

Brooks, Richard W.; Stremitzer, Alexander; Tontrup, Stephan

**Working Paper**

## Stretch It but Don't Break It: The Hidden Cost of Contract Framing

Law & Economics Research Paper Series, No. 13-22

*Suggested Citation:* Brooks, Richard W.; Stremitzer, Alexander; Tontrup, Stephan (2017) : Stretch It but Don't Break It: The Hidden Cost of Contract Framing, Law & Economics Research Paper Series, No. 13-22, SSRN, Rochester, NY, <https://doi.org/10.2139/ssrn.2353733>

This Version is available at:

<https://hdl.handle.net/10419/335575>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

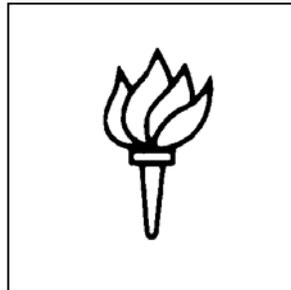
*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# NEW YORK UNIVERSITY

## SCHOOL OF LAW

LAW & ECONOMICS RESEARCH PAPER SERIES  
WORKING PAPER NO. 13-22



### Stretch It But Don't Break It: The Hidden Cost of Contract Framing

*Richard R. W. Brooks, Alexander Stremitzer, Stephan Tontrup*

September 2017

# Stretch It but Don't Break It: The Hidden Cost of Contract Framing\*

Richard R. W. Brooks<sup>†</sup>  
Columbia Law School

Alexander Stremitzer<sup>‡</sup>  
UCLA Law School

Stephan Tontrup<sup>§</sup>  
New York University School of Law

August 22, 2017

## Abstract

Recent research suggests that loss-framed contracts are an effective instrument for principals to maximize the effort of their agents. Framing effects arise from defining quantity and quality thresholds that vary the salience of losses and gains even while preserving the payoff equivalence of the underlying contract. While plausible interpretations of Prospect theory's loss-aversion insight suggest that a loss frame would lead to more effort, we show that contract thresholds also exert a norm-framing effect on performance that can trump the impact of loss aversion. Loss framing therefore carries a risk: poorly selected thresholds may reduce effort. Principals may prefer to avoid this risk by offering contracts that impose no threshold at all.

*Keywords:* contracts, loss-framing, norm-framing, loss aversion, prospect theory, expressed thresholds, worker productivity.

*JEL-Classification:* C91, D02, J33, K12.

---

\*The authors are grateful to Johannes Abeler, Thomas Baumann, Joe Doherty, Florian Ederer, Christoph Engel, Florian Englmaier, Robert Gold, Jason Oh, Peter Kuhn, Bentley McLeod, Ted Parson, Matteo Rizzoli, Andreas Roeder, Patrick Schmitz, and Rebecca Stone. We are also grateful to Seminar Audiences at UC Santa Barbara, the University of Mannheim, George Mason University, and at the ALEA conference at Vanderbilt. We thank Tyler O'Brien, Jaimini Parekh, James Davis, Yijia Lu, Sarvenaz Tarhani, Ben Nyblade and Henry Kim for excellent research assistance. Financial Support by the Lowell Milken Center for Business Law and Business Policy is gratefully acknowledged.

<sup>†</sup>Columbia Law School, 435 West 116th Street, New York, NY 10025, United States, rbrooks@law.columbia.edu.

<sup>‡</sup>UCLA Law School, 385 Charles E. Young Drive, 1242 Law Building, Los Angeles, CA 90095, United States, stremitzer@law.ucla.edu.

<sup>§</sup>New York University, School of Law, 40 Washington Square S, New York, NY 10012 s.tontrup@googlemail.com.

# 1 Introduction

Law sanctions conduct falling short of what is normatively established, but seldom rewards behavior that goes beyond the legal threshold. It is a basic asymmetry that pervades in public and private law. Contract law, for instance, protects buyers from downward deviations in sellers' performances, whether in terms of quantity, quality or timing of delivery among other nonconforming shortfalls of contract stipulations. Aggrieved buyers in these cases are typically afforded monetary compensation or other relief, such as a right to demand that sellers cure their nonconforming performances or a right to rescind the contract. There is, however, no right to additional compensation for a seller who produces a greater quantity or quality than required or who delivers earlier than stipulated to the buyer's advantage.<sup>1</sup> As a consequence, a kink results in performance incentives, where the seller has high-powered incentives below the contractual specifications but weak or no incentives to perform above and beyond what is specified in the contract.

Buyers may remove the kink by offering contractual terms stipulating performance at very high levels. If the buyer offers a high enough price, sellers will accept, and the contract preserves the sellers' performance incentives over a greater range of possible performance. Notwithstanding the simplicity of this solution, it does present some challenges.<sup>2</sup> Breach – with all its associated transaction costs – may occur more frequently. Some of these costs may be avoided by stipulating damages in the contract, but courts may still refuse to enforce

---

<sup>1</sup>Under certain strict conditions, an “overperforming” seller may have a right to recovery under a theory of restitution, quantum meruit, quantum valebat, quasi-contract etc., all of which largely rely on the buyer being unjustly enriched by the overperformer. Without *unjust* enrichment, there is generally no right to recover compensation for engaging in conduct not legally required. “If a performance is rendered by one person without any request by another, it is very unlikely that this person will be under a legal duty to pay compensation.” Arthur Corbin, *Contracts* §234 (1950).

<sup>2</sup>Take, for example, the case of Oeresund A/S, a consortium overseeing the construction of a \$10 billion dollar bridge and tunnel link connecting the cities of Copenhagen, Denmark and Malmo, Sweden. Since the planned link would accommodate rails, as well as cars, Oeresund A/S feared that vibrations caused by trains, particularly on the suspension bridge section of the link, would make drivers feel unsafe and thereby reduce demand. To address this concern Oeresund A/S contracted with a firm (hereafter “supplier”) competent in the design and use of advanced cable and damper technology. Oeresund specified an extremely low vibration level in the contract, which was impossible to achieve with current technology. Although Oeresund offered a high up-front price for meeting the specified vibration level, all parties anticipated that the supplier would in all likelihood fail to achieve this specification, which would trigger a schedule of stipulated damages paid to Oeresund based on how far short of the quality level the supplier's performance fell.

stipulated damages as “penalties.” Moreover, a seller’s behavioral motivation to perform may be weakened by knowing that, despite his effort, breach and penalties are almost inevitable.<sup>3</sup>

To avoid these potential costs, the parties may choose an alternative arrangement. Rather than allow only for sanctions when performance falls short, whether through stipulated damages or by relying on background legal remedies for downward deviations, the parties may agree to an intermediate performance level with a bonus for going beyond: i.e., rewarding overperformance in addition to punishing shortfalls. This was the contractual scheme that Caltrans, the California transportation agency, offered to its contractors following the 1994 Northridge earthquake, which resulted in the collapse of two bridges on the Santa Monica Freeway, one of the world’s busiest roadways. Caltrans offered a contract with substantial performance incentives and penalties: a \$200,000 per day bonus for completing the project ahead of schedule and a \$200,000 penalty for each day the project was behind schedule. Under this incentive scheme, reconstruction of the bridges was completed in a little over two months, 74 days ahead of the stipulated deadline (Eggers, 1997).<sup>4</sup> The contract created a symmetric incentive scheme, where the threshold stipulated in the contract no longer defined a cut-off point beyond which additional performance goes uncompensated, but instead became a framing device, allowing the drafting party to vary the salience of losses and gains.

To illustrate the point, consider a contract calling for 50 units of specified goods for 50 dollars. Under the default regime, if the seller delivers 40 units the buyer is entitled to damages of 10 dollars (assuming a linear apportionment between the shortfall and damages.) However, if the seller delivers 60 units, she cannot demand 60 dollars from the buyer. The incentive scheme is asymmetric, providing compensation up to the threshold but not beyond.

---

<sup>3</sup>The initial thought at Oeresund was that a nearly impossible-to-reach threshold matched with an astronomical contract price would give the supplier high-powered incentives to produce at its best. Yet, when the contract was underway, Oeresund became aware that the contract design may have sapped the motivation of the supplier to exert its greatest effort, knowing that failure was practically inevitable. Following this experience, officials at Oeresund A/S came to believe that it might have been more effective to specify an intermediate quality threshold and offer a bonus for exceeding it, rather than a contract where a penalty was certain no matter how hard the supplier worked.

<sup>4</sup>To complete the project so early, the contractor used up to 400 workers a day and kept crews on the job 24 hours a day. The \$13.8 million the contractor received in performance bonuses was more than offset by the estimated \$74 million in savings to the local economy and \$12 million in contract administration savings thanks to the shortened schedule.

By increasing the number of delivered units up to the threshold, the seller can increase her payoff by reducing damage payments, but beyond the threshold any increase in the number of units will not increase her payoff. If, however, the contract stipulates 50 units for 50 dollars and a bonus of 1 dollar per unit delivered beyond the threshold, the seller is entitled to an amount which equals the number of units times \$1. The same would hold true if the threshold were set at 70 or 100 units. The threshold no longer has any effect on the parties' monetary payoffs. In such a symmetric payoff regime the threshold becomes a pure framing device, expressing only a reference point or a norm. The key question is then which frame produces the best performance incentives. Our analysis seeks to answer this question.

We propose a novel theory of contract thresholds. We posit that contract thresholds produce framing effects not only through loss aversion (*loss-framing*) as was previously argued in the literature (Hossain and List, 2012), but also through a separate mechanism, which we label the "norm effect of thresholds" (*norm-framing*). In addition to serving as a pivot to frame a loss or gain, contract thresholds have a suggestive effect which influences behavior. Thresholds may communicate a norm of either how others tend to perform (a positive norm) or what performance is expected (a prescriptive norm) or both.

