

Paying with Self-Chosen Goals: Incentives and Gender Differences*

Patricio S. Dalton[†] Victor H. Gonzalez[†] Charles N. Noussair[†]

March 19, 2015

Abstract

To boost employees' performance, firms often offer monetary bonuses when production goals are reached. However, the evidence suggests that the particular level of a goal is critical to the effectiveness of this practice. Goals must be challenging yet achievable. Computing optimal goals when employees have private information about their own abilities is often not feasible for the firm. To solve this problem, we propose a compensation scheme in which workers set their own production goals. We provide a simple model of self-chosen goals and test its predictions in the laboratory. The evidence we find in the laboratory confirms our model's predictions for men, but not for women. Men exert greater effort under the self-chosen goal contract system than under a piece rate contract. In contrast, women perform worse under the self-chosen goal contract. Further analysis suggests that this is because women fail to set goals that are challenging enough, because they are less likely to update their goals to take into account their improving performance as they repeat the task.

JEL: C91, C92, J16, J24.

Keywords: Contracts, Bonus, Endogenous Goals, Productivity, Intrinsic Motivation, Challenge Seeking, Gender Differences.

*We are grateful for helpful comments to participants in seminars at Tilburg University, the French Experimental Economists Association (ASFEE) conference, the Tilburg Institute of Economic and Behavioral Research (TIBER) symposium, the Economic Science Association (ESA) meetings and the Nordic conference on Behavioral and Experimental Economics.

[†]Tilburg University, Department of Economics and CentER, Warandelaan 2, 5037 AB, Tilburg, The Netherlands. E-mail: p.s.dalton@uvt.nl, v.h.gonzalezjimenez@uvt.nl and c.n.noussair@uvt.nl.

1 Introduction

Monetary bonuses for achieving performance milestones are used to incentivize employees in a wide range of industries, including finance, insurance, retailing (Banker et al., 2000), manufacturing (Enis, 1993), energy services (Rajagopalan, 1996) and charities (Baber et al., 2002). According to the last WorldatWork’s “Survey of Bonus Programs and Practices” more than eighty percent of American firms use at least some type of bonus program (WorldatWork, 2014).¹ On average, American firms pay bonuses to their executives equal to twenty-three percent of their base pay (WorldatWork, 2014).

The main theoretical rationale for bonuses is to stimulate performance. As long as standard conditions on preferences and costs of effort hold, a monetary bonus works as any pay-for-performance compensation: as the bonus increases, the effort exerted increases. Empirically, a large number of studies have shown that bonuses are positively correlated with employees’ performance (Groves et al., 1994; Baker et al., 1988; Banker et al., 2000; Enis, 1993; Jones and Kato, 1995; Kahn and Sherer, 1990). Furthermore, beyond providing monetary incentives, experiments also show that a bonus can also trigger gift-exchange considerations and induce some agents to expend more effort than under an incentive contract Fehr et al. (2007).

Typically, bonuses are awarded only if a certain performance target is reached. This threshold feature can give this target level the status of a reference point. Thus, reference-or goal-dependence can potentially be brought to bear to better understand the way bonuses affect performance. There is evidence that once an individual has a goal, she may undertake more effort on, exhibit greater persistence with, and direct more attention to, the endeavor (Locke and Latham, 1990; Locke, 1996; Locke and Latham, 2002). However, there is also evidence that the level of the goal is critical for its effectiveness: goals must be challenging yet achievable. Goals that are too easy or too difficult to attain are not effective. Wu et al. (2008) substantiate this intuition in a model in which individuals respond to exogenously-set goals. Their model predicts that performance is an inverted V-shaped function of the goal level, implying that there is an optimal *challenging yet attainable* goal that will boost performance to the maximum.

Computing optimal goals for employees is generally not feasible for a firm. If a principal does not have perfect information about her employees’ abilities, choosing an optimal goal is very difficult, if not impossible. To overcome this problem, we propose a compensation scheme in which workers set their own production goals. With the right incentives, workers set challenging goals that induce them produce as much as they can, given their own abilities.

¹According to Joseph and Kalwani (1998), 72% of the firms use bonuses to incentivize sales. See also Lemieux et al. (2009) who document evidence on the increasing use of bonuses in the US labor market.

This aligns the incentives of both employers and employees and solves the asymmetric information problem.

In our system, the agent chooses her own goal from a menu that is prespecified by the principal. The proposed menu of contracts provides sufficient monetary incentives for the agent to announce challenging yet attainable goals. The contract includes a monetary bonus for achieving the goal, and this bonus increases monotonically with the magnitude of the goal.²

In Section 2 we introduce a formal model of goal setting. The model builds on a standard piece rate contract and considers the consequences of the addition of a bonus for achieving a goal that is set by the agent himself. The model initially assumes preferences over only monetary outcomes and effort. The model is then extended to include goal-dependent preferences of a specific form, in which the goal serves as a reference point defining domains of gains and losses (see Goerg and Kube and Gómez-Miñambres et al. (2012)).

The model yields a number of testable predictions. Agents increase their output as the piece rate increases. They also increase their output if they are allowed to set their own goal. Output always strictly exceeds the goal. This last prediction is a consequence of goal-dependent preferences. In their absence, output is exactly equal to the goal.

As described in Section 3, we design a laboratory experiment to test the model’s predictions. In the experiment, participants engage in a real effort task, which consists of counting the number of zeros in a table with approximately 150 zeros and ones. Output is measured as the number of tables completed correctly.³ We compare three contracts with regard to the output they generate. Two of these are piece rate contracts: one with a relatively low piece rate (LOPR) and another with a high piece rate (HIPR). The third contract is a *self-chosen goal* contract (GOAL), which includes both a piece rate monetary compensation and a monetary bonus that is paid if and only if a goal is reached. The novelty of this contract is twofold. First, the size of the bonus increases monotonically with the goal, i.e. the higher the production goal, the higher the monetary bonus paid if the bonus is achieved. In this way, setting ambitious goals is incentivized monetarily. The second novel feature has to do with who sets the goal. In our setting, rather a manager setting the goal for the worker, it is the latter who sets the goal for himself.

The experimental results are reported in Section 4. We find that there are sharp gender differences in the way in which subjects respond to the incentives of the contracts. First, either allowing agents to set their own goals, or specifying more high-powered marginal

²There is evidence that people tend to overstate their abilities if the goal is not monetarily incentivized (Goerg and Kube). This provides further justification for introducing monetary incentives in setting self-chosen goals.

³This task has been previously used by other scholars in the literature. See for example Abeler et al..

piece-rate incentives, induce men to produce higher output, as predicted by our model. However, they do not have the same effect on women, whose output is unaffected by these changes in compensation structure.

The gender differences observed in the lab cannot be explained by the model. However, we find evidence of a plausible explanation of the gender effect. We observe that women are relatively conservative when setting their goals in that their actual performance tends to exceed their goal by a larger margin than for men. Given that relatively modest goals translate into smaller bonuses, the conservative goal setting has a cost in terms of earnings. This behavior, in the first round, could be explained by greater risk aversion on the part of women than men. However, thereafter, we observe that women insufficiently update their goals as they gain experience, failing to internalize the trend of improvement in their performance. Male subjects, in contrast, tend to increase their goals from one period to the next, and thus tend, to a greater extent than women, to internalize and anticipate the learning acquired in the task over time. The consequence is lower goal setting on the part of females and therefore lower earnings in the set-your-own-goal condition.

Our paper is intended to contribute to several strands of literature. First, it contributes to the literature on mechanism design, originating in the work of Hurwicz (1973) and Mirrlees (1976), by offering a contract in which an incentive compatible elicitation of a production goal aligns the incentives of both principal and agent. Our contract allows the agent to self-select into a bonus scheme that is well-suited to her particular individual cost profile of effort. In this manner, the choice of one's own goal resembles the choice of one own linear contract (Laffont, 1994), by adapting the marginal incentives in the contract to the agents' type, increasing the payoff to both worker and employer.

Second, it adds to the literature on personnel economics (see Prendergast (1999) for a review) by proposing a novel pay-for-performance incentive scheme. There is a large literature on performance-payment schemes (e.g. Lazear (1986), Lazear (2000)) and bonus contracts, (Fehr et al., 1998, 2007), but the type of bonuses that this literature studies are exclusively those exogenously set by the firm.

Third, the paper contributes to the literature on goals as performance enhancers started by Locke (1996) and continued by Heath et al. (1999). This literature argues that goals act as reference points, and due to the properties of the prospect theory value function, they boost performance. Wu et al. (2008) show this intuition in a model in which individuals respond to exogenously-set goals. Gómez-Miñambres (2012) and Gómez-Miñambres et al. (2012) develop this idea further. Our extension is to endogenize the goals. Wu et al. (2008) and Gómez-Miñambres et al. (2012) show that under perfect information, an optimizing principal chooses a challenging but attainable goal for an average-ability worker, and that

this optimal goal increases worker’s performance.⁴

Gómez-Miñambres et al. (2012) test the predictions of their model in the laboratory, and to ensure that participants who act as principals have the right information about the average performance of workers, they choose participants who have previous experience in performing the task themselves. While the protocol of Gómez-Miñambres et al. (2012) for creating familiarity with the task on the part of the principal is feasible in the laboratory, it may not be practical to apply in many organizations, where managers may not be able to have hands-on experience with the task workers have to perform. In such contexts, asking the manager to set challenging but attainable goals for each worker may be a difficult task.⁵

Fourth, a number of recent theoretical papers (e.g., Koch and Nafziger (2011), Koch and Nafziger and Hsiaw (2013)) have considered the effects of endogenous goals in attenuating self-control problems in other contexts. Like us, they assume that goals are endogenous reference points and that agents are loss averse. Our paper differs from this literature in two distinct ways. First, in this literature the goals are not incentivized with money. Second, the focus is on decision problems where time plays a key role. The key insight in this literature is that an increase in the goal level set today raises an individual’s motivation to work hard in the future. If the individual has present bias, he may be tempted to shirk in the future. However, if he is sophisticated enough, he can use a goal as a self-motivating device to attenuate future shirking. The role that a goal plays in this paper is different. Present bias is not a factor in our model because we work in a static framework and the motivation to set a goal emerges directly from the incentives of the contract. The individual sets an optimal goal in order to maximize his earnings.

The paper closest to ours is Goerg and Kube, who report an experiment that includes a treatment in which workers are paid a bonus conditional on reaching a pre-specified self-chosen goal. However, their contract is different from ours in at least two important respects. First, it doesn’t include the piece rate payment. They offer a fixed payment plus a bonus if the goal is reached, so workers have no monetary incentives to continue working once the goal is reached. Second, they consider a one shot setting. Our experimental protocol, in contrast, is a multi-period setting in which the worker can update his knowledge about his ability, allowing him to adapt his goal levels accordingly. With our design, we are able to

⁴Gómez-Miñambres (2012) also studies a principal-agent model, where the principal sets a goal for the agent, and the agent derives a sense of pride from accomplishing the goal. They show that the agent’s production and the goal set by the principal both increase with the agent’s personal standards.

⁵The challenge is further complicated by the fact that workers themselves have differing abilities and interact with each other. If the manager sets a goal based on his perception of each worker’s ability, then a goal becomes a signal which can itself be an extra motivator if the signal is good enough, as in Benabou and Tirole (2003). However, it can also be discouraging if it is low relative to the signal received by other workers, or to one’s expectations about one’s own ability.

study not only how self-chosen goals boost workers' performance, but also the dynamics of self-chosen goals in a context in which the worker learns about his ability. As Latham and Locke (1991) argue, goal setting is most effective when there is feedback showing progress in relation to the goal.

2 The model

In this section, we develop a theoretical framework for the analysis of self-chosen goals from which we derive the predictions that we test in the laboratory. We use elements of Gómez-Miñambres et al. (2012)'s theoretical framework, in which the principal chooses an optimal goal for the agent. However, in our model we study the case in which the principal leaves the decision of setting a goal and producing output entirely to the agent.

