participants were assigned either to the GOAL or GOAL+BONUS treatment, where the latter is the benchmark category for the regression. ``Loss Averse'' is a dummy variable that indicates whether a participant is loss averse or not. A participant is loss averse when at least four of her $\lambda_j$, where $\lambda_j \equiv x_j/z_j$, are greater than one. ``Concave Utility" is a dummy variable that equals 1 if a participant exhibits concave utility and zero otherwise. A participant exhibits concave utility when at least four of her variables Δ_j, where $\Delta_j \equiv x_j - EV_J$, are less than zero. Standard errors are in parenthesis. * indicates a p-value <0.1, ** indicates a p-value <0.05, ***indicates a p-value <0.01.

Figure C.2 Effect of treatment and loss aversion terciles on performance and goals



Appendix D. Experimental Instructions

D.1 Welcome

This is an experiment in the economics of decision-making. The instructions are simple and if you follow them and make good decisions, you might earn a considerable amount of money, which will be paid to you at the end of the experiment. The amount of payment you receive depends on your decisions, your effort, and partly on luck. The currency used throughout the experiment is Dollars.

Once the experiment has started, no one is allowed to talk to anybody other than the experimenter. Anyone who violates this rule will lose his or her right to participate in this experiment. If you have further questions when reading these instructions please do not hesitate to raise your hand and formulate the question to the experimenter.

The experiment consists of two parts. A **counting task** part and a **decision-making** part. Both tasks will count towards your final earnings.

D.2 Part A: Counting task

This part of the experiment consists of a sequence of **6 rounds** of **5 minutes** each. In each round you need to complete as many tasks as possible. A task is completed when you count the correct number of zeros in a table that contains 100 zeros and ones.

As soon as you know the correct number of zeros contained in a table, you have to input your answer using the keyboard. Once you have entered the number, click "Next". Immediately afterwards a new table will be displayed and, again, you have to count the number of zeros in this new table. This procedure is repeated until the time is up. A timer is displayed in the upper part of your screen. After each round is over you receive feedback about your performance in that round.

Counting Tips: Of course you can count the zeros in any way you want. Speaking from experience, however, it is helpful to always count two zeros at once and multiply the resulting number by two at the end. In addition, you miscount less frequently if you track the number you are currently counting with the mouse cursor.

you will see an example when you press "Next".

(example is displayed)

Payments

(LOPR treatment)

For each correct task that you complete you receive 0.20 Dollars, but if your answer was wrong you receive nothing.

The formula we use to calculate your earnings is Earnings = # correct tasks * 0.20 Dollars.

(HIPR treatment)

For each correct task that you complete you receive 0.50 Dollars, but if your answer was wrong you receive nothing.

The formula we use to calculate your earnings is Earnings = # correct tasks * 0.50 Dollars.

(GOAL treatment)

For each correct task that you complete you receive 0.20 Dollars, but if your answer was wrong you receive nothing.

Goal: At the beginning of each round you will be asked to set a goal in terms of the number of tables you aim at solving correctly in that specific round. Provide this target at this best ability, we would like to see how accurate is your prediction

The formula we use to calculate your earnings is Earnings = # correct tasks * 0.20 Dollars

(GOAL+BONUS treatment)

For each correct task that you complete you receive 0.20 Dollars, but if your answer was wrong you receive nothing.

Goal: At the beginning of each round you will be asked to set a goal in terms of the number of tables you aim at solving correctly in that specific round. If you achieve your goal or you surpass it, you will be paid an additional bonus. The bonus is larger the large the goal is set. If your goal is high and you achieve it or surpass you will be given a high monetary reward. But, if your goal is low and you accomplish it or surpass it you will be given a low monetary reward. Also be aware that if you set a very high goal and you cannot accomplish it you will get no bonus.

The formula we use to calculate your earnings is Earnings = # correct tasks * 0.20 Dollars + goal * 20 cents.

Are you ready to start now? As soon as you press "Next" the task will start.

D.3 Part B: decision-making part

the following, you will face a series of decisions. Your task in each decision is to choose among two possible alternatives. Your earnings on this part of the experiment depend on your choices.

You will be faced with 10 decision sets. Each decision set contains several choices. In each of decision you need to choose between the option R, that delivers a sure amount of money, and the option L that results in one of two different monetary amounts.

Note that in each decision set you need to choose between L and R multiple times. But, you need to be careful since the offered amounts of money could change from one decision to the next.

you will see an example when you press "Next"

Payments

At the end of this part of the experiment **one** randomly chosen decision will be played and paid. Hence, a random number chosen by the program will be drawn to determine which decision counts towards your earnings.

This means that each choice that you make might be chosen and paid.
If it is clear what you have to do in this part of the experiment, press "Next" to start.

D.4 Exit Questionnaire.

This is the last part of the experiment. Please answer the following questions at your best ability.

Enter the computer (Letter plus digit) you are at.
Enter your age.
What is the program you are studying?
Enter your gender.
Male Female
What is your nationality?