Bias-Self-Management and the Role of Autonomy in Motivating Contract Performance

Stephan Tontrup and Christopher Jon Sprigman

Abstract

Loss-framed contracts are designed to spur productivity but they can also undermine autonomy and depress performance. Instead of imposing loss-framing our experiment provides workers with the autonomy to choose among a variety of loss and gained-framed contracts. The choice enables workers to use their own loss aversion as a commitment device. To improve their performance our subjects employ a sophisticated self-commitment strategy: they nudge their future self to work harder by entering into the loss-framed contract they expect will optimize their performance without failing the contract and suffering a loss. We call this phenomenon "Bias-Self-Management (BSM)," and we assess its utility as a tool for improving outcomes across a variety of policy settings such as retirement savings plans and labor contracts. BSM may also have implications for structuring legal commands as rules or standards.

In the second part of the paper, we isolate an independent positive effect autonomy has on performance: this autonomy premium fosters productivity independently from preference matching (choice allowing workers to better adjust work tasks to their preferences and abilities). The distinction has important implications: while firms can learn to design performance norms that encourage workers' productivity, the autonomy premium elevates performance only when workers are in fact provided with choice and experience self-determination. The economic value of a behavioral contract design that preserves autonomy and provides loss framing as a commitment device was significant: using monetary incentives alone, we needed to pay the subjects more than twice as much for them to achieve approximately the same productivity.

I. Introduction

Social scientists have systematically documented the many ways in which human behavior differs from a strict model of rational choice. First, people are only boundedly rational: they make mistakes in judgment for example weighing losses more than gains in decisions under risk. Second, people have bounded willpower: they have trouble following through on rational plans to complete their work, or to save for their retirement (Rabin 1998; Mullainathan and Thaler 2001).

By developing more empirically grounded models of human behavior, behavioral economics has helped policymakers to identify when policy objectives can be frustrated by individuals' cognitive biases and bounded willpower. Behavioral economics offers strategies for intervention that typically attempt to prevent bounded rationality from becoming consequential or that aim at debiasing decision makers (Jolls and Sunstein, 2006). More recently, a third type of policy intervention – nudging – has been added to the toolkit (Sunstein and Taylor 2008). Nudging changes decision environments and choice architecture to trigger heuristics, automatic processes and decision-making routines that are expected to

lead to (social) welfare. These policy strategies typically assume that the decisionmakers are either unaware of their bias or at least will not overcome it without a supporting intervention.

This Article follows a complementary approach. We assume that parties can be self-aware of their biases and that they can manage and even employ their own biases to their advantage. This approach, to which we refer as *Bias-Self-Management* (BSM), provides an alternative perspective on how to structure policies in response to bounded rationality. A response that enlists BSM provides individuals with some element of decision-making infrastructure that allows those who are aware of their biases to address, or even to harness their biases, to make better choices. As we shall see BSM can have considerable advantages: it allows individuals to use their private information and act upon their preferences, while imposing costs only on those who are in fact biased. Further, BSM preserves autonomy and thereby fosters intrinsic motivation. In this Article, we analyze a BSM-based behavioral contract design – one that provides workers with choice among different loss-framed contracts. Workers with time-inconsistent preferences can use the framing to harness their own loss aversion as a commitment device.

A. Benefits and Harms of Loss-Framing

Contracting parties use a variety of instruments to influence one another's performance. Often, they provide for bonuses if goods and services are delivered on time and in specified quality, or they stipulate penalties should work not meet expectations. In recent years, however, public and private contractors have attempted to induce desired performance behavior using means other than plain incentives. Building on empirical work in behavioral economics, contract drafters try to exploit common biases like loss aversion in order to extract more effort from their workers or partners (Madrian 2014).

For most people losses loom larger than gains. Contract drafters can trigger loss aversion by specifying performance thresholds – i.e., by offering an ex ante payment conditioned on productivity meeting a certain level. If a worker's performance falls short, she faces a schedule of penalties. The prior evidence regarding work productivity under these sorts of contracts largely bears out the predictions of Prospect Theory: individuals invest more effort under loss-framed contracts, because they want to avoid the loss of their initial entitlement (Hossain and List 2012; Fryer et al. 2012).

Some business contracts have pushed loss-framing to an extreme, where the counterparty is not expected to ever reach the threshold. For example, a consortium of engineering firms designed a deeply loss-framed contract when engaging construction firms to build the Øresund bridge between the cities of Copenhagen and Malmo (Alterbaum et al. 2011). An important element of the project was a suspension design that would efficiently reduce vibrations. The consortium specified an efficiency level for the suspension they knew could not be achieved with existing technology. While the contract terms promised a high payment conditioned on meeting this target, the consortium expected the contractor to fall short of the threshold and stipulated a schedule of penalties. The resulting deep loss frame was meant to push the contractor to develop the most efficient suspension it possibly could.