Consider an agent who chooses a level of production $y \geq 0$. The cost of production is increasing and convex, and for simplicity we assume $c(y) = \frac{(y/\theta)^2}{2}$. The parameter $0 \leq \theta \leq 1$ is interpreted as the agent's ability. We assume that θ is known by the agent.

2.1 The case of goal-independent preferences

We first consider the standard case of an agent with goal-independent preferences. Suppose the agent faces a piece rate compensation $w(y) = \alpha y$, where α is the compensation for each unit of output. Her utility is thus:

$$U(y) = \alpha y - \frac{(y/\theta)^2}{2} \quad (1)$$

Choosing y to maximize $U(y)$, the agent sets the optimal output under a piece rate contract $y_p^* = \theta^2 \alpha$.

Suppose now that in addition to the piece rate compensation, the principal introduces a bonus B for attaining an exogenously given output goal g . In that case, the payoff of the agent is given by

$$w(y, B(y, g)) = \alpha y + B(y, g) \quad (2)$$

where we assume a bonus of the form:

$$B(y, g) = \begin{cases} \beta g & \text{if } y \geq g \\ 0 & \text{if } y < g \end{cases} \quad (3)$$

$\beta > 0$ is a parameter indicating the monetary compensation for reaching the goal g . The employer gives a larger bonus when he requires a more ambitious goal, and the bonus is not awarded if the target level of output is not attained.

Under such a bonus scheme, the worker must choose whether to work toward the bonus or not. Thus, he must compare the payoff from the optimal level of output below the target level with that resulting from equaling or exceeding such target level. The optimal choice of production in the region below the target maximizes the utility function within that region. Denote the optimal output in this region as \underline{y} , and the resulting utility level as:

$$U(\underline{y}) = \alpha \underline{y} - \frac{(\underline{y}/\theta)^2}{2} \quad (4)$$

Similarly, the optimal output in the region of output above or equal the target level is equal to \bar{y} , and the corresponding utility level is :

$$U(\bar{y}) = \alpha \bar{y} + \beta g - \frac{(\bar{y}/\theta)^2}{2} \quad (5)$$

It is optimal not to try to achieve the goal, that is $U(\underline{y}, g) > U(\bar{y}, g)$, when the goal is set so high that it becomes too costly to attain. This is the case when $\bar{y} > \underline{y}$ and the marginal cost of achieving the additional output required to surpass the goal, $\frac{(\bar{y}/\theta)^2}{2} - \frac{(\underline{y}/\theta)^2}{2}$, is greater than the marginal benefit $\beta g + \alpha(\bar{y} - \underline{y})$. Also note that if $\frac{(\bar{y}/\theta)^2}{2} - \frac{(y_p^*/\theta)^2}{2} \leq \beta g + \alpha(\bar{y} - y_p)$, the worker exerts extra effort to attain the goal beyond the amount she would have exerted if there were no bonus in place. Therefore, a challenging and achievable exogenously set goal has the potential to increase output.

Generically, however, the principal will not have perfect information about the type of the worker. Hence, he will not know the level at which he should set the production (optimal) goal. For this reason, we consider the case in which the worker can set her own goal. We model the goal and effort as set simultaneously. The following proposition characterizes the optimal choice of goal and output level made by the agent. We show that when the worker can choose her own goal, she will set it at the optimal level of output and she will then work exactly as much as it is needed to achieve her goal

Proposition 1: *The agent chooses a goal equal to $g^* = (\alpha + \beta)\theta^2$ and exerts output $y^* = g^*$.*

Proof. We first show that in the optimum, it must be the case that the optimal output level is equal to the optimal goal chosen, i.e. $y^* = g^*$. Second, we show that $g^* = \theta^2(\alpha + \beta)$.

First we show that if $y \neq g$, either y or g is not optimal. Consider a given y_l such that $y_l < g$. Then $U(y_l, g) = \alpha y_l - \frac{(y_l/\theta)^2}{2}$. By reducing her goal and setting $g = y_l$ the agent can achieve a strictly greater payoff $U(y_l, y_l) = \alpha y_l - \frac{(y_l/\theta)^2}{2} + \beta y_l$. Now consider a given y_h such that $g < y_h$. The agent's payoff is $U(y_h, g) = \alpha y_h - \frac{(y_h/\theta)^2}{2} + \beta g$. Then, by increasing the

goal to $g = y_h$, the agents' payoff increases to $U(y_h, y_h) = \alpha y_h - \frac{(y_h/\theta)^2}{2} + \beta y_h$. Therefore, y^* must be equal to g .

The second step is to derive the optimal g . By the first step of the proof, we know that the agent will always work exactly as much as needed to receive the bonus. If so, her earnings are given by $(\alpha + \beta)g - \frac{(y/\theta)^2}{2} = (\alpha + \beta)y - \frac{(y/\theta)^2}{2}$. To derive the optimal goal, we consider the first order condition for the maximization of earnings with respect to output $y^* = (\alpha + \beta)\theta^2$, which directly implies the optimal goal: $g^* = \theta^2(\alpha + \beta)$

□

The optimal output level for both the piece rate and the self-chosen goal contracts are illustrated in Figure 1. The diagonal line indicates the marginal cost of output, $\frac{y}{\theta^2}$, with slope $\frac{1}{\theta^2}$. The horizontal lines represent the marginal benefit of output under each contract. Under a standard piece rate contract the marginal benefit expressed in terms of output is $\alpha\theta^2$. Under the endogenous bonus contract, the marginal benefit expressed in terms of output is $(\alpha + \beta)\theta^2$.⁶ From the figure, it is clear that increasing the piece rate α or adding a self-chosen bonus increases effort and thus output.

2.2 Goal-dependent preferences

Now suppose that the worker has goal-dependent preferences (Heath et al., 1999; Wu et al., 2008), which are represented by the following utility function:

$$U(y, g) = V_E(y, B(y, g)) + V_I(y, g) - \frac{(y/\theta)^2}{2} \quad (6)$$

The first term $V_E(y, B(y, g)) = w(y, B) = \alpha y + B(g, y)$ represents the utility the worker gets from receiving the monetary compensation. The second term represents the intrinsic (non-monetary) utility from attaining the goal (as in Köszegi and Rabin (2006), or Gómez-Miñambres et al. (2012)). We assume that V_I satisfies the properties of the value function of prospect theory (Kahneman and Tversky (1979)), and it is defined as follows:

$$V_I(y, g) = \begin{cases} (y - g)^{1/2} & \text{if } y \geq g \\ -\lambda(-(y - g))^{1/2} & \text{if } y < g \end{cases} \quad (7)$$

⁶An analogy that is perhaps useful in understanding Figure 1 is to long-run and short-run costs in producer theory. In the short-run, labor employed can be changed, and in the long run, capital can be adjusted as well. Here, we can consider the choice of output as a factor that can be varied in the short run, and the goal as a choice that cannot be adjusted in the short run. The horizontal line at $\alpha + \beta$ in Figure 1 represents a long run marginal payoff function, the upper envelope of short-run payoff functions for different fixed goal levels.

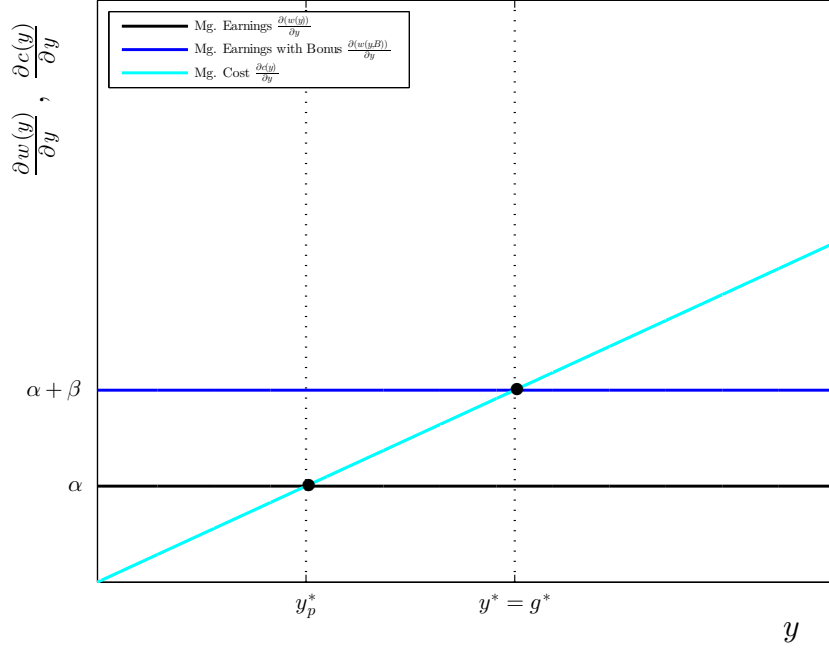


Figure 1: Optimal Output

The goal g acts as a reference point dividing the output space into gains, where the goal is attained or exceeded, and losses, where the goal is not attained. The parameter $\lambda > 1$ measures the degree of loss aversion. Because the exponent of $(y - g)$ is less than one, the function V_I also satisfies diminishing sensitivity. It is concave in the domain of gains and convex on the domain of losses.

Taking these conditions into account, the overall utility of the worker can be summarized as follows:

$$U(y, g) = \begin{cases} \alpha y + \beta g + (y - g)^{1/2} - \frac{(y/\theta)^2}{2} & \text{if } y \geq g \\ \alpha y - \lambda(-(y - g))^{1/2} - \frac{(y/\theta)^2}{2} & \text{if } y < g \end{cases} \quad (8)$$

The goal enters the utility function in two ways. On one hand, it enters positively as a conditional monetary bonus that is increasing in the magnitude of the goal. For this reason, the worker prefers higher to lower goals provided that the goal is attainable. Furthermore, the goal divides the output space into psychological loss and gain domains. Due to loss aversion ($\lambda > 1$), the marginal utility of output in the loss domain is higher than in the gain domain for a given absolute difference from the goal. Strict concavity in the domain of losses encourages the agent to work harder as she approaches her goal. The cost of effort, together with the ability of the worker and his degree of loss aversion, all influence the

optimal goal for a particular worker.⁷

Before we define the optimal choice of goal and output level of a worker with goal-dependent preferences, we note that given the strict convexity of the marginal utility function in the domain of losses and the linearity of the marginal cost of effort, it is possible to have multiple output levels $\{y_1^{**}, y_2^{**}, \dots, y_n^{**}\}$, in which the marginal utility equals the marginal cost of effort, for a given goal. We assume that the agent stops working at lowest effort level, i.e. $y^{**} = \min\{y_1^{**}, y_2^{**}, \dots, y_n^{**}\}$, at which one of the first order conditions (9) or (10), and the appropriate second order sufficient condition for a local optimum, $\frac{\partial^2 U(y, g)}{\partial^2 y} < 0$ is satisfied.