In order to test our theory we conducted an online experiment involving real effort. Participants were presented with a table with 200 digits between 1 and 9. The task consisted in counting how often a specified digit, e.g., "1" or "4", occurred in the table. After entering the correct number, participants could decide whether to continue or to stop the task. If they chose to continue they were shown a new screen, with a different table of digits.<sup>5</sup> Participants earned €1 for every completed screen. The framing of their earnings, however, was varied across four treatment conditions. The first condition offered participants a "plain-vanilla" bonus contract, which promised participants €1 per completed screen. The contract did not prescribe any target quantity. By contrast, the three other treatments offered contracts with expressed thresholds—5, 15, or 50 screens, each associated with earnings of €5, €15,

---

<sup>5</sup>Participants could continue with the experiment indefinitely if they so chose, without any arbitrary upper limit imposed by time or anything else. The "effort" consists only of the subject's willingness to keep counting tables.

and €50, respectively. We label the easy-to-meet threshold (i.e., 5 screens) the “low-bar” contract, the more demanding intermediate threshold (15 screens) the “stretching” contract, and the highest threshold (50 screens) the “extreme-effort” contract.<sup>6</sup> Under each of these three contracts, participants earned a bonus of one additional Euro for each screen they completed beyond the stipulated threshold and they faced a penalty of one Euro for each screen they fell short of the stipulated threshold. In other words, under all four contracts that constitute our treatment conditions, participants were offered exactly the same payoffs, namely €1 per completed screen.

To illustrate, assume that a participant completes 12 screens. Under the plain-vanilla contract he gets €12 (12 x €1). Under the low-bar contract, with a threshold of 5, he gets €5 plus a bonus of €7. Under the stretching contract, with a threshold of 15, he gets €15 minus a penalty of €3. Under the extreme-effort contract, stipulating a threshold of 50, he gets €50 minus a penalty of €38. At every level of performance, participants receive exactly the same payoff under each contract. They face the same linear incentive scheme in each condition; only the framing of the contract differs.

The basic results from our experiment indicate that contract framing does indeed matter, as observed in the prior literature. Our findings, however, suggest something largely overlooked in the literature: there is a non-monotonic relationship between the threshold and the performance. Compared to plain-vanilla terms, which state no explicit thresholds, our manipulations produce significantly different mean effort levels, damping or increasing effort depending on the suggested threshold. Specifically, the low-bar and extreme-effort treatments are associated with lower mean effort and the stretching treatment results in greater mean effort than observed under the plain-vanilla condition.

A trade-off is apparent. While setting a threshold at a demanding yet realistic level seems to increase the effort level of an agent, a threshold too low or too high will lead to a lower level of effort. Given the risk associated with selecting a threshold which is too

---

<sup>6</sup>These labels were not shared with the subjects. We use them here to conveniently distinguish the four treatment conditions: “plain-vanilla”, “low-bar”, “stretching” and “extreme-effort”.

low or too high, it may be better not to specify a threshold at all. Hence, framing is not a costless strategy for a principal. When the principal lacks good information about an agent's production function, the best strategy may be to play it safe and offer a linear plain-vanilla contract without expressing a threshold. By highlighting this tradeoff, our paper qualifies the suggestion that loss-framing generally leads to better performance (Hossain and List, 2012).<sup>7</sup> Moreover, through an additional treatment designed to disentangle loss-framing and norm-framing, we find that our results are driven by both a loss-framing and a norm-framing effect.

Our paper contributes to the larger literature on self-control incentives where firms may impose work targets like production minimums or artificial deadlines in order to help their workers overcome time inconsistencies in preferences (Kaur et al., 2010). However, unlike this literature, we study pure framing in which agents do not suffer disproportionate monetary penalties for failing to meet their targets.

Our paper proceeds as follows: Section 2 presents a simple model with regard to the effect of different contract thresholds on agents' effort, from which we derive our hypotheses. Section 3 presents our experimental design. Our results are presented and discussed in Section 4. Section 5 considers the application of these findings to various legal settings.

## 2 Theory and Hypotheses

The effectiveness of contract framing is usually explained by appealing to Prospect theory's concept of loss aversion (Hossain and List, 2012). Prospect theory tells us that 'losses loom larger than gains,'—that is, individuals prefer avoiding losses to acquiring commensurate gains—a concept commonly referred to as loss aversion (Kahneman and Tversky, 1979). Therefore, framing a transaction in terms of a loss should provoke a stronger response than an economically equivalent transaction framed in terms of a gain. Building on this insight, researchers have recently turned their attention to the framing of contracts. By establishing

---

<sup>7</sup>Our observations are nicely summarized in Voltaire's observation that perfection is the enemy of the good ("Le mieux est l'ennemi du bien").

thresholds in contracts around which earnings are presented as losses or gains, researchers have found that individuals (Brooks, Stremitzler, and Tontrup, 2011; Fryer, Levitt, List, and Sadoff, 2012) and teams (Hossain and List, 2012) exert greater effort under loss-framed contracts than under payoff-equivalent contracts framed in terms of gains.

Prospect theory predicts that contracts imposing reasonable thresholds should (weakly) *increase* observed effort compared to the plain-vanilla contract, because they express a threshold that frames performance levels below it as losses, while the plain-vanilla contract frames such performance levels as neutral accomplishments or gains. Only for extreme thresholds would Prospect theory predict lower effort. This is because incremental losses carry less weight as they grow more distant from the agent's reference point. As the reference point is shifted to the right, the agent therefore faces portions of the utility function with ever lower marginal utility, causing the agent to exert lower effort.

Besides Prospect theory, we hypothesize an additional influence operating through norm-framing. Norm-framing assumes that subjects interpret the thresholds as communicating expectations about what they should do. Subjects therefore will want to conform with the norm and will derive additional utility from doing so. However, deviating from the norm, either by falling short of the norm or by exceeding it, will reduce this additional utility. As the agent exceeds the norm, norm-framing therefore will have a negative effect on marginal utility. By contrast, as the agent approaches the norm from below, norm-framing will have a positive effect on marginal utility. This will be true unless the agent finds himself too far away from fulfilling the norm, in which case the norm might discourage the agent such that norm-framing would have a negative effect on marginal utility. Norm-framing—if strong enough—will therefore cause the agent to meet the norm set by the thresholds. It follows that low thresholds will tend to lead to a decrease in an agent's effort level relative to the plain-vanilla contract, while high but realistic thresholds will tend to lead to an increase in an agent's effort.

**Prospect theory.** Prospect theory’s main property, that losses loom larger than gains, can be captured by assuming a kink in the utility function at the reference point. As an example, consider the following utility function:

$$U_{PT}(y, \bar{y}) = \begin{cases} y & \text{if } y \geq \bar{y} \\ \frac{7}{5}y - \frac{2}{5}\bar{y} & \text{if } \frac{3}{5}\bar{y} \leq y < \bar{y} \\ \frac{4}{5}y - \frac{1}{25}\bar{y} & \text{if } y < \frac{3}{5}\bar{y} \end{cases} . \quad (1)$$

where  $y$  is the output chosen by the agent and  $\bar{y}$  is the reference point set by the threshold stipulated in the contract. Note that the marginal utility below the reference point is greater than above the reference point ( $\frac{7}{5} > 1$ ). Also, marginal utility for very low values of  $y$  is low ( $\frac{3}{5} < 1$ ). This captures another key property of Prospect theory, namely, that incremental losses carry less weight as they grow more distant from the agent’s reference point.<sup>8</sup> Finally, we assume that an agent who accepts a loss-framed contract experiences a disutility when choosing zero effort, which can be avoided by rejecting the contract ( $-\frac{1}{25}\bar{y}$ ).<sup>9</sup> As this disutility increases with higher thresholds, there is also a greater chance that the agent will reject the contract.

Assuming that the agent’s cost function is given by  $c(y) = \frac{1}{20}y^2$ , an agent will choose the effort maximizing his payoff,  $\Pi_{PT}(y, \bar{y}) = U_{PT}(y, \bar{y}) - c(y)$  (labeled PT in Figure 1). We can derive the following lemma:

**Lemma 1** *An agent with preferences corresponding to Prospect theory will chose an effort level of 10 under the low-bar contract, an effort level of 14 under the stretching contract, and an effort level of 8 under the extreme-effort contract.*

**Proof.** The lemma follows from the first-order conditions for each of the three segments of our utility function (the function is concave on each segment). As  $U(y, \bar{y})$  is not differentiable over the entire interval we must also check the corner solutions,  $\bar{y}$ ,  $\frac{3}{5}\bar{y}$ , and 0. Finally,  $\Pi_{PT}(y, \bar{y})$  is positive for each of the maximizers, so that the agent’s participation constraint is met. ■

<sup>8</sup>In more elaborate models it is also common to assume concavity of the utility function above the threshold. However, we do not need this property for our argument.

<sup>9</sup>This makes the loss-framed contract a self-commitment device (see Imas et al., 2015). We chose the specific terms ( $-\frac{2}{5}\bar{y}$  and  $-\frac{1}{25}\bar{y}$ ) to assure continuity of  $U(y, \bar{y})$  at  $\bar{y}$  and  $\frac{3}{5}\bar{y}$ .

In other words, Prospect theory predicts an inverse U shape for the relationship between effort choice and the stipulated threshold. Putting the agent in a loss frame initially (weakly) increases the agent's effort. However, as the loss frame becomes too extreme, loss-framing dampens the agent's effort level.