As in the Øresund case, loss-framing is typically employed to push the counterparty to invest more effort. Yet trying to override a partner's effort choice can be counterproductive if the contract drafter possesses information insufficient to stipulate effective performance thresholds. People value losses and gains relative to a reference point (Kahneman & Tversky

1980). The further distant perceived gains or losses are from that reference point the lower the value individuals place on them. Thus, when contract drafters establish a loss frame but set the threshold too high, workers derive less utility from the same performance. The loss frame reduces their effort instead of pushing it. A threshold also communicates what performance is expected, and how other workers may perform. Individuals tend to conform their effort choice to these social and prescriptive norms. Accordingly, a threshold set too low dampens performance, because it communicates a social norm of low productivity that workers conform to. Thus, loss framing will only enhance effort if thresholds are well adjusted to worker`s preferences and abilities (Brooks, Stremitzer and Tontrup 2017).

B. Self-Nudging and the Autonomy Premium

In this study we look for an alternative approach for behavioral contracting: one that does not assist parties in trying to exploit biases of their counterparty, but that aims to make both contractual parties better off. Instead of suppressing parties' autonomy by imposing loss-framed contracts, we aim to preserve contract autonomy and increase productivity by offering workers a choice between various loss- and gain-framed contracts. The most obvious benefit of this approach is access to private information. Parties typically possess private knowledge about their production function. A contract design that provides choices can reveal and benefit from this private information. The risk that an employer under-challenges or overstretches workers is minimized.

But why would workers prefer a loss-framed over a gain-framed contract? If a worker is loss averse she generally should prefer a gain-framed contract: it leads her to a higher net-outcome as gain-loss utility below the loss-frame's threshold is negative. Rational types should not choose a loss-framed contract either: they should be indifferent. Luft (1994) and others present survey results in line with this prediction, suggesting that workers resent being exposed to losses and therefore prefer gain-framed contracts. However, the empirical evidence regarding whether workers would deliberately enter a penalty contract, and, if so, why is contradictory. Quidt (2014) proposes that people prefer loss-framed contracts as they appear to pay better initially. And, in Imas et al. (2017) subjects reveal a willingness-to-pay for entering a loss-framed instead of a gain-framed contract, evidence the authors interpret either as a demand for self-commitment or, in line with Quidt, as a preference for receiving the reward for their work immediately as workers do under a loss-framed but not a gain-framed contract.

We want to show that present-biased workers value a loss frame because it provides them with a nudge: they deliberately choose penalty contracts to instrumentalize their own loss aversion as a commitment device to improve the performance of their future self. Preferences about work are often inconsistent over time: when engaged in a work task, individuals frequently prefer a lower-effort choice than both before and after performance (Augenblick et al. 2015, Cadena et al. 2011, Kaur et al. 2010, Sadoff et al. 2015). Individuals have been shown to be aware of this time inconsistency (Giné et al. 2012, Hsiaw 2013, 2015, Kaur et al. 2015, Royer et al. 2015, Schwartz et al. 2014). Choosing a loss-framed contract could therefore be a rational strategy of workers who anticipate that their preferences will be inconsistent over time.

We developed a novel experimental design that allows us to analyze subjects' motivations for choosing a loss-framed contract. We offer participants a menu of different

gain- and loss-framed contracts and elicit the personal performance expectations they have for each of the contracts. This study design reveals the strategy subjects use for their contract choice. We assume that workers are sophisticated about their loss aversion and present bias and make a behaviorally rational decision: we expect them to select, from the menu of loss-framed contracts offered to them, the threshold that nudges their future self into its optimal performance. Our data confirm this theory: our subjects anticipate that the loss frame and its size will strongly affect their productivity, and they enter the loss-framed contract they expect will optimize their performance and net outcome. We used a standard psychometric measure – a lottery game – to distinguish subjects manifesting more or less intense loss aversion. In support of our theory, we find that the types with a stronger bias can harness their own loss aversion more effectively for self-commitment: They increase their performance and payoff significantly more than participants who are only mildly loss averse or not at all.