We now proceed to derive the optimal effort. Given the functional form of the utility function and its discontinuity at the reference point, we distinguish between two cases. Case 1 is when current output is below the goal and case 2 is when it exceeds the goal. Maximization of (8) for a given g leads to the following first order conditions.

$$\frac{\partial U(y, g)}{\partial y} = \alpha + \frac{1}{2}(y - g)^{-\frac{1}{2}} - \frac{y}{\theta^2} = 0 \quad \text{if } y > g \quad (9)$$

$$\frac{\partial U(y, g)}{\partial y} = \alpha + \frac{\lambda}{2}(g - y)^{-\frac{1}{2}} - \frac{y}{\theta^2} = 0 \quad \text{if } y < g \quad (10)$$

Equation (9) applies to the case in which $y > g$ and equation (10) governs the case when $y < g$. Note that because we assume loss aversion (i.e. $\lambda > 1$), failing to reach a goal by a given amount is more costly than the benefit of surpassing the goal by the same amount. In addition, output increases with the difficulty of the goal as long as the goal is attained or surpassed (i.e. $y > g$)⁸. However, effort decreases with goal difficulty conditional on the goal not being attained (i.e. $y < g$)⁹. In other words, goals and effort are complements in the domain of gains but substitutes in the domain of losses. This captures the idea that an agent's output is increasing in the level of the goal when it is attainable. It also implies that a challenging but attainable goal works better than goals that are either too easy or too difficult.¹⁰

From Figure 2, we see that the marginal utility of output is increasing when the goal is not reached, decreasing once the goal is reached, and tends to infinity asymptotically

⁷In the model of Gómez-Miñambres et al. (2012), the goal enters the utility function only through a psychological utility. We also allow the goal to enter the utility as a monetary bonus for two main reasons. First, although non-binding goals are interesting on their own, in this paper we want to focus on goals which yield monetary bonuses, since they are common practice in firms. Second, we wanted to guarantee that we provided incentives for our participants to truthfully reveal their goals in the experiment.

⁸Formally, $\frac{\partial^2 U(y, g)}{\partial y \partial g} = \frac{1}{4}(y - g)^{-\frac{3}{2}} > 0$ if $y > g$.

⁹Formally, $\frac{\partial^2 U(y, g)}{\partial y \partial g} = \frac{-\lambda}{4}(g - y)^{-\frac{3}{2}} < 0$ if $y < g$.

¹⁰This property is also present in the models of Wu et al. (2008) and Gómez-Miñambres et al. (2012).

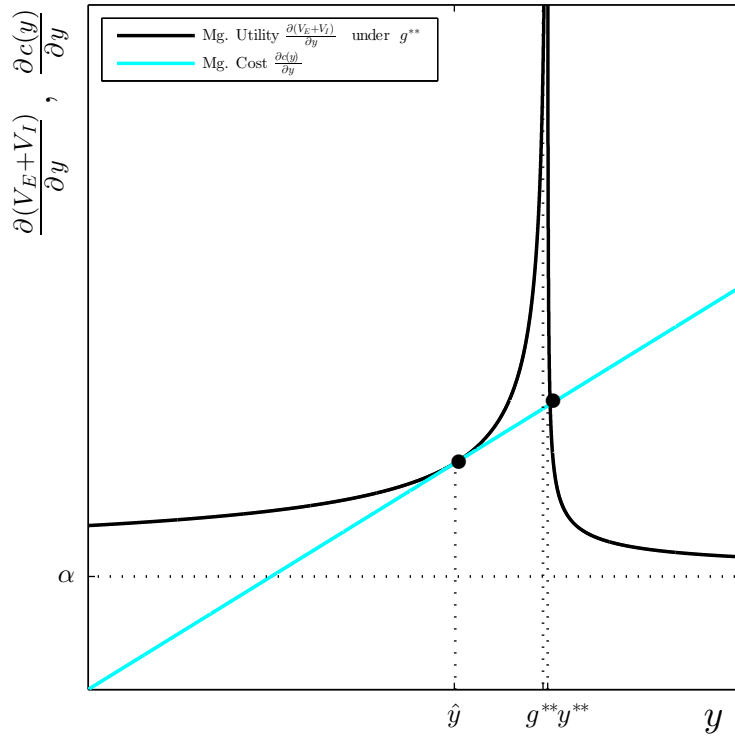


Figure 2: Optimal Output y^{**} when the goal is optimal g^{**}

as the goal is approached. This particular inverted-V shape of the marginal utility is due to two reasons. First, the goal acts as a reference point, dividing the space of outcomes into gains and losses. The second is diminishing sensitivity, in that the utility function is concave over gains but convex over losses.

Figure 3 illustrates that if the goal is set too high relative to the optimal goal, it can be detrimental to performance. Under such a goal, the agent will only produce output y_H . Beyond that level, marginal cost of effort exceeds marginal utility, which is relatively low due to the overambitious goal. Figure 4 illustrates that goals that are set too low are also counterproductive. The agent experiences an opportunity cost in terms of utility whenever he sets a goal that is not challenging enough, because the marginal utility of output decreases rapidly after the goal is attained.

Suppose that the agent sets the optimal goal from her perspective. We calculate this optimum in three steps. We first show that, at an optimum, (10) must be satisfied by the goal which induces a point of tangency between the marginal utility and the marginal cost of output, as shown in Figure 2. This condition is similar to those derived in Wu et al. (2008) and Gómez-Miñambres et al. (2012). Then, we use this condition to compute the optimal goal. Finally, we calculate the optimal output and show that it exceeds the optimal goal.

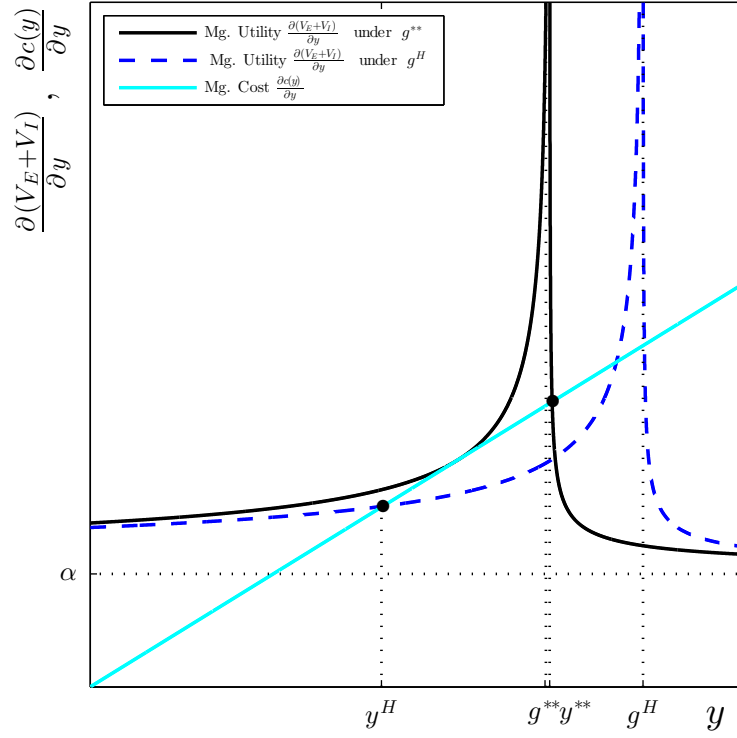


Figure 3: Output y^H when the goal is too ambitious g^H

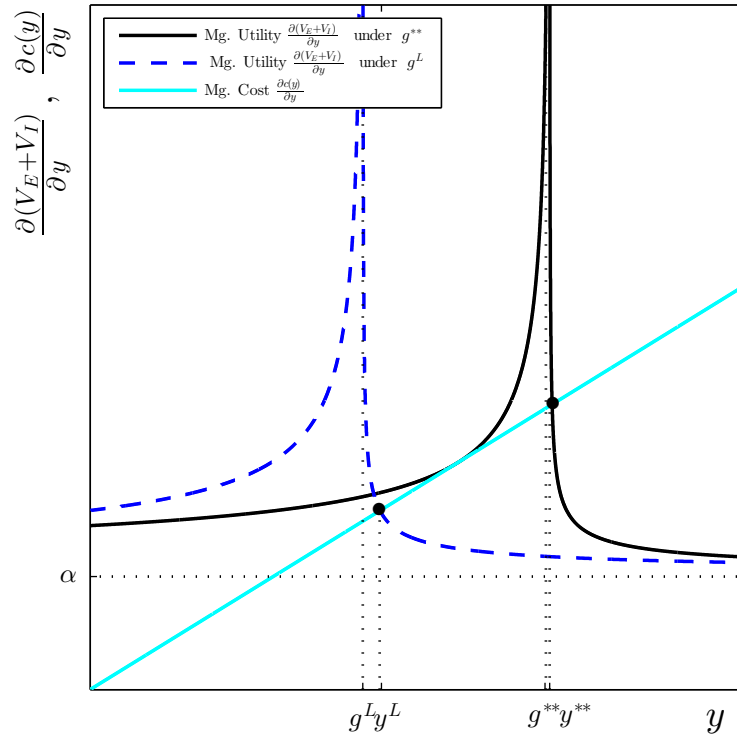


Figure 4: Output y^L when the goal is too low g^L

We first prove Lemma 1 that shows that there is only one goal that satisfies (10) at a point of tangency:

Lemma 1: *There exists a unique goal \hat{g} such that $\frac{\partial^2(V_E(y,\hat{g})+V_I(y,B))}{\partial y \partial g} = 1/\theta^2$ with an associated output level $\hat{y}(\hat{g})$ that satisfies (10).*

Proof. The (a) strict convexity of $\frac{\partial(V_E(y,B)+V_I(y,g))}{\partial y}$ with respect to y in the region $y < g$, indicated in (10) and illustrated in Figure 2, (b) the continuity of $V_I(y, g)$ with respect to g in (10) and (c) the property that $\frac{\partial^2(V_E(y,B)+V_I(y,g))}{\partial y \partial g} < 0$ in (10), ensure that there is a unique level curve in the domain of losses $\frac{\partial(V_E(y,B)+V_I(y,\hat{g}))}{\partial y}$ for a unique goal \hat{g} that yields a point of tangency with respect to y/θ^2 .

Strict convexity ensures that any goal g generates at most one point of tangency. Continuity ensures that there exists at least one goal g that has a point of tangency, since as g varies continuously, (9) does not change discontinuously. The uniqueness of the goal level \hat{g} that generates a point of tangency is guaranteed by $\frac{\partial^2(V_E(y,g)+V_I(y,B))}{\partial y \partial g} < 0$ in (10), which indicates that the marginal utility is strictly monotonically decreasing in g for any given value of y . In other words, as g increases, the marginal utility function moves in a strictly south-easterly direction. \square

We now show that the goal \hat{g} that generates the point of tangency is the optimal goal g^{**} .¹¹ We will require the following condition.

Assumption 1: For all, g' and y' , such that $g' < \hat{g}, y' < \hat{y}, g' < y'$ and $\hat{g} < \hat{y}$ hold, the following inequality also holds:

$$\frac{(\hat{y}/\theta)^2}{2} - \frac{(y'/\theta)^2}{2} < \alpha(\hat{y} - y') + \beta(\hat{g} - g') + \lambda[(g' - y')^{1/2} - (\hat{g} - \hat{y})^{1/2}] \quad (11)$$

This condition is a requirement that α , β and λ , for which higher values motivate greater output, be sufficiently large relative to the marginal cost of output. This condition ensures that monetary incentives in the worker's contract are sufficiently powerful so that the utility from achieving the goal is not the dominant motivation. It requires that the benefit from reducing one's goal from an optimal level g^{**} to a lower level g' not exceed the costs of doing so. The benefit is in terms of lower cost of effort $\frac{(\hat{y}/\theta)^2}{2} - \frac{(y'/\theta)^2}{2}$, and the costs are in terms of lost piece rate wages $\alpha(\hat{y} - y')$, foregone bonus $\beta(\hat{g} - g')$ and lost intrinsic utility $V_I(y', g') - V_I(\hat{y}, \hat{g})$, which equals $\lambda[(g' - y')^{1/2} - (\hat{g} - \hat{y})^{1/2}]$. We are now in a position to prove

¹¹We denote the optimal goal under reference dependence as g^{**} to distinguish it from g^* , the optimal goal in the absence of reference dependence.

Proposition 2, which characterizes the optimal goal under reference-dependent preferences.