**Norm-framing.** Norm-framing assumes that an agent interprets the threshold as expressing a norm with which he will seek to conform. He therefore derives an additional utility from meeting it. As the agent exceeds the norm, this additional utility is reduced as he does not like to deviate from the norm. However, as the agent approaches the norm from below, he experiences extra motivation beyond the monetary payoff (for example, because he anticipates the satisfaction of fulfilling the norm). This is true, unless the agent is so far away from the norm that he feels discouraged by it, knowing that there is no realistic way to meet the norm. As an example, consider the following utility function  $\Delta U_{NF}$  capturing the *additional* effect of norm-framing (which can be thought of as the effect of norm-framing if the agent derived no utility from his monetary payoff):

$$\Delta U_{NF}(y, \bar{y}) = \begin{cases} -\frac{3}{5}y + \frac{4}{5}\bar{y} & \text{if } y \geq \bar{y} \\ \frac{3}{5}y - \frac{2}{5}\bar{y} & \text{if } \frac{2}{5}\bar{y} \leq y < \bar{y} \\ -\frac{3}{10}y - \frac{1}{25}\bar{y} & \text{if } y < \frac{2}{5}\bar{y} \end{cases} . \quad (2)$$

The additional utility from meeting the threshold is  $\Delta U_{NF}(\bar{y}, \bar{y}) = \frac{1}{5}\bar{y}$ .<sup>10</sup> As the agent exceeds the norm, norm-framing has a negative effect on marginal utility ( $-\frac{3}{5}$ ). As the agent approaches the norm from below, there is a positive effect on marginal utility ( $\frac{3}{5}$ ). However, as the agent finds himself too far away from fulfilling the norm he feels discouraged, resulting in a negative effect on marginal utility ( $-\frac{3}{10}$ ). Finally, we assume that an agent who chooses zero effort experiences a disutility from facing a norm-framed contract, which can be avoided by rejecting the contract ( $-\frac{1}{25}\bar{y}$ ).<sup>11</sup>

Let us consider an agent who has preferences conforming to norm-framing. Assuming the agent is not also loss-averse, his utility is given by the sum of the utility derived from

<sup>10</sup>The additional utility increases with the threshold, capturing the idea that an agent derives a greater satisfaction from meeting a more ambitious norm.

<sup>11</sup>We chose the particular terms  $-\frac{2}{5}\bar{y}$  and  $-\frac{1}{25}\bar{y}$  to assure continuity of  $\Delta U_{NF}(y, \bar{y})$  at  $\bar{y}$  and  $\frac{2}{5}\bar{y}$ .

the monetary payoff,  $y$ , and the effect of norm-framing:<sup>12</sup>

$$U(y, \bar{y}) = y + \Delta U_{NF} = \begin{cases} \frac{2}{5}y + \frac{4}{5}\bar{y} & \text{if } y \geq \bar{y} \\ \frac{7}{5}y - \frac{2}{5}\bar{y} & \text{if } \frac{2}{5}\bar{y} \leq y < \bar{y} \\ \frac{7}{10}y - \frac{1}{25}\bar{y} & \text{if } y < \frac{2}{5}\bar{y} \end{cases} . \quad (3)$$

Assuming the above cost function,  $c(y) = \frac{1}{20}y^2$ , the agent will choose the effort maximizing  $\Pi_{NF}(y, \bar{y}) = U(y, \bar{y}) - c(y)$  (labeled NF in Figure 1). We can derive the following lemma:

**Lemma 2** *An agent with preferences corresponding to prospect theory and norm-framing will choose an effort level of 5 under the low-bar contract, an effort level of 15 under the stretching contract, and an effort level of 7 under the extreme-effort contract.*

**Proof.** First note that the first-order conditions give us 4, 16, and 7 as interior maximizers of the functions in each segment. For  $\bar{y} = 5$ , it follows that  $\Pi_{NF}(y, 5)$  is strictly decreasing on the interval  $[5, \infty)$ , and strictly increasing on the intervals  $[2, 5)$  and  $[0, 2)$ . It then follows from the continuity of  $\Pi_{NF}(y, 5)$  that the corner solution 5 maximizes the agent's utility. For  $\bar{y} = 15$ , it follows that  $\Pi_{NF}(y, 15)$  is strictly decreasing on the interval  $[15, \infty)$  and the function is strictly increasing on the intervals  $[6, 15)$  and  $[0, 6)$ . It then follows from the continuity of  $\Pi_{NF}(y, 15)$  that the corner solution 15 maximizes the agent's utility. Finally, for  $\bar{y} = 50$ ,  $U(y, \bar{y})$  is strictly decreasing on the interval  $[50, \infty)$  and on the interval  $[30, 50)$ . It follows from the continuity of  $U(y, \bar{y})$  that the interior solution 7 must maximize  $\Pi_{NF}(y, 50)$ . Finally,  $\Pi_{NF}(y, \bar{y})$  is positive for each of the maximizers, so that the agent's participation constraint is met. ■

The lemma reflects the strong tendency under norm-framing to conform with the norm. Mathematically, the intuition is that the agent's optimal effort choice will be given by the corner solution  $\bar{y}$  if the kink of  $\Delta U_{NF}$  at threshold  $\bar{y}$  is strong enough.

**Plain-vanilla contract.** We assume that the plain-vanilla contract, which does not stipulate any explicit threshold, does not set a norm. We also assume that the plain-vanilla

---

<sup>12</sup>If the agent was both loss-averse and susceptible to norm-framing his utility function would be given by  $U_{PT}(y, \bar{y}) + \Delta U_{NF}$ .

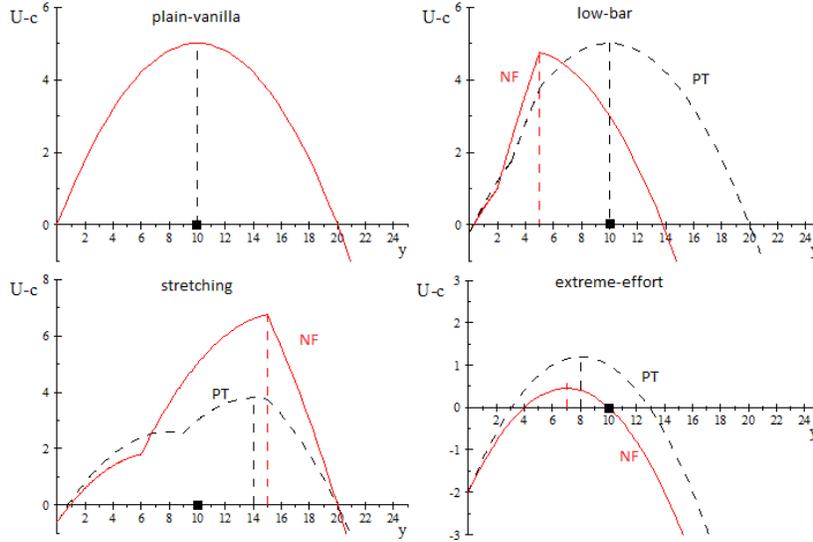


Figure 1: Equilibrium Effort Choices Over Treatments: PT vs NF.

contract either does not set a reference point or sets the reference point at the status quo.<sup>13</sup> The utility function induced by the plain-vanilla contract  $U_{PV}(y)$  is therefore given by:

$$U_{PV}(y) = y = U_{PT}(0, \bar{y}) \quad (4)$$

Assuming the above cost function, an agent will choose his effort to maximize  $\Pi_{PV} = U_{PV}(y) - c(y)$  (see Figure 1). We can derive the following lemma:

**Lemma 3** *Under the plain-vanilla contract, the agent chooses effort level 10.*

Summarizing the lemmas we can say that both norm-framing and loss-framing predict that the stretching contract should *increase* mean effort and that the extreme-effort contract should *decrease* mean effort compared to the plain-vanilla contract (to facilitate the comparison, we mark the effort induced by the plain-vanilla contract with a black dot in Figure 1). However, in the low-bar condition, only norm-framing predicts that the low-bar contract decreases mean effort relative to the plain-vanilla contract.<sup>14</sup> As we assume both elements to be operative, we can derive the following hypothesis:

<sup>13</sup>These two assumptions are the same in our example,  $U_{PV} = y = U_{PT}(0, \bar{y})$ . Below we discuss an additional option, namely, that the reference point is set at the agent's expectation of the effort level he is going to exert.

<sup>14</sup>If we assume concavity of the utility function above the threshold, Prospect theory would predict that effort levels under the low-bar contract increase relative to the plain-vanilla contract (assuming that the plain

**Hypothesis 1** *In the low-bar treatment, mean effort is lower than under the plain-vanilla contract (H1.1). In the stretching treatment, mean effort is higher than under the plain-vanilla contract (H1.2). In the extreme-effort condition, the mean effort level decreases relative to the stretching condition and even falls below the level observed in the plain-vanilla contract (H1.3).*

If we were to find that mean effort is lower under the low-bar contract than under the plain-vanilla contract (H1.1) we may conclude that this is driven by norm-framing, given that loss-framing predicted no effect or the opposite effect. However, this argument depends crucially on the assumption that the plain-vanilla contract sets the reference point at the status quo,  $\bar{y} = 0$ , or does not set a reference point at all. While these assumptions seems plausible, there are other plausible assumptions about the location of the reference point that would account for the pattern predicted by H1.1 within the framework of Prospect theory.<sup>15</sup>

In order to establish that the drag-down effect under the low-bar contract is indeed driven by norm-framing, we run an additional treatment designed to “switch off” the norm-framing effect of the contractual threshold. We label this the “random-threshold” treatment, wherein subjects are offered either a randomly selected low-bar or extreme-effort contract. Since the subjects know the threshold is determined randomly, they should not infer that the threshold communicates a norm of what is expected of them. If only norm-framing causes the drag-down effect under the low-bar condition, then under a randomly chosen low-bar threshold mean effort should rise to the same level as under the plain-vanilla contract. This leads us to the following hypothesis:

**Hypothesis 2** *The drag-down effect predicted by Hypothesis 1.1 disappears under the random-*

---

*vanilla contract sets the reference point to the status quo level). The effect predicted by Prospect theory and the effect predicted by norm-framing would then be countervailing. As mentioned in the result section we indeed see such an effect.*

<sup>15</sup> Assume, e.g., that the plain-vanilla contract sets the reference point at an agent’s payoff expectations, as suggested by Koszegi and Rabin (2006). If those payoff expectations happen to be sufficiently high ( $\bar{y} \geq 20$ ), Prospect theory would predict that the low-bar contract leads to *lower* mean effort than the plain-vanilla contract. In other words, it can be argued that Prospect theory’s prediction about how mean effort under the plain-vanilla contract compares to the mean effort under the low-bar contract is ambiguous. This reflects a general concern with Prospect theory, which is very sensitive to different assumptions about the location of the reference point.