Self-commitment to thresholds that match preferences and abilities is the first element that makes the autonomy-preserving contract design we implement effective. The second element is the experience of autonomy itself. We assume that autonomy enhances worker motivation and productivity independently of the performance gains that arise from the better match between tasks and preferences. The idea that outcome control and responsibility affect both motivation and effort has received longstanding attention in multiple disciplines. In economic philosophy, Sen (1985) and Nussbaum (2000) argue that control and choice are capacities valued by individuals beyond outcomes. Similarly, Frey (1997) provides evidence that people place inherent value on inner goods such as autonomy and try to produce and preserve them. In the field of organizational psychology, Deci and Ryan (2000) speak of self-determination and suggest that the basic desire to be in control of one's fate fuels intrinsic motivation for all sorts of work productivity.

To our knowledge our study is the first to separate two elementary effects of choice: (1) matching tasks well to preferences and abilities and (2) the experience of autonomy. Our study provides the first empirical evidence that both choice effects improve contract performance independently of each other. Earlier work has either not distinguished between these two effects or at least could not clearly separate them. For behavioral contracting, however, this distinction is crucial and has important implications. With today's technology firms can track their workers and measure their performance better and easier than ever before; over time, firms can learn to adjust performance norms such that they effectively encourage workers' productivity without granting them contractual choice and autonomy. But while effective preference matching depends on information that can be learned, the autonomy premium does not. The autonomy premium elevates performance only when workers are in fact provided with and experience self-determination and choice. So even if companies have the information to adjust work norms optimally, effort and productivity could still be improved when contract terms offer choice and autonomy.

In summary, our study suggests that offering loss-framed contracts as a commitment device is an effective strategy for behavioral contract design. Workers derive self-determination utility from being granted autonomy, and this raises their performance. They use their autonomy productively to nudge their future self towards a better performance by adjusting the loss frame to their preferences and abilities. The economic value of the behavioral contract design we propose is significant: to make subjects reach the same

performance under a linear gain-framed contract, we would have had to provide more than twice the monetary incentive.

III. Experimental Design

1. The Real Effort Task

In our experiment subjects can enter into a contract governing the performance of a real effort task. The task presents them with a table that contains 200 digits. Subjects are instructed to count how often a specified digit (for example "2" or "3") occurs in the table. An answer is considered correct within the range of $-2 \mid +2$ of the true value. Subjects must repeat the same task if their answer is incorrect. To prevent participants from trying to reach the solution by trial and error, they cannot make a new entry for 15 seconds following an incorrect input. As it takes participants around one minute to complete a single table, the buffer between entries makes guessing time-inefficient.[1] When subjects enter a correct answer, they can decide whether or not they want to continue the task with a new table.

We do not induce the agreement by instructing subjects to assume they have a contract. Participants are free to agree to or to reject the contract. However, subjects are aware that if they reject the contract offer, they will receive no payment.

2. The Treatments

a. Baseline and Double-Incentive Baseline

Our *Baseline* condition presents subjects with a *Linear* contract that does not stipulate a performance threshold. Subjects are paid €1 for each table they complete. In *Double-Incentive Baseline* subjects are paid €2 per table completed. The *Double-Incentive* treatment allows us to estimate the monetary value of offering subjects autonomy to employ loss framing as a commitment device, in comparison to a contract regime based on monetary incentives alone.

b. Single Threshold Contract Treatments

All other treatments set out a threshold – i.e., a specification of the number of tables that subjects are expected to complete. As we will explain in the theory section, the threshold is a pure framing device that does not affect subjects' actual payment. The contracts that include a threshold specify a penalty for any performance below the threshold and a bonus for any performance exceeding it. The underlying payoff scheme, however, is always linear and held constant across all treatments: subjects earn €1 per table, regardless of what threshold is specified in the contract.

We used the results of our *Baseline* condition to attune the thresholds of the loss-framed contract treatments. In *Baseline* participants completed an average of 10.4 tables. Based on this performance, we calibrated three basic contracts. The first, which we label the *Low Bar* contract, expresses a threshold of 5 tables. We set that threshold such that we would observe a significant effect if a quarter of the subjects would lower their average performance from the level we found in the *Baseline* condition to the threshold. The *Stretching* treatment offers a contract with a threshold of 15 tables. The *Stretching* contract should increase performance significantly above the *Baseline* treatment if a quarter of the subjects would lift

their effort up to meet the 15-table threshold. Finally, we offered subjects a contract that stipulated an *Extreme Effort* threshold of 50 tables that obliged subjects to invest five times the effort they had exerted in the *Baseline* treatment.

c. Choice Treatments

The next treatments provide subjects with a choice between the loss-framed and gain-framed contracts that we presented above.