Proposition 2: *Suppose that an agent has the preferences given in equation (8). Then, the optimal goal g^{**} satisfies:*

$$g^{**} = \alpha\theta^2 + 3\left(\frac{\lambda\theta^2}{4}\right)^{\frac{2}{3}} \quad (12)$$

and the optimal output y^{**} satisfies

$$y^{**} = \alpha\theta^2 + \frac{\theta^2}{2}(y^{**} - g^{**})^{-\frac{1}{2}} \quad (13)$$

Proof. We first show that a goal that specifies $g > \hat{g}$ are not optimal. Consider an individual who sets $\bar{g}'' > \hat{g}$. Recall that we maintain the assumption that the agent will stop exerting effort in the first instance for which the marginal cost of output exceeds the marginal utility of output. Let \hat{y} satisfy Lemma 2 and consider $\bar{y}'' > \hat{y}$. By the strict convexity and monotonicity of (10) in g , we have that $c(\bar{y}'') - c(\hat{y}) > V_E(\bar{y}'', (\bar{g}'')) + V_I(\bar{y}'', (\bar{g}'')) - V_E(\hat{y}, \hat{g}) + V_I(\hat{y}, \hat{g})$. This results in strictly lower earnings that under goal \hat{g} . This is illustrated in Figure 3 .

We now show that it is also not optimal to set $g < \hat{g}$. Consider an individual who sets $\underline{g}'' < \hat{g}$. She then reduces her piece rate compensation by $\alpha(\hat{y} - \underline{y}'')$. She further reduces her bonus, since she has now attained a less ambitious goal, by $\beta(\hat{g} - \underline{g}'')$. By shifting the goal to the left, she also losses $V_I(\underline{y}'', \underline{g}'') - V_I(\hat{y}, \hat{g}) = \lambda[(\underline{g}'' - \underline{y}'')^{1/2} - (\hat{g} - \hat{y})^{1/2}]$. On the hand the lower output entails lower cost, so that she gains $\frac{(\hat{y}/\theta)^2}{2} - \frac{(\underline{y}''/\theta)^2}{2}$. However, Assumption 1 guarantees that $U((\underline{y}''), \underline{g}'') < U((\hat{y}), \hat{g})$.

We have thus shown that the optimal goal $g^{**} = \hat{g}$. Calculating the optimal goal for our parameters, we have, differentiating (10) with respect to g ,

$$\frac{\theta\lambda}{4}(\hat{y} - g)^{-\frac{3}{2}} - \frac{1}{\theta} = 0 \quad (14)$$

Solving for g^{**} , we obtain the expression in (12). The production level y^{**} associated with the optimal goal level is found using (12) to solve (9), which yields the solution in (13). \square

Figure 2 shows the optimal goal level and its associated production level. From the figure it is evident that the optimal production level exceeds the optimal goal. This is a result of the assumption of reference dependence, since it does not occur if such dependence is not assumed. Thus, under reference dependence, performance will exceed the goal, though perhaps only by a modest amount.

From (12) it is clear that the optimal goal is strictly positive and bounded above. The model also generates the following comparative statics. First, an increase in the piece rate

increases the optimal goal ($\frac{\partial g^{**}}{\partial \alpha} > 0$) and the optimal output ($\frac{\partial y^{**}}{\partial \alpha} > 0$), as in the case of no reference dependence. Second, greater loss aversion leads to higher goals ($\frac{\partial g^{**}}{\partial \lambda} > 0$) and more output. The higher the ability level θ , the higher the goal ($\frac{\partial g^{**}}{\partial \theta} > 0$). Furthermore, it is straightforward to show that $\frac{dy^{**}}{dg^{**}} > 0$.¹² Therefore, for a given cost profile, increasing α , λ or θ increases the optimal goal and consequently the resulting production level. Thus, if agents are behaving optimally, output increases with higher goals.

3 The experiment

3.1 The general setting

Our dataset consists of 25 sessions. The sessions were conducted at the CentERLab in Tilburg University and all of the subjects were students at the university. We used Z-Tree (Fischbacher, 2007) to implement the experiment. Subjects were recruited via an online system. On average each session lasted approximately one hour. Between five and eighteen subjects took part in each session. No subject participated more than once in the experiment. The currency used in the experiment was Euros. A total of 232 subjects took part in the experiment.

In the experiment, subjects performed a time consuming real effort task under either a self-chosen goal incentive scheme or a piece rate incentive scheme, as described in Section 2. The real effort task was the one employed by Abeler et al., which consisted of counting the number of zeros in tables composed of 150 randomly ordered zeros and ones. The task was unfamiliar to all participants. It entailed a cost of effort in terms of attention and patience. In addition, the output of the task was of no use for the experimenter so that the impact of any social preferences with regard to the experimenter were minimized.

Each session was divided into nine five-minute periods. A subject's performance in each period counted toward her earnings. There were three treatments, called Self-chosen goal (GOAL hereafter), Low piece rate (LOPR) and High piece rate (HIPR). Thirty-nine subjects participated in LOPR, 93 in HIPR, and 100 in GOAL. The differences between the treatments are described in the following subsection. We use LOPR as the benchmark of comparison against which we evaluate the two other treatments.

3.2 The Three Treatments

In the GOAL treatment, subjects had to choose a target level of corrected solved tasks per round g the beginning of each period. Achieving that goal would yield a monetary bonus

¹²Implicit differentiation of (13) yields that $\frac{dy^{**}}{dg^{**}} = \frac{\frac{\theta}{4}(y^{**}-g^{**})^{-\frac{3}{2}}}{\frac{\theta}{4}(y^{**}-g^{**})^{-\frac{3}{2}} + \frac{1}{\theta}}$, thus it is clear that $0 < \frac{dy^{**}}{dg^{**}} < 1$, given that $(y^{**} - g^{**}) > 0$.

B , increasing in the goal level at a rate $\beta = 20$ (20 Eurocents). That is, setting the goal one successfully completed unit of the task higher yielded 20 cents more, conditional of the goal being attained. In addition to the bonus if reached, subjects receive a piece rate of $\alpha = 20$ (20 Eurocents) per correctly solved table or unit of output y . The payoff function of a participant assigned to this treatment in each round was:

$$w^{GOAL} = 20 \cdot y - 20 \cdot \lfloor \frac{1}{3}i \rfloor + B(g, y) \quad (15)$$

with

$$B(g, y) = \begin{cases} 20 \cdot g & \text{if } y \leq g \\ 0 & \text{if } y > g \end{cases} \quad (16)$$

For every third incorrectly solved table, i , a participant had 20 cents subtracted from her earnings. This punishment was introduced to reduce guessing on the part of participants, as well as to capture a situation in which there is a cost to the worker of making errors, such as for example producing defective units of output.

In the LOPR treatment, subjects were paid a constant piece rate of $\alpha = 20$ (20 Eurocents) for each unit of output. A penalty of 20 cents for every third incorrectly solved table was also in effect. The per-period payoff function of a participant assigned to this treatment was the following:

$$w^{LOPR} = 20 \cdot y - 20 \cdot \lfloor \frac{1}{3}i \rfloor \quad (17)$$

The HIPR treatment was identical to LOPR except that subjects received a piece rate of $\alpha = 50$ (50 cents) for each unit of output. The per-period payoff function in this treatment was:

$$w^{HIPR} = 50 \cdot y - 50 \cdot \lfloor \frac{1}{3}i \rfloor \quad (18)$$

Similarly to under the LOPR treatment, for every third incorrectly solved table, a participant had 50 cents subtracted from her earnings.

Figure 5 illustrates the incentives in effect in each treatment. The horizontal axis indicates the number of units of output, and the compensation a participant receives is given on the vertical axis. The figure shows that for any level of production the earnings associated with HIPR always dominate the payments associated to LOPR and GOAL. In turn, the earnings in GOAL weakly dominate those of LOPR for any output profile.

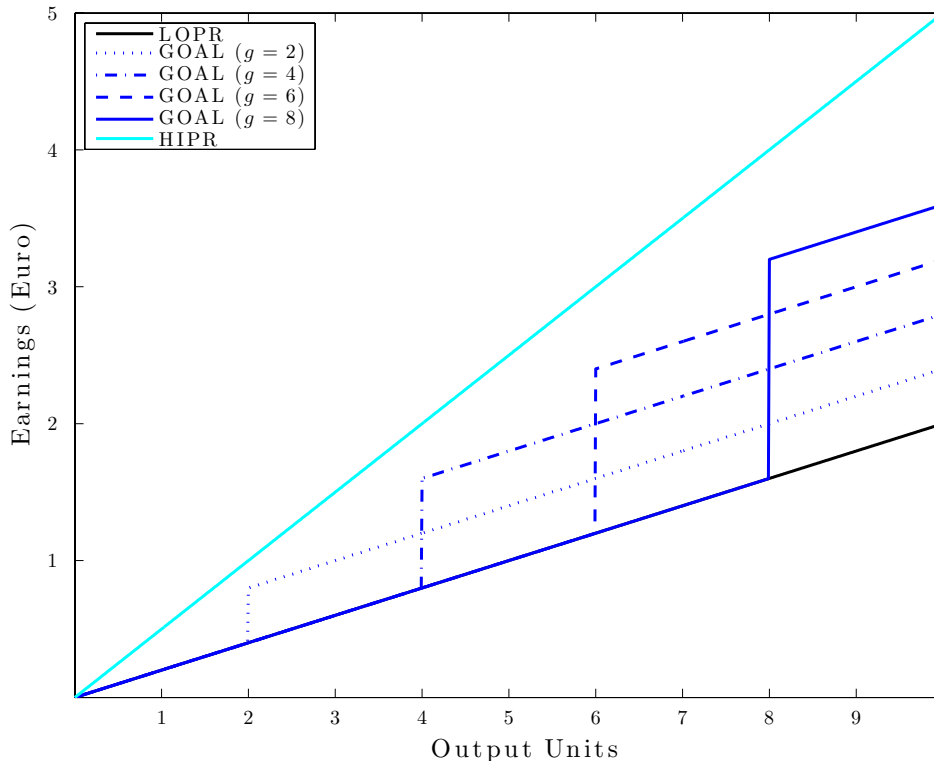


Figure 5: Incentives by Treatment

4 Results

We first consider differences in output between the GOAL, LOPR and HIPR contracts. In order to do so, we estimate the following Poisson regression of output, measured as the number of correctly solved tables, on the treatment dummies, some demographic characteristics (domestic or foreign student and program of study) and period dummies.

$$Correct_{rsj} = \beta_0 + \beta_1 \cdot HIPR_j + \beta_2 \cdot GOAL_j + \Gamma' Controls_{rsj} + \epsilon_{rsj} \quad (19)$$

$Correct_{rsj}$ denotes the number of correctly solved tables by subject j in round r of session s . As Table 1 shows, when we consider the whole sample, the different contracts do not seem to affect output significantly. However, when we split the sample by men and women, the gender differences become clear. According to Table 1 and Figure 6, men increase their output by a roughly similar proportion, compared to the reference condition LOPR, when they are offered a higher piece rate, under HIPR, or when they can set their own goal, under GOAL.