*low-bar treatment.*

### 3 Design

Participants were given the opportunity to enter into a contract to perform a real effort task in exchange for money.<sup>16</sup> They were instructed that if they rejected the contract they would receive no payment based on the real effort task. Subjects who rejected the contract were immediately directed to the second stage of the experiment where they and all other participants were asked to complete a pair of monetarily incentivized economic and psychological tests (see more details below).

Participants who accepted the contract offer were directed to begin the task. The task consisted of counting how often a specified digit, e.g., “1” or “4”, occurred in a table containing 200 digits ranging from 1 to 9. Given that the task did not require any special skills, we assume that performance was a function only of effort (see Abeler, Falk, Goette, and Huffman, 2011). After scanning the table and counting the occurrences of the specified digit, participants entered the count into a screen dialog box. An answer was considered correct if it fell within a range of  $\pm 2$  of the true value. For example, if the correct number of “3” digits was 42 while the participant counted 40, the result was treated as correct. If participants gave an answer outside of this tolerated margin, they could retry counting the table as often as they wished, but had to wait 15 seconds after each failed trial, before they could make a new input.<sup>17</sup> Participants were required to make some screen input at least every three minutes; failure to do so terminated the experiment. After each successfully completed screen, participants were asked whether they wanted to continue on to the next screen. If they decided to go on, a new table was displayed asking participants to count another randomly specified digit; if subjects chose to stop the task, they were directed to

---

<sup>16</sup>Subjects were given instructions that emphasized the experiment involved real contracts. As contract law requires, but unlike in most other experiments, we did not impose the contract; rather, participants were free to accept or reject the offers. After being presented with the contract terms, participants decided whether to accept or reject the offer (see contractual terms described in detail below).

<sup>17</sup>We introduced this wait time after each failed trial in order to dissuade participants from guessing repeatedly without actually counting any digits. Counting one table takes between 45 seconds and one minute such that a wait time of 15 seconds seems sufficient to achieve this goal.

the incentivized economic and psychological tests.

As previously described, we implemented four treatments that varied the threshold and one baseline condition—the so-called plain-vanilla condition—which did not express a threshold, but simply offered €1 for each successfully completed screen. We conducted the plain-vanilla treatment in advance and found that participants completed an average of 10.4 screens. We used this data to calibrate the thresholds of the treatments, which have the same linear payment scheme as the plain-vanilla condition. The low-bar threshold (5 screens) was set clearly below the mean effort participants exerted under the plain-vanilla contract. The stretching threshold (15 screens) obliged participants to increase their efforts by nearly 50% in comparison to the average performance under the plain-vanilla contract. The extreme-effort threshold (50 screens) called for an extremely high level, demanding approximately 5 times the mean performance observed under the plain-vanilla contract.<sup>18</sup> Finally, in the random-threshold treatment, the threshold was randomly determined. Subjects were offered either a low-bar or an extreme-effort contract. Subjects learned that the threshold was determined randomly before they had to decide whether they wanted to accept or reject the offer. However, they were not informed that only thresholds 5 and 50 could be selected in order to avoid the possibility that they would infer some target quantity the principal might expect them to reach (e.g., some quantity between 5 and 50).

We evaluated the effort participants exerted under the different treatments along two dimensions: quantity and quality. Quantity is measured by the number of successfully completed screens. Quality of performance is determined by the accuracy with which participants completed the task. We recorded each count that participants entered and calculated by how much the entry deviated from the true number of digits. We counted both failed trials (where participants had to repeat, if they wanted to continue) and successful, but not necessarily accurate, trials (where participants came within  $\pm 2$  of the correct answer).

---

<sup>18</sup>Furthermore, it required participants to invest notably more time than is usually expected in online experiments. Online studies are typically completed within approximately 15 to 20 minutes rather than a full hour. Note, though, that participants were informed about the duration of the experiment in the email that invited them to the study. They also learned that the actual time needed would depend on their performance (see Musch and Reips, 2000).

In addition to effort measures, we also elicited participants' willingness to enter into the contract. The efficiency of a particular contract depends not only on how people perform under its terms, but also on how likely they are to agree to incur the contractual obligation in the first place. Unlike in most other experiments, but in line with the formation requirements of a legally enforceable contract, participants were free to accept or reject the contract offers.<sup>19</sup> This real contract setting also allows us to study the impact of different terms on acceptance rates. If rates differ, it follows that participants are more willing to accept some terms than others, which is relevant for a principal who wants to reduce the likelihood that offers are rejected.

The experiment was conducted online using the server of the Max Planck Institute of Economics. We decided to conduct the experiment online because we wanted participants to have real opportunity costs when deciding whether to continue with the task. We measured effort by eliciting the point where participants prefer some other activity over continuing with the experiment. In a laboratory setting, participants have low opportunity costs because they cannot leave until the session is over.<sup>20</sup> By contrast, at home, participants can easily stop and pick up a preferred activity. Even if we had conducted individual sessions in order to allow participants to leave without influencing others, coming to the laboratory creates sunk costs that make participants less likely to reject a contract without earning adequate compensation for their participation.<sup>21</sup>

All 251 participants were students of the University of Münster, Germany, with various majors; 90% of them had not participated in an economic experiment before. Participants were sent an invitation by email via the mail-server of the university. The email did not de-

---

<sup>19</sup>Most of the literature does not endogenize the acceptance decision but assumes that participants have entered into a contract under given terms (see, e.g., Falk and Kosfeld, 2006; Fehr, Klein, and Schmidt, 2007). In our design we want to account for the possibility that behavior changes depending on whether parties explicitly assent to the contract (see Hoppe and Schmitz, 2011, for a similar design.) Lazear, Malmendier, and Weber (2012) have recently shown in an experimental study how the possibility to opt out of an economic environment create selection effects (in their case the willingness to share money with other participants).

<sup>20</sup>If participants were allowed to leave the session early in a laboratory setting, this would be observed by the other participants. Observing others leaving would likely influence participants' decisions about when to stop the task. The only way this could have been avoided is by running 251 individual sessions.

<sup>21</sup>Show-up fees in lab experiments can mitigate this problem but can never be tailored to the individual participant.

scribe the purpose of the study in order to avoid participants self-selecting into an experiment in which they were interested. The invitation contained a link that directed participants to the website of the experiment. It was active only for a single login. Participants had to complete the stages of the experiment within strict time limits and were kept informed of this fact with constant screen messages. If participants logged out or did not finish stages within these time limits, the experiment was automatically aborted and participants were notified that they were excluded from the experiment. We set those time limits to force participants to focus on the task and block internet distractions that can easily distort results of online studies. Participants were informed up front about the amount of time they would need to complete the whole study, thereby reducing the likelihood that participants would have to break off the experiment because they ran out of time. The online instructions given to participants are presented in Appendix C.<sup>22</sup>

The real effort experiment was followed by two psychometric tests. The first test measures loss aversion by giving the participants an opportunity to participate in two lotteries.<sup>23</sup> The first lottery presents participants with a 1/2 probability of winning €8 and a 1/2 probability of losing €5. The second lottery has the same payoffs, but is repeated six times and thus lowers the probability of suffering an overall loss. An unbiased participant should play both lotteries, since participating yields a gain in expectations. We classify participants into two different categories: 1) the “loss averse” type who rejects at least one of the lotteries, and 2) the unbiased “rational” type who participates in both lotteries. The second test, a cognitive reflection test, measures participants’ impulsiveness (CRT, see Frederick, 2005). The test consists of a set of three questions, which participants have a total of 90 seconds to complete. Questions are designed such that participants’ initial impulses lead them to an incorrect answer. The test therefore measures the participants’ ability to think beyond their initial impulse and reach the correct but counter-intuitive answer.

---

<sup>22</sup>We confirmed that participants could understand the instructions by posing control questions in an offline pilot session. Participants had no difficulties in correctly calculating their earnings based on the different contract terms. We programmed the experiment using the open-source survey application LimeSurvey. See <http://www.limesurvey.org/>.

<sup>23</sup>The test was developed and used by Goette, Hoffman, and Fehr (2004).

All aspects of the experiment, including the follow-up behavioral tests, were incentivized and payoffs depended on the participants' own decisions. When calibrating the payoffs we made sure that participants could expect to earn slightly more than in a regular student job in and outside of the university in order to make incentives comparable to a normal working environment. A student job in Münster would offer approximately €8 for an hour. In our real effort task, participants needed less than one minute to count a table and move to the next screen. They could read the instructions in 5 minutes which would leave them with an hourly wage of €55. Since only every 5th participant was randomly selected for payment participants earned €11 in expectation. On average, our participants therefore earned 1/3 more than in a regular student job.<sup>24</sup>

## 4 Results

### 4.1 Treatment Effect on Quantity

We first present how participants' mean effort levels differed across treatments (Table 1 and Figure 2). Observe that the plain-vanilla contract, which does not express a quantity threshold, leads to a mean effort of 10.4 screens, the low-bar contract (with expressed quantity of 5) leads to a mean effort of 6.2 screens, the stretching contract (with expressed quantity of 15) leads to a mean effort of 14.3 screens, and the extreme-effort contract (with expressed quantity of 50) leads to a mean effort of 8.4 screens.

Compared to the plain-vanilla contract, mean effort decreases under the low-bar contract (H1.1), increases under the stretching contract (H1.2), and decreases under the extreme-effort contract (H1.3). These differences are statistically significant (Mann-Whitney ranksum, Table 1 reports Z and p values).