α) Self-Commitment

Our *Self-Commitment* treatment analyzes whether subjects are capable of using loss-framed contracts as a commitment device when they are granted a choice of contracts. The treatment offers subjects choice among the three threshold contracts *Low Bar*, *Stretching* and *Extreme Effort*. In addition, subjects can select a *Linear* contract, which does not impose a threshold or a loss frame. Subjects must pay to enter a threshold contract: for the *Low-Bar* contract subjects must pay 20 cents from their earnings; for the *Stretching* contract 60 cents and for the *Extreme Effort* contract, €1. Entering the *Linear* contract in contrast is costless. By choosing a threshold contract, subjects reveal that they have a willingness-to-pay for committing to a particular level of performance.

In addition, we elicit subjects' performance expectations for each of the four offered contracts. This allows us to analyze the strategy subjects use for choosing a contract. To control for order effects, we randomize the order in which the four contracts are presented to participants when they indicate the number of tables they expect to complete under each contract. Subjects are incentivized to reveal their true expectations. If their actual performance meets their expectations for the contract they entered, they are paid an additional 10 cents for each table they completed. However, if they fall short of their expectations, they get no additional payment. For example, if they indicate an expectation of 30 tables, they are paid an additional 300 cents if they complete at least 30 tables; they receive no additional payment if they fall short. This does not change the performance payment of €1 per table.

Because the incentives we provide for the expectation measure may affect behavior, we apply them randomly. We instruct the participants that we will randomly select one in five subjects. The selected subjects are then randomly assigned one of the contracts and paid according to the expectations they indicated and their actual performance under the assigned contract. All other participants are permitted to choose between the four contracts and perform the task under the contract they prefer, but are not paid for the expectation measure. Participants learn whether they are selected or not after they indicate their expectations and before they begin working on the real effort task. This design ensures that all subjects have a monetary incentive to reveal their true performance expectation. On the other hand, the measure does not change the performance incentives of the subjects who are not selected. For our analysis we exclude from the sample all selected subjects who were randomly assigned a contract.

β) Autonomy

For isolating the autonomy premium's effect on contract performance we designed a treatment that we can compare directly with the single threshold treatments. This *Autonomy* treatment offers participants a choice among the three threshold contracts *Low Bar*, *Stretching*

and *Extreme Effort*. It allows us to analyze the effort subjects invest under the very same contract terms when they freely selected these terms, versus when offered only one particular contract without being given a choice.

γ) Preference Matching

We designed a control treatment, *Preference Matching*, to separate adjustment from autonomy effects. The treatment presents subjects with the three threshold contracts (*Low Bar*, *Stretching* and *Extreme Effort*). Participants indicate which contract of the three they prefer. The treatment provides them with an incentive to reveal their true contract preference: We instruct the participants that 1 in 5 of them will be offered the contract they indicate they prefer; if they are not selected, one of the threshold contracts (including the one they prefer) would be randomly picked and offered to them. Thus, if subjects indicate their true preferences, they are more likely to actually receive their preferred contract.

The treatment allows us first to reveal the effect of a better match of subjects' production functions with the contract's threshold. We can compare the performance of participants who prefer the same terms and who received the contract they preferred versus when they are assigned a contract they would not have chosen.

The comparison of the *Autonomy* and *Preference Matching* treatments enables us to isolate the autonomy premium; i.e. the independent effect of the experience of autonomy on performance. For this comparison, we consider only subjects who were randomly offered their preferred contract in *Preference Matching*, such that all subjects perform under the contract terms they favor; however only in *Autonomy* they can directly *choose* their preferred contract, while in *Preference Matching* the favored contract is assigned randomly.

Credibility of the random assignment in *Preference Matching* is important, in particular when subjects happen to receive the contract they prefer. If participants suspect they may have received the contract automatically upon indicating they have a preference for it, the differences between the *Preference Matching* and *Autonomy* treatments would be blurred. To ensure credibility, we prepared two sets of photographs picturing buildings in New York City. The entryways of each of the buildings are marked with a street number -- either 5, 15 or 50 -- corresponding to the three contract thresholds. The numbers are pixelated. Participants pick one of the photos and are assigned the contract threshold corresponding with the street number of the building they selected. We provide a link to Google Maps for each building subjects are presented with. This allows subjects to rule out a manipulation by checking that the buildings indeed have the street numbers we indicate.

δ. Random Choice

We implement a *Random Choice* treatment to identify self-determination utility as the driving force of the autonomy premium, as opposed to positive reciprocity. The treatment either provides subjects with the autonomy to choose one of the three threshold contracts, or it randomly offers them one of the three contracts without giving them a choice. Which of the two conditions is implemented is determined randomly. The random assignment prevents subjects from perceiving the grant of autonomy as an act of kindness in response to which they may wish to reciprocate. We use the same methodology as in the *Preference Matching* treatment to ensure the credibility of the random assignment.