Under HIPR, men's output is on average 23% higher than under LOPR, ($p=0.052$). The GOAL contract, in turn, yields higher output on the part of male participants by 29 % on average compared to LOPR ($p<0.01$). There is no significant difference in performance

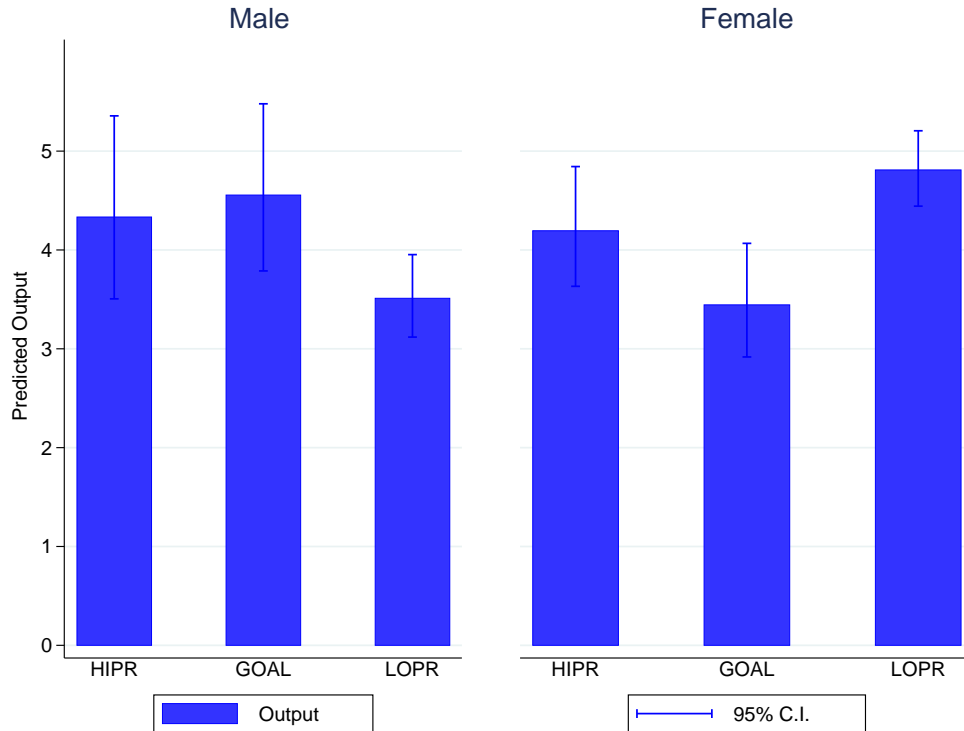


Figure 6: Predicted Output by Condition and Gender

between the HIPR and GOAL for male participants ($\chi^2(1)=0.19$, $p=0.664$). Hence, if the goal of the employer is to increase male workers' performance, she would be better off imposing either the HIPR or the GOAL contract than under LOPR. Each of the high-powered contracts would result in more output relative to the low-powered piece rate contract, with the GOAL contract inducing even higher performance at lower cost than HIPR.

The results are very different for women. Both relatively high-powered incentive contracts, the HIPR and the GOAL, generate lower average output than LOPR. Under HIPR, women's output is on average 13 % lower than under LOPR, *ceteris paribus*, ($p=0.062$). GOAL also decreases output by 33 % on average compared to LOPR ($p<0.001$). Women's performance in the HIPR or GOAL is about the same ($\chi^2(1)=4.17$, $p=0.041$). Thus from the point of view of the employer whose target is to increase a female worker's output, both increasing the piece rate or imposing a goal contract would be counterproductive.

We summarize this discussion of the gender differences on the effect of incentive change on effort as our first result.

Result 1 (Gender differences in output): *The GOAL and the HIPR contracts yield more production from men, but lower output from women, than under LOPR.*

Table 1: Determinants of Output

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Sample		Male Sample		Female Sample	
	Correct	Correct	Correct	Correct	Correct	Correct
HIPR	-0.003 (0.064)	-0.009 (0.063)	0.210 (0.108)	0.227* (0.109)	-0.137 (0.073)	-0.164* (0.071)
GOAL	-0.100 (0.066)	-0.130* (0.066)	0.260** (0.094)	0.222* (0.093)	-0.334*** (0.085)	-0.355*** (0.085)
constant	1.451*** (0.037)	1.137*** (0.051)	1.256*** (0.060)	0.908*** (0.079)	1.571*** (0.040)	1.289*** (0.062)
Demographics	NO	YES	NO	YES	NO	YES
Round dummies	NO	YES	NO	YES	NO	YES
Session FE	YES	YES	YES	YES	YES	YES
N	1926	1926	927	927	999	999
Pseudo Log-likelihood	-3967.658	-3889.242	-1910.538	-1870.165	-2011.267	-1974.703

Note: This table presents estimations of the model $Correct_{rjs} = \beta_0 + \beta_1 \cdot HIPR_j + \beta_2 \cdot GOAL_j + \Gamma' Controls_{rjs} + \epsilon_{rjs}$ with $\epsilon_{rjs} \sim poisson(\mu)$ and $Controls_{rjs}$ contains session dummies, round dummies and demographic variables such as Dutch nationality, following studies related to economics and degree level. All estimations used Poisson count data regressions and robust standard errors, which are presented in parentheses. Regression (1) are the estimates excluding session dummies and demographic variables. Regression (2) are the estimates of the complete model. Regression (3) are the estimates of the model excluding demographics and the period dummies when the sample is composed of male subjects. Regression (4) are the estimates of the complete model when the sample is uniquely composed of male subjects. Regression (5) are the estimates of the model excluding demographics and the period dummies when the sample is composed of female subjects. Regression (6) presents the estimates of the complete model when the sample is composed by female subjects. *** denotes significance at the 0.1 percent level, ** denotes significance at the 1 percent level and * denotes significance at the 5 percent level.

We compare the genders to each other with regard to their output by estimating the following model using a Poisson count regression.

$$Correct_{rjs} = \beta_0 + \beta_1 \cdot HIPR_j + \beta_2 \cdot GOAL_j + \beta_3 \cdot HIPR_j \cdot gender + \beta_4 \cdot GOAL_j \cdot gender + \beta_5 gender + \Gamma' Controls + \epsilon_{rjs} \quad (20)$$

As Table 2 and Figure 6 show, women produce more output (approximately 36% more correct tables) than men under the low piece rate contract ($p < 0.001$). In contrast, men produce more output than women when the piece rate is higher HIPR ($\chi^2(1) = 7.31$, $p < 0.01$) and when they face the GOAL contract ($\chi^2(1) = 17.14$, $p < 0.001$).

In Section 2, we showed that if agents have reference-dependent preferences, they will outperform their goals, though perhaps by only a small amount. This is precisely what we observe in the GOAL treatment. While the average output was 4.09 tables per period, the average goal level was 3.75. This gap of 0.35 tables is statistically significant ($t(809) = -5.35$,

Table 2: Determinants of Output with Gender Interaction Effects Included

	(1)	(2)
	Correct	Correct
HIPR	0.210 (0.108)	0.227* (0.109)
GOAL	0.260** (0.094)	0.222* (0.093)
HIPR · gender	-0.347** (0.131)	-0.391** (0.130)
GOAL · gender	-0.594*** (0.127)	-0.577*** (0.126)
gender	0.315*** (0.073)	0.381*** (0.100)
Constant	1.256*** (0.060)	0.908*** (0.079)
Demographics	NO	YES
Round dummies	NO	YES
Session FE	YES	YES
N	1926	1926
Pseudo Log-Likelihood	-3921.805	-3844.867

Note: This table presents estimations of the model $Correct_{rjs} = \beta_0 + \beta_1 \cdot HIPR_j + \beta_2 \cdot GOAL_j + \beta_3 \cdot HIPR_j \cdot gender + \beta_4 \cdot GOAL_j \cdot gender + \beta_5 \cdot gender + \Gamma' Controls_{rjs} + \epsilon_{rjs}$ with $\epsilon_{rjs} \sim poisson(\mu)$ and $Controls_{rjs}$ contains session dummies, round dummies and demographic variables such as Dutch nationality, following studies related to economics and degree level. All regressions used poisson count data regressions and robust standard errors, which are presented in parentheses. Regression (1) presents the estimates of the model excluding demographic and round dummies. Regression (2) presents the estimates of the complete model. $gender = 1$ if the subject is a female. *** denotes significance at the 0.1 percent level, ** denotes significance at the 1 percent level and * denotes significance at the 5 percent level.

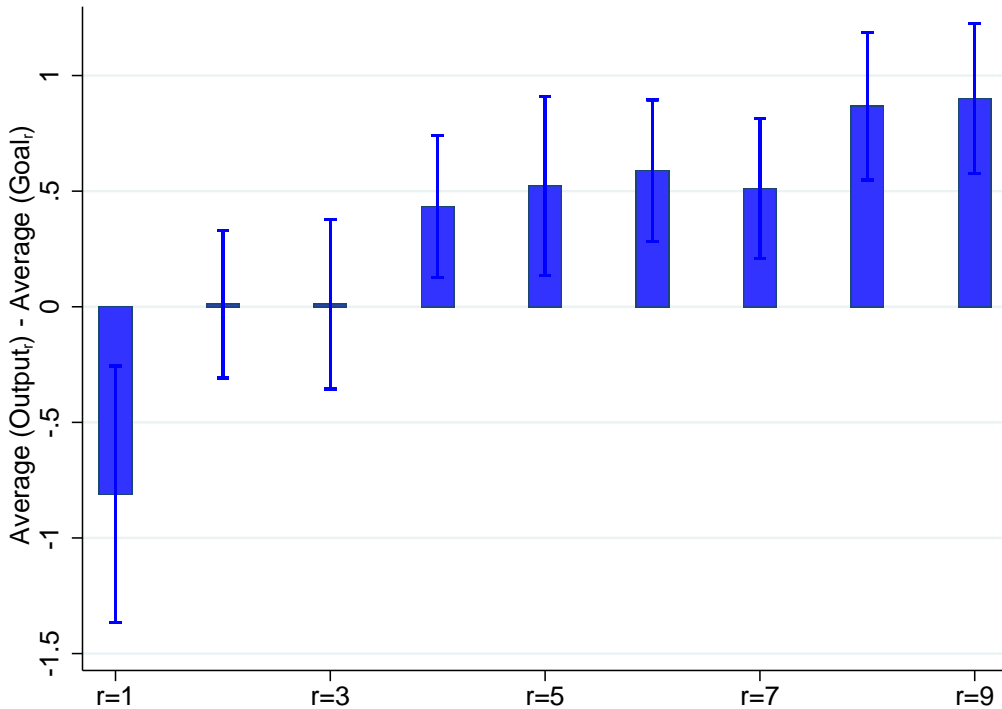


Figure 7: Average Output minus goal in each round

$p < 0.001$) and relatively small (performance being only 9% higher than the goal). The pattern was present for both women and men. The gap between output and goals is always positive, although it only becomes statistically significant from the third round onward. The increase in the difference between output and goals with experience suggests that the difference is not due to risk aversion, since it can be presumed that beliefs about one's own performance become more precise with practice. This would mean that an agent with a given level of risk aversion, in the absence of other effects, would exhibit a decreasing excess of output over his goal with experience. The empirical pattern shown in Figure 7 suggests that the piling up effect is a feature of long-run behavior. This discussion is summarized as our second result.