In certain situations, a principal might care not about maximizing agents' mean effort but rather about minimizing the variance of effort levels induced by the contract. This

---

<sup>24</sup>Stochastic payouts are routinely used in experimental economics as there is evidence that paying out larger amounts at a lower probability simulates high stakes: that is, incentives for participants are stronger than if they are paid smaller amounts at a higher probability, even though expected payoffs are equivalent (Laury, 2006; Laury and Holt, 2008).

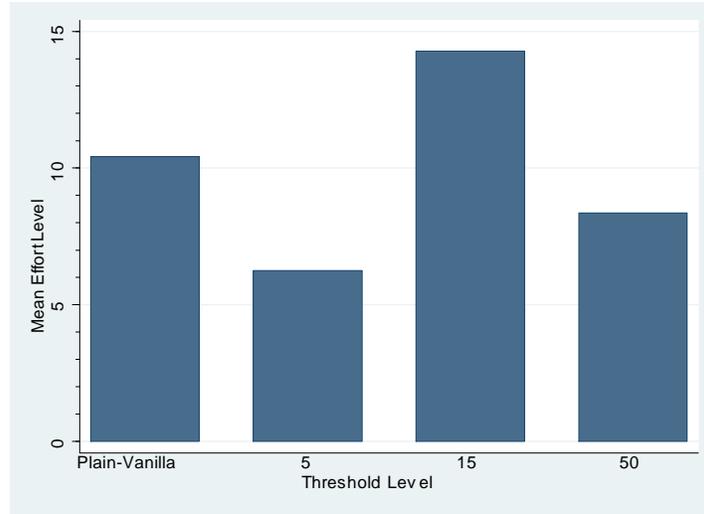


Figure 2: Mean Effort Levels Depending on Specified Quantity Threshold.

Table 1: Mean Effort Levels Compared Across Treatments (Mann-Whitney)

contract type	threshold	mean [Std. Dev.]	contract type			
			plain-vanilla	low-bar	stretching	extreme-effort
plain-vanilla	—	10.4 (N=50) [7.3]	—	$Z = 3.3^{***}$ $p < 0.01$	$Z = -2.5^{**}$ $p = 0.01$	$Z = 2.2^{**}$ $p = 0.03$
low-bar	5	6.2 (N=50) [3.9]		—	$Z = -5.2^{***}$ $p < 0.01$	$Z = 0.1$ $p = 0.96$
stretching	15	14.3 (N=54) [8.1]			—	$Z = 3.6^{***}$ $p < 0.01$
extreme-effort	50	8.4 (N=39) [8.5]				—

seems particularly plausible when setting safety regulations when the principal likely wants to ensure that every worker reliably exerts a particular effort level. Our results suggest that the variance in effort levels is lowest for the low-bar contract. (See standard deviations in Table 1 and the full distributions in Figure 3). Levene’s test is significant at below the 1% level when comparing the variance under the low-bar treatment with the variances under all other treatments. The variance under the plain-vanilla treatment seems to be slightly lower than under the stretching and extreme-effort treatments, but the difference in variance is

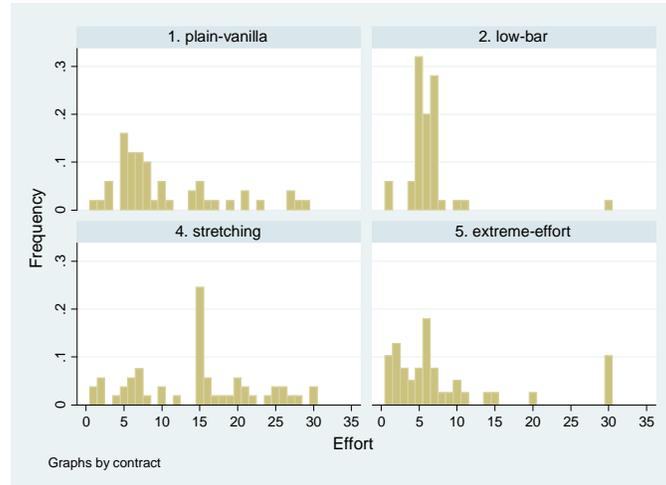


Figure 3: Distribution of effort levels across treatments.

not significant.<sup>25</sup>

When we consider only those participants who accepted the contract our analysis reflects the perspective of a principal who faces an unlimited supply of labor and does not care whether potential workers are turned off by the terms of a particular contract. There are, however, situations where a principal wants to increase the mean effort level of a fixed group of people. Then, participants rejecting the proposed contract impose a cost on the principal. We incorporate this scenario in our analysis by treating all participants who reject the contract as exerting zero effort and find that our qualitative results are unchanged (see Table 5 in Appendix A).<sup>26</sup>

<sup>25</sup>These results do not change when running the Brown-Forsythe test and a non-parametric permutation version of the Levene test (Good, 2006). See Appendix A, Table 6 for individual p-values and Appendix B for a description of the resampling procedure.

<sup>26</sup>Also in this scenario, the variance in effort levels is lowest in the low-bar treatment. Levene's test is significant at below the 1% level. The Brown-Forsythe test confirms this result. It leads to statistical significance below the 1% level when comparing the variance under the low-bar treatment with the variance under the plain-vanilla and the stretching treatment. When comparing the variance under the low-bar contract with the variance under the extreme-effort contract we get statistical significance at the 5% level ( $p=0.02$ ). Note that treating all participants who reject the contract as exerting zero effort artificially dampens the variance under the extreme-effort contract given the high rejection rate.

## 4.2 Disentangling loss-framing from norm-framing

Beyond our hypotheses about the aggregate pattern of mean effort levels we also had a specific prediction about the behavioral channels driving our effect. We argued that only norm-framing explains the drag-down effect under the low-bar condition. The support we found for H1.1 offers some evidence that this conjecture is correct. However, as we explained, this conclusion rests on the plausible assumption that the plain-vanilla contract sets the reference point at the status quo.<sup>27</sup> In order to find direct evidence that does not rely on this assumption, we ran the *random*-threshold treatment. This treatment is designed to switch off norm-framing by suppressing subjects' beliefs that the principal communicates an expected level of performance. This allows us to cleanly disentangle loss-framing from norm-framing. As loss-framing is switched off, the drag-down effect should disappear under a randomly chosen low-bar threshold. This is what we predicted in H2.

Conditional on opt in, we find that mean effort under the random-low-bar treatment is 10.7 screens, while it was 6.2 screens under the low-bar treatment. This effect is statistically highly significant (Mann-Whitney ranksum,  $p < 0.01$ ).<sup>28</sup> As subjects on average complete 10.4 screens under the plain-vanilla contract, the drag-down effect disappeared consistent with H2. Treating all participants who rejected the contract as exerting zero effort, we even find evidence for a reversal of the drag-down effect as the mean effort level under the random-low-bar treatment (12.2 screens) is higher than under the plain-vanilla treatment (8.4 screens). This effect is statistically significant (Mann-Whitney ranksum,  $p = 0.03$ ).<sup>29</sup>

---

<sup>27</sup>We report evidence consistent with this assumption in Appendix A. Norm-framing assumes that people want to conform to the norm expressed by the threshold. It has been demonstrated that cognitive reflection correlates with people's tendencies to conform to norms (see, Altmann and Falk, 2009). We would therefore assume that subjects with low scores on the cognitive reflection test would be more likely to follow their initial impulse and match the threshold than participants with high scores. We report weak suggestive evidence for this in Tables 7 and 8.

<sup>28</sup>Treating all participants who reject the contract as exerting zero effort, the mean effort levels are 12.2 and 5.0 screens, respectively. This difference is also statistically significant (Mann-Whitney ranksum,  $p < 0.01$ ).

<sup>29</sup>Looking at the random-extreme-effort treatment is less interesting as it does not allow us to disentangle norm-framing from loss-framing. Nonetheless, what we observe is consistent with our model's prediction. Under the random-extreme-effort treatment, mean effort—conditional on opt-in—is 11.8, while it was 8.4 under the extreme-effort treatment. Treating all participants who reject the contract as exerting zero effort, the mean effort levels are 10.3 and 5.3 screens, respectively. These two differences are statistically significant (Mann-Whitney ranksum,  $p < 0.01$ ). However, in both cases, effort levels are not significantly different

Our argument rests on the assumption that the random-threshold manipulation only affected norm-framing, while leaving loss-framing fully operative. If this is true, we can indeed conclude that the drag-down was entirely due to norm-framing because the effect disappears under the random-low-bar treatment. However, it could be argued that the random treatment not only switches off norm-framing but also affects loss-framing, as it might weaken subjects' perception of the threshold as a reference point. This would introduce a confound as our manipulation would have simultaneously weakened loss-framing *and* norm-framing. The vanishing of the drag-down effect under the random-low-bar treatment might then be entirely due to the weakened effect of loss-framing.

Table 2: Mean Effort: Loss Averse vs. Rational

	Loss-Averse	Rational
plain-vanilla	9.7 ( $N = 29$ )	11.5 ( $N = 21$ )
random-low-bar	12.4 ( $N = 23$ )	8.4 ( $N = 16$ )
Mann-Whitney	$Z = -2.3^{**}$ $p = 0.02$	$Z = 0.7$ $p = .49$

To counter this challenge, we compare the mean effort in the plain-vanilla treatment with the mean effort in the random-low-bar treatment for subjects scoring high on the psychometric test of loss aversion (loss averse types). We find that the mean effort of loss averse types is lower in the plain-vanilla treatment (9.7) than in the random-low-bar treatment (12.4). This difference is statistically significant (Mann-Whitney ranksum,  $p=0.02$ ), suggesting that loss-framing is operative under the random-low-bar treatment and pushes mean effort up compared to the plain-vanilla treatment rather than dragging it down (see Table 2).<sup>30</sup> This

from the effort levels induced by the plain-vanilla contract (Mann-Whitney ranksum,  $p=0.35$  and  $p=0.17$ ). Looking at Equation 3 and Figure 1, we can see that switching off norm-framing should cause mean effort to increase under the extreme-effort contract. Similarly, we would expect a decrease in contract rejection rates.