Result 2 (Piling-up): *Subjects tend to outperform their goals. The difference between goals and output does not decrease with repetition.*

As predicted by the model, we also observe a positive correlation between the goal level and effort ($\rho_{correct_j, goal_j} = 0.4437, p < 0.01$). Notably, this positive relationship appears when the goal lies in an intermediate range, where the goals are presumably challenging and

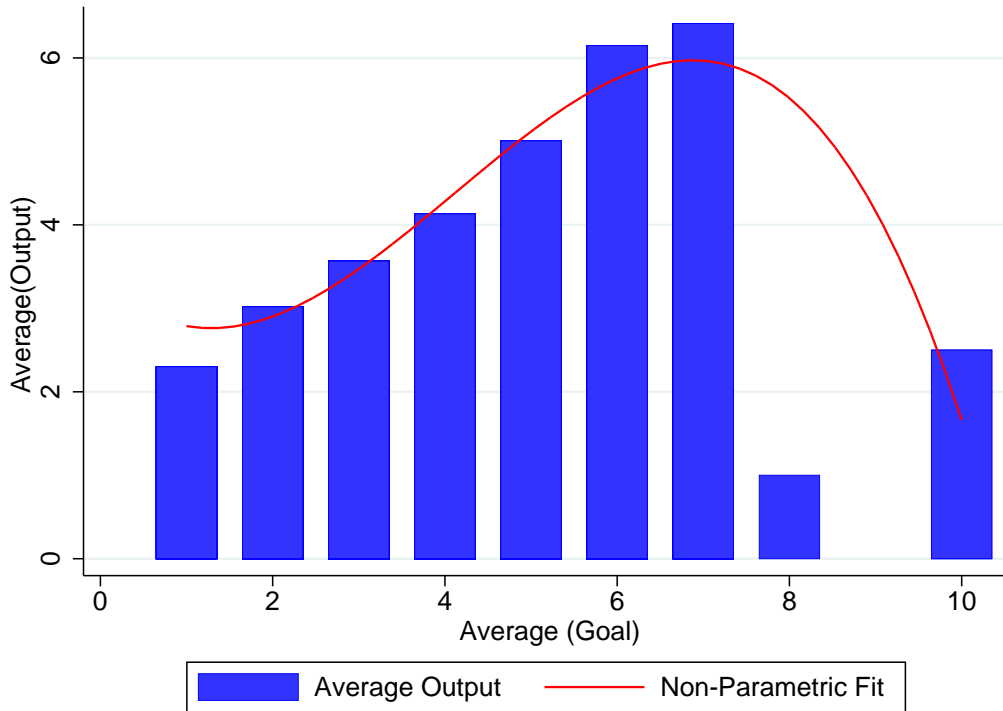


Figure 8: The relation between Goal level and Output

attainable. More precisely, when goals are very high ($g > 7$) the correlation between effort and goal level is not significant and possibly even negative. Figure 8 displays a graphical representation of this phenomenon.

5 Understanding the Gender Differences

The model predicts that the GOAL contract increases workers' output. While we observe this effect for men, we find exactly the opposite for women. Women produce significantly less under the self-chosen goal contract, even compared to a contract that offers less money in expected terms. In this section, we investigate potential reasons for the gender difference, and we report evidence suggesting that the main reason seems to be that women systematically underestimate their potential, because they do not take into account the fact that their performance improves with practice.

Women set significantly lower goals than men (See Table 3).¹³ This effect is present in all nine rounds of the experiment. Moreover, we observe that the extent to which performance exceeds goals is greater for women than for men. Women outperform their goals by a 0.29 tasks more than what men do ($p < 0.05$).

¹³The average goal set by women is roughly half of a table lower than that set by men ($p < 0.001$).

Table 3: Average Behavior by Gender in GOAL

Gender	Goal	Margin	Prob. Achiev.	Goal (1 st round)
Male	4.049 (1.395)	.2380 (1.862)	.739 (.439)	4.204 (1.925)
Female	3.557 (1.246)	.527 (1.694)	.773 (.419)	3.352 (1.230)
Total	3.798 (1.343)	.385 (1.783)	.756 (.4293)	3.77 (1.656)

Note: This table presents the averages and standard deviations of goal levels, the difference between the goal level and the produced output, the probability of achievement of the self-chosen goal and the goal set in the first round in the GOAL treatment. Standard deviations are in parentheses.

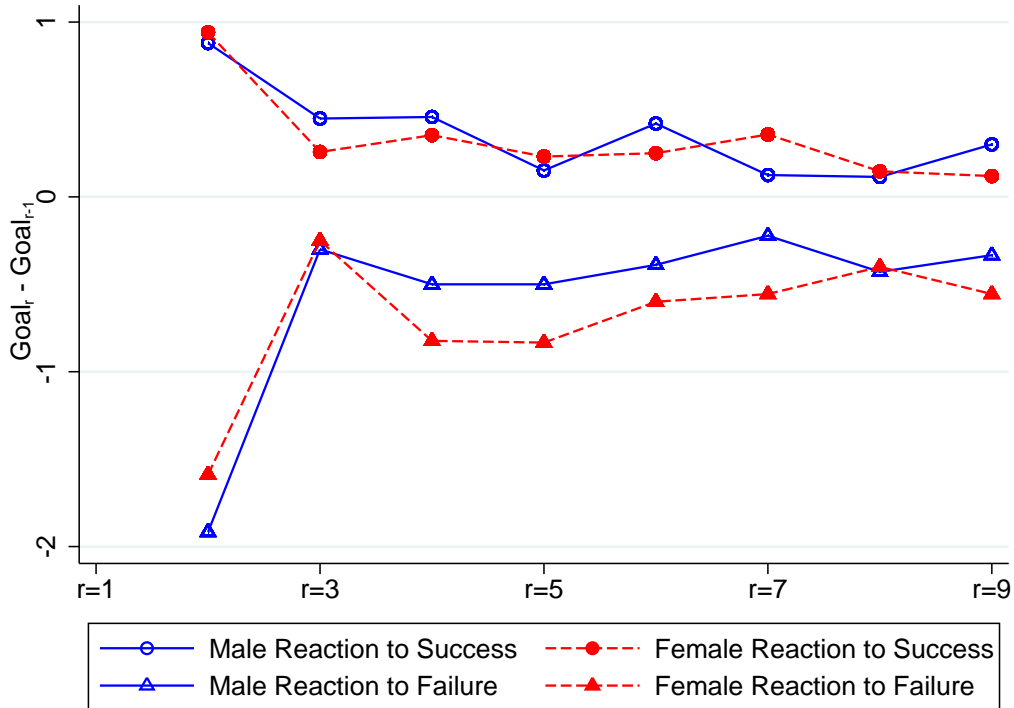


Figure 9: Goal setting as reaction to success or failure by gender

Is this evidence suggesting that women do not set challenging enough goals? Although we do not know the actual maximum potential of individuals to perform in the task, the potential of women is not likely to be lower than that of men, since we observe that women perform better than men under the baseline LOPR contract. If anything, we would expect women to set higher goals than men. However, the fact that we observe the contrary, suggests that women set goals that are not challenging enough. Moreover, because the monetary bonus is strictly increasing in the goal chosen, setting goals below one’s own potential is costly in terms of foregone earnings.

The question that follows is why women set goals that are suboptimally low. Two candidate explanations are that the pattern is due to women having lower self-efficacy beliefs, or greater risk aversion, than men (Eckel and Grossman, 2008). In fact, empirical evidence shows that women exhibit systematically lower general self-efficacy than men (Scholz et al., 2002) and are more risk averse (Croson and Gneezy, 2009; Eckel and Grossman, 2008).¹⁴ However, while low self-efficacy beliefs or high risk aversion can explain the *initial* low level of goal for females, they cannot explain the fact that women set low goals systematically over the course of the nine rounds, even after they have the chance to update their beliefs with full information about their own performance.

We consider gender differences in the way individuals update their goals in response to their performance and goal level in the previous round, by estimating the following regression ¹⁵ ¹⁶:

$$\begin{aligned} \log(g)_{jr} = & \theta_0 + \theta_1 \cdot \log(g)_{j,r-1} + \theta_2 \cdot \max(0, g_{j,r-1} - y_{j,r-1}) + \\ & \theta_3 \cdot \max(0, y_{j,r-1} - g_{j,r-1}) + \Gamma' Controls_j + \nu_{jr} \end{aligned} \quad (21)$$

¹⁴Self-efficacy is the extent of one’s belief in one’s own ability to complete a task and reach a goal (Bandura (1977); Bandura and Cervone (1986)). Self-efficacy can be conceptualized as either general or task-specific. General self-efficacy represents an optimistic general sense of personal competence. The task-specific type represents one’s perceived beliefs about competence in a particular task. For example, females tend to exhibit higher language self-efficacy than males, but lower mathematical self-efficacy (Huang, 2013). In our experimental setting, we can interpret the goal level as a measure of subjects’ self-efficacy. This is possible because the elicitation of the goal is incentivized monetarily in such a way that a rational subject would announce what she believes is close to the highest production level that she can achieve. As we have seen, the higher the belief in her ability to perform the task (i.e. the higher the self-efficacy), the higher the goal that the person will announce.

¹⁵Note that this specification allows for the possibility that the goal is updated asymmetrically in response to success or failure.

¹⁶Including only a one-period lag in the goal setting process has a better fit than models including lags of two or more periods.

Table 4 presents the Arellano-Bond estimation of this model. The estimates show that men and women adopt the same adjustment behaviour in reaction to success or failure to reach their goals. They both decrease their goal (by 8%) after failing to achieve their goal ($\chi^2(1)=0.29$, $p=0.587$) and increase their goal (by 7 %) ($\chi^2(1)= 0.80$, $p=0.3719$) when the previous round goal is achieved. However, in addition, men exhibit an increasing trend in the goals they set. On average, the per round rate of goal growth for the male subsample (before the effect of prior success and failure to reach the goal in the last period is taken into account) is 41.5%. However, there is no evidence of a trend in goal setting for women. These empirical observations can be summarized in the following result.

Result 3 (Gender Differences in Goal Setting): *Both men and women tend to increase their goals after successfully attaining them and decrease their goals after failing to do so. However, controlling for these effects, men tend to increase their goals over time, while women do not.*

This systematic failure to internalize the trend of improvement in performance comes at a cost for women. They earn in average 0.87 euros less than men in our experiment. They could have earned on average 7.5% more money had they set and updated their goals in a similar manner to men.¹⁷

Women’s underperformance in a self-chosen goal contract is essentially explained by their failure to update their goals upwards, even when they learn that they can perform better than what they initially thought. They remain low even when it is monetarily costly to do so, and when all the information needed to change them is available.

6 Conclusion

This paper studies the properties, both theoretical and behavioral, of a compensation scheme in which agents set their own production goals. We first proposed a model in which goals assume the role of endogenous reference points. We then test the model’s predictions with a laboratory experiment. We observe (a) evidence for pilling-up, that is, subjects systematically outperform their goals, but only by a small amount, and (b) that goals do motivate greater effort, at least on the part of male participants.

However, the effectiveness of the self-chosen goal incentive scheme depends crucially on the gender of the worker. While the self-chosen goal contract makes men increase

¹⁷This result is not minor, considering, for example, that the gender gap in earnings in the United States for teenagers between 19-24 years old (the age of the subjects in our sample) is about 9% (US Department of Labor, 2010).

Table 4: Determinants of Goal Setting

	(1)	(2)	(3)
	Total Sample	Male	Female
	Goal (Present round)	Goal (Present round)	Goal(Present round)
Goal (lag round) ($\log(g_{s,r-1})$)	0.343*	0.415**	0.189
	(0.138)	(0.154)	(0.167)
Success (lag round) ($\max(0, y_{j,r-1} - g_{j,r-1})$)	0.068***	0.073***	0.060***
	(0.011)	(0.014)	(0.014)
Failure (lag round) ($\max(0, g_{j,r-1} - y_{j,r-1})$)	-0.078***	-0.080***	-0.076*
	(0.019)	(0.020)	(0.032)
gender	-0.031		
	(0.028)		
Constant	0.821***	0.795***	0.953***
	(0.175)	(0.193)	(0.189)
Demographics	YES	YES	YES
N	630	308	322

Note: This table presents the Arellano Bond Panel Data estimation of the model $\log(g)_{j,r} = \theta_0 + \theta_1 \cdot \log(g)_{j,t-1} + \theta_2 \cdot \max(0, g_{j,t-1} - y_{j,t-1}) + \theta_3 \cdot \max(0, y_{j,t-1} - g_{j,t-1}) + \beta_1 \cdot \text{gender} + \Gamma' \text{Controls}_j + \nu_{j,r}$, with *Controls* containing demographic variables such as Dutch nationality and degree level. All regressions used robust standard errors which are presented in parentheses. Regression (1) presents the estimation of this model using the whole sample. Regression (2) presents the estimation of this model using male subjects. Regression (3) presents the estimation of the model using female subjects. *** denotes significance at the 0.1 percent level, ** denotes significance at the 1 percent level and * denotes significance at the 5 percent level.

their production compared to a low piece rate contract, it is counterproductive for women. Women produce significantly less under the self-chosen goal contract than when they are offered a simple piece rate contract that offers equal or lower earnings at any level of production. This is the case despite the fact that women perform the task better than men in the baseline LOPR condition, indicating that there is no gender bias in favor of men in the task. Self-chosen goals, in contrast, are very effective for men, as they produce the same output and effort than when they are offered high piece rate payments. The pattern of performance of women that we observe, lower output at higher levels of compensation in HIPR and GOAL relative to LOPR, is consistent with satisficing behavior with regard to income or a backward-bending labor supply curve for labor reflecting strong income effects.