<sup>30</sup>It might seem odd that, for rational types, the mean effort level under the random-low-bar treatment (8.4) seems to be lower than under the plain-vanilla treatment (11.5). However, this effect is not statistically

result suggests that we need not worry about the possibility of a confound. A weakening of loss-framing could only account for the vanishing of the drag-down effect if loss-framing had a negative effect on mean effort.

### 4.3 Treatment Effect on Quality

In addition to effort measured in terms of *quantity* (that is, screens successfully completed), we also explore *quality* effort (measured by the accuracy with which participants perform the counting task). One concern could be that quantity and quality may be in conflict. In other words, contracts leading to higher quantity might lead to lower quality (lower accuracy of performance). In our experiment, we can distinguish two cases of inaccuracy. The first case of error occurs if a participant’s answer deviates from the true value by a margin of less than or equal to  $\pm 2$ . In this case, he can proceed to the next screen. In the second case, the participant’s answer falls outside the  $\pm 2$  margin of error, which forces the participant to re-count the table if he wants to proceed to the next screen. We form an inaccuracy score for each participant by adding up the deviations in both successful and failed attempts and dividing them by the number of successfully completed screens.<sup>31</sup> Table 3 shows how the mean inaccuracy score differs across treatments.

Table 3: Quality Levels Across Treatments

contract type	threshold	inaccuracy score	
		mean	median
plain-vanilla	– (N=50)	0.75	0.33
low-bar	5 (N=50)	1.23	0.77
stretching	15 (N=54)	0.60	0.21
extreme-effort	50 (N=39)	1.45	1.82

We find that effort quality is the lowest under the extreme-effort contract, closely followed by the low-bar contract, with mean inaccuracy scores of 1.45 and 1.23, respectively. Participants are significantly more accurate under both the plain-vanilla contract (0.75) and

significant (Mann-Whitney ranksum,  $p=0.49$ ).

<sup>31</sup>As we do not record the exact deviation for failed attempts, we multiply every failed attempt by 3, which equals the minimum amount of deviation which counts as a failure.

the stretching contract (0.60). The differences are statistically significant at below the 5% level (Mann-Whitney ranksum).<sup>32</sup> Moreover, comparing Tables 1 and 3, we observe that the plain-vanilla and stretching contracts induce higher effort levels than the low-bar and the extreme-effort contract along both the quantity and the quality dimensions. The positive relationship between quality and quantity is significant (Linear regression, Coef=-1.35,  $p < 0.051$ ).<sup>33</sup> This finding seems to run counter to the intuition of the multitasking model (Holmstrom and Milgrom, 1991), which would predict that stronger incentives to increase quantity lead to more shirking on the quality dimension.<sup>34</sup> An account consistent with our results might be that contract frames which motivate participants more make them exert more effort on both the quality and the quantity dimension of effort.<sup>35</sup>

#### 4.4 Contract Rejection Rates

Thresholds might also communicate information about the difficulty of the task. Participants might infer from a low threshold that they are only expected to count very few screens, suggesting that the task is difficult. In this scenario, subjects would take a high threshold to mean that the task is easy and a low threshold to mean that the task is difficult. This effect should only be present at the time of the opt-out decision whereas later, when subjects have started the task, they will have learned about its difficulty. We would therefore expect that opt-out rates would increase for lower thresholds as participants avoided the more difficult task.

Table 4 reports the contract rejection rates for the different treatments. Rejection rates

---

<sup>32</sup>Participants under the plain-vanilla treatment make fewer mistakes than under the extreme-effort treatment (Mann-Whitney test,  $p < 0.01$ ) and the low-bar treatment ( $p = 0.01$ ). The same holds true when comparing the stretching contract to the extreme-effort ( $p < 0.01$ ) and the low-bar treatment ( $p < 0.01$ ).

<sup>33</sup>Note that our quality measure is an inaccuracy score. Therefore a negative value of the coefficient implies a positive relationship between quantity and quality. The number of observations is 193.

<sup>34</sup>Note that in our context, incentives become stronger because framing affects the mapping of payoffs into agent's utility, while in the original Holmstrom and Milgrom (1991) paper, incentives become stronger because of a direct manipulation of payoffs.

<sup>35</sup>However, our design is not well suited to test the multi-tasking theory. While participants might be able to save time by lowering the accuracy with which they perform the task, it might also slow them down if participants cannot perfectly control their level of accuracy. This is because lower accuracy might increase the probability of having to recount screens as answers fall outside the margin of error.

under the plain-vanilla and the low-bar contract are 19.4%. They decrease for the stretching contract (16.9%) and then increase dramatically for the extreme-effort contract (37.1%). While rejection rates under the stretching and the low-bar contract do not differ statistically from one another or from those under the plain-vanilla contract, we find that the extreme-effort contract leads to a significantly higher rejection rate (37.1%) than any other contract. This behavior cannot be explained by assuming that participants infer from a high threshold that the required task is likely to be easy.

Table 4: Rejection Rates Across Treatments

contract		
type	threshold	rejection rate
plain-vanilla	– (N=50)	19.4%
low-bar	5 (N=50)	19.4%
stretching	15 (N=54)	16.9%
extreme-effort	50 (N=39)	37.1%

Imas et al. (2015) find that subjects might prefer loss contracts over gain contracts, suggesting that subjects select into loss contracts as a commitment device to improve performance. Our (statistically not significant) finding that subjects reject stretching contracts less often than low-bar contracts is consistent with this result. However, the fact that the rejection rate shoots up for the extreme-effort contract suggests that preferences for loss contracts might be non-monotonic.

## 5 Conclusion

Prospect theory suggests that if principals frame incentives as losses rather than gains, agents will exert greater effort. We show that this effect is not so simple. When thresholds are used to trigger loss aversion, contract framing entails a risk. In this paper we demonstrate that people tend to meet expressed thresholds around which earnings are framed as losses or gains. This norm-framing effect can trump the impact of loss aversion. By stipulating a demanding but reasonable stretching threshold, a principal can maximize the mean effort of

her employees. However, demanding too little may dampen the effort level compared to the expected performance under a simple contract that does not impose a threshold. In addition, a threshold pushed too high can similarly depress effort: an extreme loss frame may push the agent into an area of his value function where the slope is lower than under the plain-vanilla contract, and may undermine intrinsic motivation because the agent is unlikely to meet the threshold. This risk of contract framing is not marginal. The optimal stretching threshold will depend on the production function of the agent. But this information is often private and unobservable. Unless the principal has nuanced information about her agent's abilities and the technology available to the agent, or can invest in obtaining such information, a second-best plain-vanilla contract will often be a better choice.

Apart from informing contract drafters on how to set quantity/quality thresholds in contracts, the insights from this study can also be applied to other areas. Companies set production targets for their manufacturing workers and sales targets for their sales personnel. Law firms set billable hours targets for associates.<sup>36</sup> Often, in these cases, there is an informal target below which disciplinary action is taken, a formal target where bonus payments kick in, and an even higher informal target above which an employee is considered for promotion. Similarly, regulators concerned about lowering health care costs often allocate prescription drug budgets to doctors. If doctors prescribe more drugs they face a penalty. If they prescribe fewer drugs, they get a bonus. Our study informs regulators about how tightly to set those budgets.

Another interesting insight from this study is a new justification for the default contract law doctrine of implied warranties. The implied warranty of merchantability (UCC §§ 2-314) grants the buyer a remedy against the seller if the quality of the good sold does not meet reasonable expectations of buyers, even if the seller has not explicitly promised this quality. Since the warranty is implied, the doctrine saves the buyer from having to explicitly define what minimum quality he would be willing to accept. Our results suggest that defining such

---

<sup>36</sup>Targets at law firms vary between 1700 and 2300 billable hours, but it is not clear whether this might be law firms trying to screen applicants [http://www.law.yale.edu/studentlife/cdoadvice\\_truthaboutthebillablehour.html](http://www.law.yale.edu/studentlife/cdoadvice_truthaboutthebillablehour.html)

a minimum standard would be detrimental for the buyer, since the seller will be strongly influenced by the defined standard, taking it as a performance norm. The doctrine of implied warranties therefore affords buyers with minimum protection without forcing them to stipulate a counterproductive low quality threshold.

An important recent example of setting targets in international law is the Paris Agreement dealing with greenhouse gas emissions mitigation, signed in 2016. Our study does not directly apply to the structure of the agreement as it does not provide for a schedule of deductions and bonuses for falling below or exceeding the stipulated targets. However, the flavor of our study suggests that such targets should be set at an ambitious but realistic level. The agreed goal of holding the increase in the global average temperature to well below 2°C above pre-industrial levels and to pursue efforts to limit the temperature increase to 1.5°C target above pre-industrial levels might well reflect this consideration. This is because the 1.5°C target is generally considered to be desirable but overly ambitious.<sup>37</sup>

Finally, our finding that setting low thresholds leads to low variance of effort levels suggests that regulators or principals should set low effort thresholds when they are concerned about consistency and uniformity of effort across agents. We would expect this to be true, for example, in establishing safety regulations. When regulating nuclear safety the government might care less about maximizing the average level of effort and more about minimizing the variance of effort. A company might care about uniformity of performance to facilitate coordination in a complex, high-stakes production process and might therefore set easy-to-reach goals. Similarly, the stereotypical observation that civil servants often face less ambitious performance targets than employees in the private sector might not be the result of rent-seeking in a sector isolated from competition but might in fact serve the Weberian notion that considers uniformity and reliability in the civil service intrinsically valuable (Weber, 1949, p. 214).