Women's underperformance relative to men under a self-chosen goal contract is, at least in part, explained by their failure to systematically update their goals upward to reflect improving performance. Previous work has shown that women tend to shy away from competitive challenges more than men(?). Here, we find that they also shy away from challenging themselves. While the behavior of men is very consistent with our model's predictions, the results indicate that gender should be taken into account when an employer is choosing how to incentivize employees.

A natural question to ask is how a self-chosen goal compares to a goal contract exogenously chosen by the employer. The chief advantage of the self-chosen goal contract is that it can exploit the private information that the employee has about his/her own ability, and potentially align the incentives of employer and employee. On the other hand, an exogenous goal set by an employer might be able to benefit from an additional motivating force for the worker, the desire to reciprocate a generous wage offer chosen by an employer (Fehr et al., 1998). This is an issue to be explored in future research.

References

- Abeler, J., Falk, A., Goette, L., and Huffman, D. Reference points and effort provision. *The American Economic Review*, (2):470–492.
- Baber, W. R., Daniel, P. L., and Roberts, A. a. (2002). Compensation to managers of charitable organizations: An empirical study of the role of accounting measures of program activities. *Accounting Review*, 77(3):679–693.
- Baker, G. P., Jensen, M. C., and Murphy, K. J. (1988). Compensation and Incentives: Practice vs. Theory. *The Journal of Finance*, 43(3):593–616.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2):191–215.

- Bandura, A. and Cervone, D. (1986). Differential engagement of self-reactive influences in cognitive motivation. *Organizational Behavior and Human Decision Processes*, 38:92–113.
- Banker, R. D., Lee, S.-Y., Potter, G., and Srinivasan, D. (2000). An empirical analysis of continuing improvements following the implementation of a performance-based compensation plan. *Journal of Accounting and Economics*, 30:315–350.
- Benabou, R. and Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies*, 70(1):489–520.
- Croson, R. and Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, 47(2):448–474.
- Eckel, C. C. and Grossman, P. J. (2008). Chapter 57 Differences in the Economic Decisions of Men and Women: Experimental Evidence. *Handbook of Experimental Economics Results*, 1(1998):509–519.
- Enis, C. R. (1993). Earnings-Based Compensation Plans, Performance, and Capital Expenditure Policy Carrier in the Motor Carrier Industry. *The Accounting Review*, 68(4):928–941.
- Fehr, E., Kirchsteiger, G., and Riedl, A. (1998). Gift exchange and reciprocity in competitive experimental markets. *European Economic Review*, 42(96):1–34.
- Fehr, E., Klein, A., and Schmidt, K. M. (2007). Fairness and Contract Design. *Econometrica*, 75(1):121–154.
- Fischbacher, U. (2007). Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178.
- Goerg, S. J. and Kube, S. Goal Setting and Monetary Incentives : When Large Stakes Are Not Enough. Preprints of the Max Planck Institute for Research on Collective Goods 2012:19, Max Planck Institute.
- Gómez-Miñambres, J. (2012). Motivation through goal setting. *Journal of Economic Psychology*, 33:1223–1239.
- Gómez-Miñambres, J., Corgnet, B., and Hernán-gonzalez, R. (2012). Goal Setting and Monetary Incentives : When Large Stakes Are Not Enough. Economic Science Institute Working Papers:12-24, Chapman University.
- Groves, T., Hong, Y., MacMillan, J., and Naughton, B. (1994). Autonomy and Incentives in Chinese State Enterprises. *The Quarterly Journal of Economics*, 109(1):183–209.

- Heath, C., Larrick, R. P., and Wu, G. (1999). Goals as reference points. *Cognitive psychology*, 38:79–109.
- Hsiaw, A. (2013). Goal-setting and self-control. *Journal of Economic Theory*, 148(2):601–626.
- Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, 28:1–35.
- Hurwicz, L. (1973). The Design Resource of Mechanisms Allocation for. *American Economic Review*, 63(2):1–30.
- Jones, D. C. and Kato, T. (1995). The Productivity Effects of Employee Stock-Ownership Plans and Bonuses: Evidence from Japanese Panel Data. *American Economic Review*, 85(3):391–414.
- Joseph, K. and Kalwani, M. U. (1998). The Role of Bonus Pay in Salesforce Compensation Plans. *Industrial Marketing Management*, 27:147–159.
- Kahn, L. M. and Sherer, P. D. (1990). Contingent Pay and Managerial Performance. *Industrial and Labor Relations Review*, 43(3).
- Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–291.
- Kőszegi, B. and Rabin, M. (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.
- Koch, A. K. and Nafziger, J. Motivational goal bracketing. IZA Discussion Paper 4471, IZA institute.
- Koch, A. K. and Nafziger, J. (2011). Self-regulation through Goal Setting. *Scandinavian Journal of Economics*, 113(1):212–227.
- Laffont, J. J. (1994). The New Economics of Regulation Ten Years After. *Econometrica*, 62(3):507–537.
- Latham, G. P. and Locke, E. a. (1991). Self-Regulation through Goal Setting. *Organizational Behavior and Human Decision Processes*, 50:212–247.
- Lazear, E. P. (1986). Salaries and Piece Rates. *The Journal of Business*, 59(3):405.
- Lazear, E. P. (2000). Performance Pay and Productivity. *The American Economic Review*, 90(5):1346–1361.

- Lemieux, T., Bentley, M. W., and Parent, D. (2009). Performance Pay and Wage Inequality. *Quarterly journal of economics*, 124(1):437–452.
- Locke, E. A. (1996). Motivation through conscious goal setting. *Applied and Preventive Psychology*, 5:117–124.
- Locke, E. A. and Latham, G. P. (1990). Work Motivation and Satisfaction: Light at the End of the Tunnel. *Psychological Science*, 1:240–246.
- Locke, E. A. and Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *The American psychologist*, 57(9):705–717.
- Mirrlees, J. (1976). Optimal tax theory. *Journal of Public Economics*, 6:327–358.
- Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *Quarterly Journal of Economics*, 122(August):1067–1101.
- Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature*, 37(1):7–63.
- Rajagopalan, N. (1996). Strategic Orientations, Incentive Plan Adoptions and Firm Performance: Evidence from Electric Utility Firms. *CEO publication*, G96-13(307).
- Scholz, U., Doña, B. G., Sud, S., and Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment*, 18:242–251.
- WorldatWork (2014). Bonus Programs and Practices. Technical Report June.
- Wu, G., Heath, C., and Larrick, R. (2008). A prospect theory model of goal behavior. mimeo.

7 Appendix

Experiment Instructions 20 cent treatment

This is an experiment in the economics of decision-making. The instructions are simple and if you follow them carefully and make good decisions, you might earn a considerable amount of money, which will be paid to you via bank transfer. The amount of payment you receive depends on your decisions. The currency is Euros and the payment to you will be in Euros.

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.

What do you have to do?

The experiment consists of a sequence of 9 rounds of 5 minutes each. In each round your task is to count zeros in a series of tables. The figure shows the work screen that you will use later.

The screenshot shows a computer interface for an experiment. At the top left, it says "Period" and "3 out of 3". On the left side, there is a table of 15 rows and 15 columns of binary digits (0s and 1s). On the right side, there is a response box with the text "Time spent 76" and "How many zeros are in the table?". Below the text is a blue input field. At the bottom right, there is a red "OK" button.

1	1	1	0	0	0	0	0	1	0					
1	1	1	0	1	1	1	0	1	0					
0	0	1	1	0	0	0	1	1	1					
0	0	1	1	1	0	1	0	0	1					
0	0	1	0	0	0	0	0	0	0					
0	0	1	0	0	0	1	1	0	1					
1	0	0	0	1	0	1	0	0						
0	0	0	1	1	0	0	1	0	0					
0	0	0	0	1	0	1	0	1	0					
0	0	0	0	0	0	0	0	0	0					
0	1	0	0	1	0	0	1	0	1					
1	0	1	1	0	0	1	0	0	1					
0	1	0	0	1	1	1	1	0	0					
0	1	0	0	0	0	0	0	1	1					
0	0	1	0	0	0	0	0	1	1					

When faced with this screen you have to insert in the box, located on the right side of the screen, the exact number of zeros that appear in the table set in the left side of the screen. After you have entered the number, click the OK button. If you entered the correct number of zeros you will receive 20 cents, but if your input was wrong you will receive nothing and you will be accounted with one "fail". Furthermore, with three "fails" 20 cents will be deducted from your total earnings. Immediately after you are notified whether you solved a table correctly or not, a new table table will be displayed and you are again expected to count the number of zeros. This procedure is repeated until the time is up. In total, you have 5 minutes per round. In other words, you have 5 minutes to solve as many tables as you can in each round and you will have 9 rounds. The elapsed time is displayed above the input box as you can see in the image displayed above.

Counting Tips: Of course you can count the zeros in anyway 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.

Total earnings per round

To clarify, your total compensation in each round will be :The number of tables you solve during the round n_t at the rate of 20 cents each. This is summarized as follows:

$$E_t = n_t * 20$$

Total earnings The total earnings of the experiment are the sum of compensations for each round. This is summarized as follows

$$E = E_1 + \dots + E_9$$

For further comprehension we provide some examples.

Example 1

Assume that in some round you solved two tables correctly. Your earnings are therefore:

$$Earnings = 2 * 20 = 40$$

Thus, you earn a total of 40 cents or 0.4 Euros.

Example 2

Assume that in some round you solved three tables correctly. Your earnings are therefore:

$$Earnings = 3 * 20 = 60$$

Thus, you earn a total of 60 cents or 0.6 Euros.

Example 3.

Assume that in some round you solved three tables correctly but you had 4 fails. Your earnings are therefore:

$$Earnings = 3 * 20 - 20 = 40$$

Thus, you earn a total of 40 cents or 0.4 Euros.

Questions

To be sure that you understood correctly this instructions you need to answer the following questions.

1. You solved four tables correctly. Your earnings are therefore:
2. You solved two tables correctly and had one "fail". Your earnings are therefore:
3. You solved four tables correctly and had seven "fails". Your earnings are therefore:

Final Remarks

After you are done please wait in your chair a questionnaire will be provided and the experimenter will go to your seat and proceed to get your bank account number to pay your total gains in the experiment. Please keep in mind throughout the experiment that you cannot use the Enter key to proceed to the next screen. Instead you are advised to use the mouse to click the respective buttons.

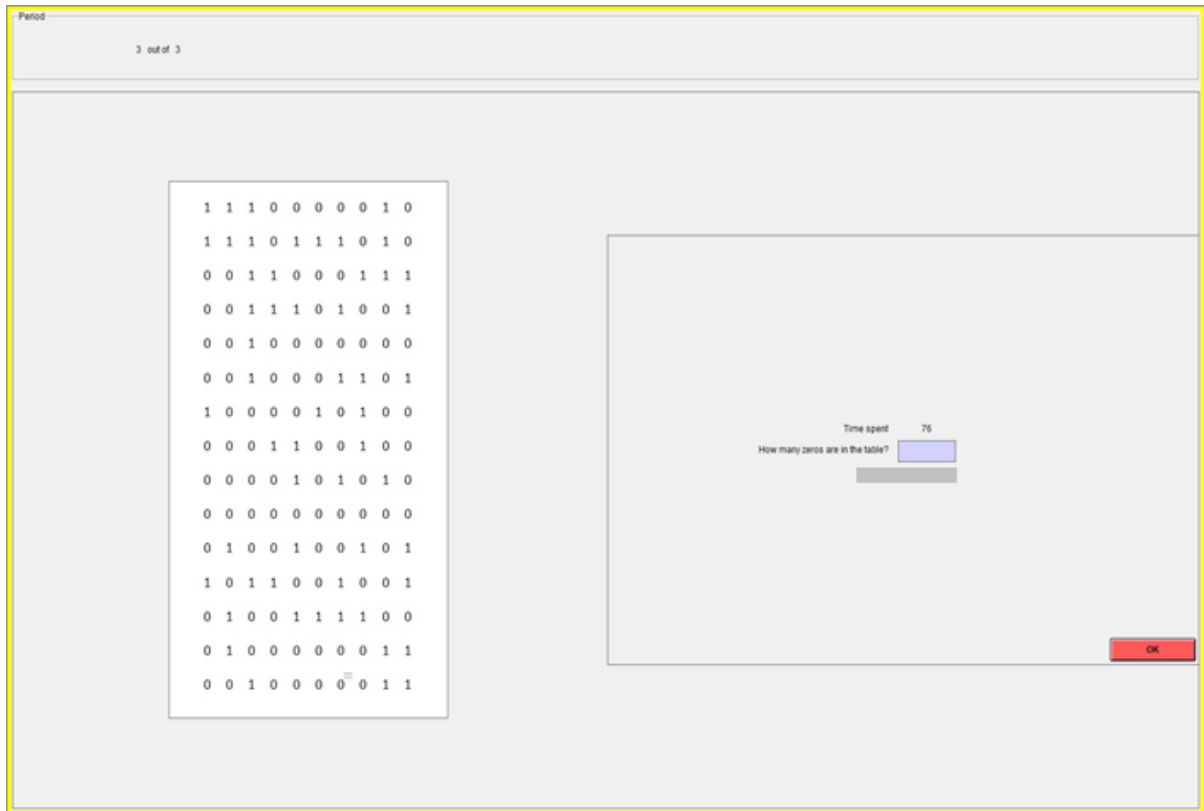
Experiment Instruction 50 cent treatment

This is an experiment in the economics of decision-making. The instructions are simple and if you follow them carefully and make good decisions, you might earn a considerable amount of money, which will be paid to you via bank transfer. The amount of payment you receive depends on your decisions. The currency used in the market is Euros. The payment to you will be in Euros.

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.

What do you have to do?

The experiment consists of a sequence of 9 rounds of 5 minutes each. In each round your task is to count zeros in a series of tables. The figure shows the work screen that you will use later.



When faced with this screen you have to insert in the box, located on the right side of the screen, the exact number of zeros that appear in the table set in the left side of the screen. After you have entered the number, click the OK button. If you entered the correct number of zeros you will receive 20 cents, but if your input was wrong you will receive nothing and you will be accounted with one "fail". Furthermore, with three "fails" 20 cents will be deducted from your total earnings. Immediately after you are notified whether you solved a table correctly or not, a new table will be displayed and you are again expected to count the number of zeros. This procedure is repeated until the time is up. In total, you have 5 minutes per round. In other words, you have 5 minutes to solve as many tables as you can in each round and you will play for 9 rounds. The elapsed time is displayed above the input box as you can see in the image displayed above.

Counting Tips: Of course you can count the zeros in anyway 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.

Total earnings per round

To clarify, your total compensation in each round will be :The number of tables you solve during the round n_t at the rate of 20 cents each. This is summarized as follows:

$$E_t = n_t * 20$$

Total earnings The total earnings of the experiment are the sum of compensations for each round. This is summarized as follows

$$E = E_1 + \dots + E_9$$

For further comprehension we provide some examples.

Example 1

Assume that in some round you solved two tables correctly. Your earnings are therefore:

$$Earnings = 2 * 20 = 40$$

Thus, you earn a total of 40 cents or 0.4 Euros.

Example 2

Assume that in some round you solved three tables correctly. Your earnings are therefore:

$$Earnings = 3 * 20 = 60$$

Thus, you earn a total of 60 cents or 0.6 Euros.

Example 3.

Assume that in some round you solved three tables correctly but you had 4 fails. Your earnings are therefore:

$$Earnings = 3 * 20 - 20 = 40$$

Thus, you earn a total of 40 cents or 0.4 Euros.

Questions

To be sure that you understood correctly this instructions you need to answer the following questions.

1. You solved four tables correctly. Your earnings are therefore:
2. You solved two tables correctly and had one "fail". Your earnings are therefore:

3. You solved four tables correctly and had seven "fails". Your earnings are therefore:

Final Remarks

After you are done please wait in your chair a questionnaire will be provided and the experimenter will go to your seat and proceed to get your bank account number to pay your total gains in the experiment. Please keep in mind throughout the experiment that you cannot use the Enter key to proceed to the next screen. Instead you are advised to use the mouse to click the respective buttons.

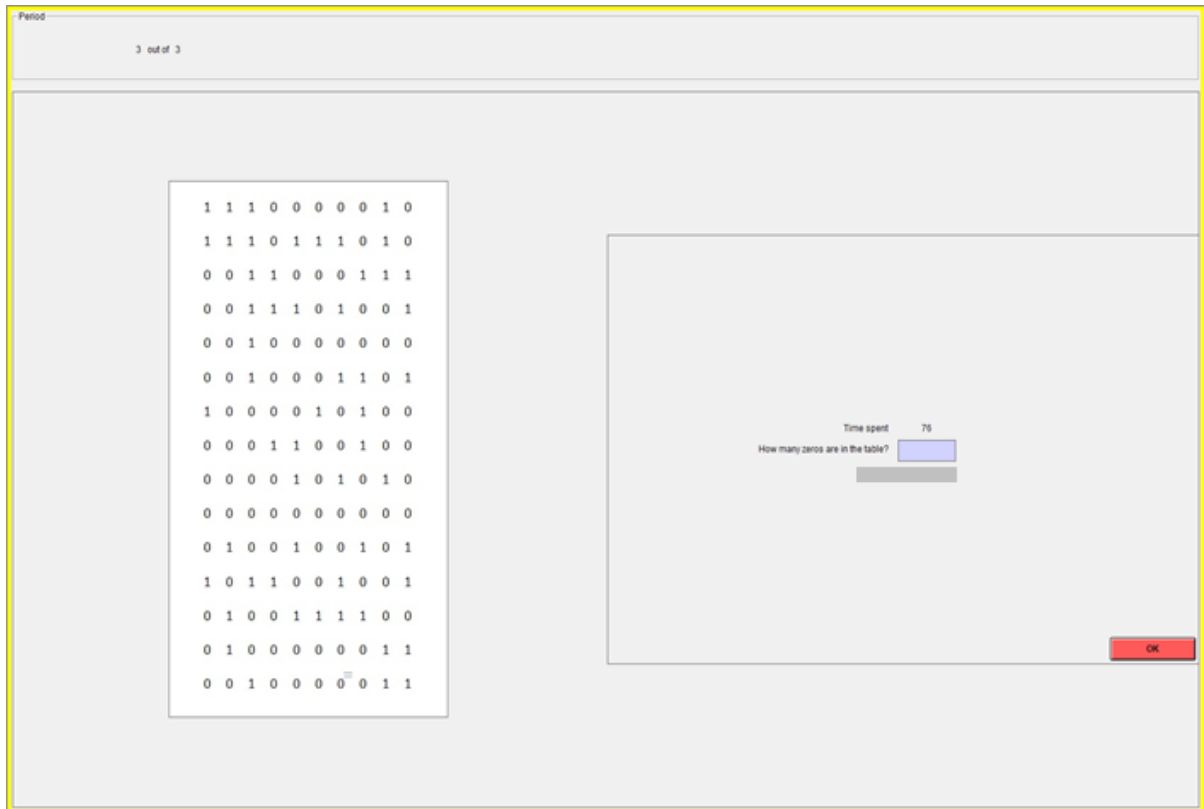
Experiment Instructions goal contract treatment

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

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.

What do you have to do?

The experiment consists of a sequence of 9 rounds of 5 minutes each. In each round your task is to count zeros in a series of tables. The figure shows the work screen that you will use later.



When faced with this screen you have to insert in the box, located on the right side of the screen, the exact number of zeros that appear in the table set in the left side of the screen. After you have entered the number, click the OK button. If you entered the correct number of zeros you will receive 20 cents, but if your input was wrong you will receive nothing and you will be accounted with one "fail". Furthermore, with three "fails" 20 cents will be deducted from your total earnings. Immediately after you are notified whether you solved a table correctly or not, a new table will be displayed and you are again expected to count the number of zeros. This procedure is repeated until the time is up. In total, you have 5 minutes per round. In other words, you have 5 minutes to solve as many tables as you can in each round and you will play for 9 rounds. The elapsed time is displayed above the input box as you can see in the image displayed above.

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.

Your prediction and its prediction:

At the beginning of each round you will be asked to predict the number of tables that you think you could solve correctly in that specific round. If your prediction is correct or you surpass it, we will provide you an additional monetary reward.

At this point it is important to point out that the monetary reward is not fixed but it increases as the prediction increases. This is, if your prediction is high and you achieve or surpass it you will be given a high monetary reward. But, if your prediction 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 prediction and you cannot accomplish it you will gain no monetary reward. The formula that we use for this monetary reward is given next with n_t the number of correct tables that you solved in a round and a_t your prediction for the same round.

$$B_t = \begin{cases} a_t * 20 & \text{if } a_t \leq n_t \\ 0 & \text{if } a_t > n_t \end{cases}$$

Total earnings per round

To clarify, your total compensation in each round will be :The number of tables you solve during the round n_t at the rate of 20 cents each. Plus the additional monetary reward in case you achieve or surpass your prediction. This is summarized as follows:

$$E_t = n_t * 20 + B_t$$

Total earnings The total earnings of the experiment are the sum of compensations for each round. This is summarized as follows

$$E = E_1 + \dots + E_9$$

For further comprehension we provide some examples.

Example 1 $a_t = n_t = 2$

Assume that in some round you predicted to solve two tables correctly and you actually solved two tables correctly. Your earnings are therefore

$$Earnings = 2 * 20 + 2 * 20 = 80$$

Thus, you earn a total of 80 cents.

Example 2 $a_t = 3, n_t = 2$

Assume that in some round you predicted to solve three tables correctly and you actually solved two tables correctly. Your earnings are therefore:

$$Earnings = 2 * 20 = 40$$

Thus, you earn a total of 40 cents.

Example 3 $a_t = 2, n_t = 3$

Assume that in some round you predicted to solve two tables correctly and you actually solved three tables correctly. Your earnings are therefore:

$$Earnings = 3 * 20 + 2 * 20 = 100$$

Thus, you earn a total of 100 cents.

Example 4 $a_t = 2, n_t = 3$

Assume that you predicted to solve two tables correctly and you solved three tables correctly. But you miscounted a fourth, fifth and sixth table. Your earnings are therefore:

$$Earnings = 3 * 20 - 20 + 2 * 20 = 80$$

Thus, you earn a total of 80 cents.

Questions

To be sure that you understood correctly this instructions you need to answer the following questions.

1. You solved three tables correctly; you predicted to solve two tables correctly. Your earnings are therefore:
2. You solved one table correctly; you predicted to solve one table correctly. Your earnings are therefore:
3. You solved two tables correctly; you predicted to solve one table correctly. Your earnings are therefore:
4. You solved two tables correctly; you predicted to solve one table correctly. You miscounted 4 tables. Your earnings are therefore:

Final Remarks

After you are done please wait in your chair, a questionnaire will be provided and the experimenter will go to your seat and proceed to pay your total gains in the experiment. Please keep in mind throughout the experiment that you cannot use the Enter key to proceed to the next screen. Instead you are advised to use the mouse to click the respective buttons.