---

<sup>37</sup>The notion that keeping global increase in temperature below 2°C is ambitious but plausible, while limiting global temperature increase to 1.5°C is most likely infeasible, is supported by the International Energy Agency's (IEA) assessment of the Paris Climate Accords. The report states on page 5: "The 2°C pathway is very tough: the road to 1.5°C goes through uncharted territory" (p. 5, 2016 World Energy Outlook).

## References

- ABELER, J.; FALK, A.; GOETTE, L., AND D. HUFFMAN (2011): “Reference Points and Effort Provision,” *American Economic Review*, 101, 470–492.
- ALTMANN, S., AND A. FALK (2009): “The Impact of Cooperation Defaults on Voluntary Contributions to Public Goods,” *mimeo*.
- BROOKS, R. W.; STREMITZER, A., AND S. TONTRUP (2011): “Contracts - Why Loss Framing Increases Effort,” *Journal of Institutional and Theoretical Economics*, 168, 62–82.
- CORBIN, A. I. (1950): *Corbin on Contracts, A Comprehensive Treatise on the Working Rules of Contract Law*. St. Paul: West Publishing Co.
- EGGERS, W. D. (1997): *How-to Guide No. 17, May 1997*. Reason Foundation (Mar. 7, 2013, 9:45 PM), <http://reason.org/files/e0f57a04efae97a83b33c7520dc65dc8.pdf>.
- FALK, A., AND M. KOSFELD (2006): “The hidden costs of control,” *The American economic review*, 96(5), 1611–1630.
- FEHR, E.; KLEIN, A., AND K. M. SCHMIDT (2007): “Fairness and Contract Design,” *Econometrica*, Vol. 75, No. 1, 121–54.
- FREDERICK, S. (2005): “Cognitive Reflection and Decision Making,” *Journal of Economic Perspectives*, 19(4), 25–42.
- FRYER, R. G. J., S. D. LEVITT, J. LIST, AND S. SADOFF (2012): “Enhancing the Efficacy of Teacher Incentives Through Loss Aversion: A Field Experiment,” *NBER Working Paper No. 18237*.
- GOETTE, L; HUFFMAN, D., AND E. FEHR (2004): “Loss Aversion and Labor Supply,” *Journal of the European Economic Association*, 2(2-3), 216–28.
- GOOD, P. I. (2006): *Resampling Methods: A Practical Guide to Data Analysis*. Birkhauser, Boston, 3rd edn.
- HOLMSTROM, B., AND P. MILGROM (1991): “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design,” *Journal of Law, Economics, and Organization*, 7 Special Issue, 24–52.
- HOPPE, E. I., AND P. W. SCHMITZ (2011): “Can Contracts Solve the Hold-Up Problem? Experimental Evidence,” *Games and Economic Behavior*, 73(1), 186–199.
- HOSSAIN T.; LIST, J. (2012): “The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations,” *Management Science*, vol. 58 no. 12, 2151–2167.
- IMAS, A., S. SADOFF, AND A. SAMEK (2015): “Do People Anticipate Loss Aversion?,” *CESifo Working Paper No. 5277*.

- KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decisions under Risk," *Econometrica*, March 1979, 47(2), 263–91.
- KAUR, S.; KREMER, M., AND S. MULLAINATHAN (2010): "Self-Control and the Development of Work Arrangements," *American Economic Review*, 100(2), 624–28.
- KOSZEGI, B., AND M. RABIN (2006): "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121 (4), 1133–1165.
- LAURY, S. K. (2006): "Pay One or Pay All: Random Selection of One Choice for Payment," *Andrew Young School of Policy Studies Research Paper Series*, No. 06-13.
- LAURY, S. K., AND C. A. HOLT (2008): "Payoff Scale Effects and Risk Preference Under Real and Hypothetical Conditions," in *Handbook of Experimental Economics*, ed. by C. A. Plott, and V. Smith, pp. 1047–53.
- LAZEAR, E. P., U. MALMENDIER, AND R. A. WEBER (2012): "Sorting in Experiments with Application to Social Preferences," *American Economic Journal: Applied Economics*, 4(1), 136–63.
- MUSCH, J.; REIPS, U.-D. (2000): "A Brief History of Web Experimenting," in *Psychological Experiments on the Internet*, ed. by M. Birnbaum, chap. 3, pp. 61–88. Verlag Academic Press.
- VOLTAIRE, F.-M. (1770): *La Begueule: Conte Morale (The Prude Woman)*. Nabu Press, May 2010.
- WEBER, M. (1946): *From Max Weber: Essays in Sociology*. Oxford University Press, New York.

## A Appendix A: Tables

Table 5: Mean Effort Levels Compared Across Treatments (Including Rejections)

contract type	threshold	mean [Std. Dev.]	contract type			
			plain-vanilla	low-bar	stretching	extreme-effort
plain-vanilla	—	8.4 (N=62) [7.7]	—	$Z = 2.4^{**}$ $p = 0.02$	$Z = -2.1^{**}$ $p = 0.04$	$Z = 3.1^{***}$ $p < 0.01$
low-bar	5	5.0 (N=62) [4.3]		—	$Z = -4.1^{***}$ $p < 0.01$	$Z = 1.8^*$ $p = 0.08$
stretching	15	11.9 (N=65) [9.2]			—	$Z = 4.2^{***}$ $p < 0.01$
extreme-effort	50	5.3 (N=62) [7.7]				—

Table 6: Differences in Effort Dispersion between Treatments (P-Values)

	Diff. in Std. Dev	p-values		
		Levene Test <sup>a</sup>	Brown-Forsythe Test <sup>b</sup>	Permutation Test <sup>c</sup>
low-bar vs. plain-vanilla	-3.4	.000	.000	.003
low-bar vs. stretching	-4.2	.000	.000	.000
low-bar vs. extreme-effort	-4.6	.000	.003	.004
plain-vanilla vs. stretching	-.8	.417	.292	.376
plain-vanilla vs. extreme-effort	-1.2	.860	.994	.669
stretching vs. extreme-effort	-.4	.628	.373	.599

a) the Levene test compares average absolute deviations from the mean. b) the Brown-Forsythe test compares average absolute deviations from the median. Both tests assume that the test statistic follows an F-distribution. c) the permutation test is a non-parametric test comparing absolute average deviations from the mean based on resampling methods.

Table 7: Cognitive Reflection and Matching Threshold

	threshold contract type		
	low-bar	stretching	extreme
High CRT			
match	4 (16%)	0	0
exceed/undercut	21 (84%)	17	15
Low CRT			
match	12 (29%)	13 (27%)	0
exceed/undercut	29 (71%)	36 (73%)	24
Fischer exact	p=0.129	p<0.01***	-

We coded as High-CRT all subjects who correctly answered 2 or 3 questions on the CRT test.

Table 8: Robustness Check: Cognitive Reflection and Matching Threshold (Recoded)

		threshold contract type		
		low-bar	stretching	extreme
High CRT				
	match	9 (20%)	5 (13%)	0
	exceed/undercut	36 (80%)	35 (87%)	24
Low CRT				
	match	7 (33%)	8 (31%)	0
	exceed/undercut	14 (77%)	18 (69%)	15
Fisher exact		p=0.09	p=0.02**	

We coded as High-CRT all subjects who correctly answered at least one question on the CRT test.

## Appendix B: Permutation Test

Let  $\theta_C$  be the estimator of the average deviation of observations  $\theta_i^c$  from the estimated sample mean  $\mu_C$ :<sup>1</sup>

$$\begin{aligned}\theta_C &= \frac{1}{n_c} \sum_{i=0}^{n_c} |\mu_C - \theta_i^c| \\ \mu_C &= \frac{1}{n_c} \sum_{i=0}^{n_c} \theta_i^c\end{aligned}$$

where  $c$  is a label representing either of two treatment groups,  $a$  and  $b$ . We observed that  $\theta_a - \theta_b = x$ . We wanted to test the null hypothesis that the two treatment groups do not differ with respect to outcome  $\theta_c$ . In other words, we wanted to determine the probability with which we would observe this difference independent of treatment assignment. We therefore randomly reassigned the  $n_a + n_b$  observations from the two treatment groups into two new synthetic treatment groups with sizes  $n_a$  and  $n_b$  and calculated the relevant statistic  $\theta_a - \theta_b$ . With a sufficiently small sample size, we would have been able to account for all permutations, generating the full probability distribution of  $\theta_a - \theta_b$  under the assumption that treatment assignment is purely random. As the sample size was too big we chose to sample a sufficiently large subset of the set of possible permutations. To achieve this we repeated the resampling procedure 10,000 times. We could then determine the probability that  $\theta_a - \theta_b \geq x$  independent of treatment assignment. If this probability turned out sufficiently small we could reject the null hypothesis.

For example, the observed difference in average deviations from the mean was  $-4.02$  between the low-bar (*LB*) and the plain vanilla (*PV*) treatment,  $\theta_{LB} - \theta_{PV} = -4.02$ . We then took the 104 observations, 50 from the low-bar treatment and 54 from the plain vanilla treatment and randomly reassigned them to two new synthetic samples with sizes 50 and 54. We then calculated  $\theta_a - \theta_b$ . We repeated the procedure 10,000 times to generate a probability distribution of  $\theta_{LB} - \theta_{PV}$  and determined the probability that  $\theta_a - \theta_b \geq -4.02$ . The simulated p-value (two-tailed) turned out to be 0.003.

---

<sup>1</sup>We use the average deviation from the mean rather than the standard deviation as this is the statistic used by the Levene test.

## **Appendix C: Instructions/Screenshots**

(English Translation from the German Original.)

## Preliminaries

Thank you very much for participating in our experiment! Please notice the time restrictions that are imposed on you when fulfilling the experimental tasks. If you do not meet the restrictions, the experiment will be automatically terminated, though you will still be paid according to your performance up until that point in the experiment

As announced in the invitation e-mail you received, every 5th participant will receive a payment. Which of the participants gets paid will be determined randomly. All additional instructions will be provided in the course of completing the experiment.



Remaining Time:  
01 min 54 sec

- 1.
- 2.

[https://mpisurvey.econ.mpg.de/limesurvey/adminadmin.php?action= ...](https://mpisurvey.econ.mpg.de/limesurvey/adminadmin.php?action=...)

## The Contract

You can form a legally valid contract with the experimenter to perform a task. Please note that the contract you can form will be legally binding under §241 BGB. Your task is to correctly count how often a particular digit is contained in a window displayed on your computer screen. Your earnings are specified in the following paragraph of your contract:

"Agreed upon is the correct counting of 15 screens for a fixed payment of €15. For each screen you count beyond the 15 you will receive a bonus payment of €1. If you count less than the agreed upon 15 screens, you breach the contract. In this case, you will owe a penalty of €1 for each screen short of the 15 agreed upon screens."

## Entering the Contract

After reading the instructions you will be asked whether you agree to the conditions presented and want to enter into the contract. You can choose between two buttons that will be displayed: one for "Yes" if you want to accept the contract; and one for "No" if you want to reject the contract. If you click on "No" your session will be terminated. If you enter into the contract it is binding by law and the experiment will start immediately.



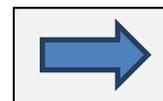
## Your Task

On the screen you will see displayed digits ranging from "1" to "9". You will be asked to count all the "1" digits in the table. Any entry falling within a range of -2 to +2 from the correct number will be considered correct.

If you indicate a correct number, you will be asked whether you want to continue with the task. If you click on "Yes", a new screen with the next table of digits will be shown, until you choose to stop completing the task by clicking on "No". If you indicate a wrong number (outside the acceptable margin of error) you will be asked whether you want to continue or stop the task. If you decide to continue, the same table will be displayed again. There is no limit to the number of times you can try counting that table. If you decide to stop the task, you will automatically be led to the last part of the experiment.

Below you see an example of a table. The contract requires that you count how often the number "1" occurs.

*I*	1	2	4	3	6	5	1	4	3	7	5	6	1	3	5	2	5	1	4	2
*II*	9	6	7	1	5	2	6	4	8	1	2	4	1	6	3	5	1	1	5	7
*III*	8	2	6	4	5	1	1	5	3	7	5	9	3	5	2	6	3	5	1	7
*IV*	3	6	4	3	6	7	1	9	0	6	0	7	0	4	7	5	6	8	2	4
*V*	3	9	4	7	1	4	1	5	1	3	2	9	9	7	5	3	6	7	6	9
*VI*	3	1	4	2	5	3	1	6	3	7	1	4	2	6	3	5	4	2	7	5
*VII*	9	4	6	3	8	5	8	1	3	1	5	4	7	3	5	2	6	4	8	1
*VIII*	4	2	5	3	7	9	5	7	1	4	2	6	4	8	9	2	3	4	2	1
*IX*	5	4	7	8	5	6	1	4	3	3	7	8	4	0	1	4	2	6	4	8
*X*	1	5	7	3	9	6	8	3	5	2	6	1	5	1	7	2	5	3	1	6



## Contract Formation

You can now enter into the contract with the terms described above. Please note that this contract is legally binding under § 241 BGB. If you want to enter into the contract with the mentioned terms, then please check the “Yes” box.

If you want to reject the contract please mark the “No” box. Your session will be terminated if you do not accept the offered contract.

Yes

No

## Table #1

Please count how often the digit "1" is contained in the table below. Please write the correct number in the small box below. You can only move on to the next screen if you enter a count that is considered

*I*	5	3	5	2	1	4	2	6	4	7	3	5	1	7	5	8	9	3	9	7
*II*	2	4	3	6	1	3	5	0	1	7	8	5	0	2	6	0	9	9	2	2
*III*	2	7	1	9	0	6	9	3	6	5	3	8	2	4	2	6	4	8	9	1
*IV*	3	4	5	2	7	1	5	2	9	5	0	7	8	8	4	6	2	2	1	0
*V*	2	0	6	8	7	7	2	5	1	7	3	0	5	3	7	1	9	0	1	6
*VI*	5	3	1	6	4	8	1	5	2	9	8	0	5	0	3	9	4	5	2	7
*VII*	1	2	6	1	7	8	3	0	4	7	2	8	3	8	6	1	7	3	8	4
*VIII*	1	5	2	8	5	0	1	6	3	8	3	9	0	0	4	7	2	8	2	7
*IX*	5	5	2	8	4	9	1	6	2	8	3	0	2	4	1	6	2	7	4	9
*X*	8	3	7	2	8	5	9	3	7	1	8	3	7	5	0	1	5	3	8	1

*You can only enter one number in the box. Please confirm your entry with the enter button!*

## Repeating the Task

Your entry was incorrect. Please decide whether you want to recount the same table or not. If you want to repeat the screen, please click on the “Yes” button. If you do not want to continue the task, please click the “No” button.

Yes

No

Remaining Time:  
01 min 23 sec

## Continuing the Task

Your entry was correct! Please decide, whether you want to continue fulfilling the contract or not. If you want to continue, please click on the “Yes” button; if you want to stop, please click on the “No” button.

Yes

No

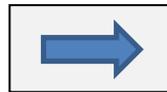
Remaining Time  
01 min 01 sec

## Participation in Lotteries

In the next part we ask you to decide whether you want to participate in two lotteries. You can increase your earnings by participating, but you can also lose some portion of whatever you may have earned in the earlier parts of this experiment. Note: Your decisions to participate in the two lotteries presented below are independent from each other. Therefore you can decline to participate in one lottery, but take part in the other; you can also reject both lotteries, or participate in both.

The experimenter will determine the outcome of the lotteries by throwing a fair coin. Your earnings will be added to your account, while your losses will be subtracted.

Please click “Yes” on the next screen if you wish to participate in the lottery and “No” if you do not wish to participate.



Remaining Time:  
02 min 56 sec

## Lottery 1

Lottery 1 has the following payoff: You win €8 with a probability of  $\frac{1}{2}$  and lose €5 with a probability of  $\frac{1}{2}$ .

If you participate in this lottery, either €8 will be credited to your account or €5 will be deducted from your account. If you reject to participate in the lottery your earnings in this part of the experiment are €0, that is, you neither win an additional payoff, nor do you lose earlier earnings.

Please click on the “Yes” button if you want to participate in the lottery and the “No” button if you do not want to participate in Lottery 1.

Yes

No

Remaining Time:  
0 min 34 sec

## Lottery 2

Lottery 2 has the same payoff as Lottery 1: You win €8 with a probability of  $\frac{1}{2}$  and lose €5 with a probability of  $\frac{1}{2}$ . However, the lottery is repeated six times, that is, there are six independent payoffs which will be consolidated in one overall payoff by adding and subtracting.

**Here is an example:** Assume you win €8 three times and lose €5 three times. The overall payoff for Lottery 2 would be €9  $[(3 \times 8) - (3 \times 5) = 9]$ .

If you reject to participate in Lottery 2 your earnings in this part of the experiment are €0, that is, you neither win an additional payoff, nor do you lose earlier earnings.

Please click on the “Yes” button if you want to participate in the lottery and the “No” button if you do not want to participate in Lottery 1.

Yes

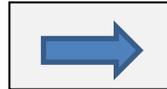
No

Remaining Time:  
01 min 09 sec

## The Quiz

On the following screen we will present you with a quiz with three questions. Note that you have no more than 90 seconds to answer all three questions. Once time has run out, you cannot answer any more questions. You can answer the questions in any order you like. You can also leave questions out. At the end of the 90 seconds, you will be paid €2 for each correct answer.

In order to begin the experiment, please click on the “Arrow” below. After 100 seconds, the quiz will automatically begin.



Remaining Time:  
00 min 03 sec

## The Questions

1. A bat and a ball together cost €1.10. The bat costs €1 more than the ball. How many cents does the ball cost?

2. If it takes 5 machines 5 minutes to make 5 bats, how long would it take 100 machines to make 100 bats?

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

Remaining Time:  
00 min 03 sec.

# Instructions for the Random Treatment

[https://mpisurvey.econ.mpg.de/limesurvey/admin/admin.php?action= ...](https://mpisurvey.econ.mpg.de/limesurvey/admin/admin.php?action=...)

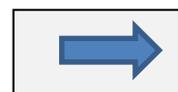
## The Contract

You can form a legally valid contract with the experimenter to perform a task. Please note that the contract you can form will be legally binding under §2: BGB. Your task is to correctly count how often a particular digit is contained in a window displayed on your computer screen. Your earnings are specified in the following paragraph of your contract:

"Agreed upon is the correct counting of a particular number of screens for a fixed payment. For each screen you count beyond this number you will receive a bonus payment of €1. If you count less than the agreed upon number of screens, you breach the contract. In this case, you will owe a penalty of €1 for each screen you count less than the agreed upon number of screens."

## The Contract Offer

The contract offer does not yet stipulate the number of screens you are obliged to count. A random device will determine this number now. Once you click on the button below the number will be drawn and will be included into the contract. Then you can decide whether you want to close the contract with this content or not.



## Random Draw of the Contract Offer

The random device has selected the following number of screens the contract obliges you to count:

5

This number will be stipulated in the contract offer. Your earnings are specified in the following paragraph of your contract:

"Agreed upon is the correct counting of 15 screens for a fixed payment of €15. For each screen you count beyond the 15 you will receive a bonus payment of €1. If you count less than the agreed upon 15 screens, you breach the contract. In this case, you will owe a penalty of €1 for each screen short of the 15 agreed upon screens."

## Entering the Contract

After reading the instructions you will be asked whether you agree to the conditions presented and want to enter into the contract. You can choose between two buttons that will be displayed: one for "Yes" if you want to accept the contract; and one for "No" if you want to reject the contract. If you click on "No" your session will be terminated. If you enter into the contract it is binding by law and the experiment will start immediately.