ε. Open Choice

Finally, in the *Open Choice* treatment subjects can select and make the contract prescribe any threshold, within the range from 1 to 50 tables they prefer.

d. Loss Aversion Measure

We use an incentivized loss aversion measure (Goette et al. 2004) that gives subjects the opportunity to participate in two lotteries. The first lottery offers subjects a 0.5 chance of winning €8 and the same probability of losing €5. The second lottery presents the exact same payoffs, but the lottery is repeated six times and subjects are paid the average outcome of all six draws. Therefore, they face a lower risk to suffer an overall loss. An unbiased participant should accept both the first and the second lottery, as their expected outcome is positive. We classify participants into the following two categories: 1) The "loss averse" type rejects one or both lotteries; 2) the "rational" type participates in both lotteries.

III. Theory and Predictions

1. Loss and Norm Framing

Loss-framing can increase contract performance. Hossain and List (2012) for example present evidence from a field study with Chinese factory workers, who lift their productivity under loss-framed contracts. Fryer et al. (2013) vary the incentives for teachers, and find that loss-framed contracts increase teaching performance.

Loss framing exploits the fact that payments can be offered as gains or losses, at equivalent monetary costs. Prospect Theory assumes that individuals value utility relative to a reference point (Kahneman and Tversky 1979, 1992; Kőszegi and Rabin 2006, 2007). Outcomes below the reference point are experienced as losses and outcomes above as gains. To account for this reference dependence we assume that subjects derive independent gain-loss utility from their performance. While consumption utility remains constant, gain-loss utility is negative below the reference point and positive when productivity exceeds the reference point. Our treatments shift the subjects' reference point when they vary the performance threshold: For instance, in the *Low Bar* treatment the reference point is five and in the *Extreme Effort* it is 50 tables.[2]

The rationale for designing loss-framed contracts is that reference-dependence is asymmetric: i.e., for most people, losses loom larger than gains. Assume an individual either falls short of the reference point or exceeds the reference point by the same margin. The negative gain-loss utility from not reaching the reference point is larger than the positive benefit individuals derive from exceeding it. As a result loss-framing can increase the costs of falling short of the performance norm and can thus induce individuals to work harder. Thus, by framing payments as losses, firms may have a stronger influence on worker motivation and performance without spending more on financial incentives (Hossain and List 2012).

However, loss-framing can also depress productivity if thresholds are not well-adjusted (Brooks, Stremitzer, Tontrup, 2017). The more distant perceived losses are from people`s reference points, the less gain-loss utility they carry and this reduces the value individuals give them (Kahneman & Tversky, 1980). This in turn reduces the motivating effect of both gains and losses on effort choices. Thus, when contract drafters push thresholds too high, the loss frame will reduce, rather than spur, work effort. A threshold set too low also dampens performance, because workers conform to it by lowering their productivity (Brooks,

Stremitzer, Tontrup, 2017). Loss-framed contracts communicate prescriptive and social norms that frame the desirability of performance. The loss frame suggests that an effort choice below the threshold is insufficient but also that exceeding the norm is neither required nor expected. Therefore, and similar to loss-aversion, workers derive larger social utility when they approach the norm from below, and less once they exceed the norm. When social desirability of continuing performance is lower, once the threshold is reached, subjects are more likely to stop their performance compared to when no norm was specified. As a consequence, if the prescribed threshold is set below the subjects' optimal effort choice absent the performance norm, loss-framing will dampen productivity. A contract drafter's attempt to exploit loss aversion may thus backfire.

2. Loss-framing as a Commitment Device

We want to demonstrate that an alternative contract design that preserves workers` autonomy by offering them a choice of loss and gain-framed contracts can be more effective, because it (a) accesses private information (b) allows workers to derive independent utility from their autonomy and (c) eliminates the potential negative side-effects of norm-framing.

But why would workers deliberately choose a loss-framed contract? Standard economic theory has always assumed that workers would not prefer a loss-framed over a gain-framed contract, if they were given a choice. A strictly rational subject should be indifferent between loss and gain-framed contracts, while a loss averse subject should prefer a gain-frame; the loss-frame earns the subject a strictly lower net outcome, as gain-loss utility is negative below the threshold and increases subjects' net-outcome only once performance exceeds the threshold. In contrast, under a gain-framed contract the status quo is the reference point; gain-loss utility is strictly positive with the first unit of effort the subject invests, increasing the subject's net outcome from the start.

The same holds when a subject can decide between contracts that stipulate loss frames of different sizes: subjects should be indifferent or choose the lightest frame; the deeper the loss-frame, the less gain-loss utility and net-outcome the subject will earn.

We assume by contrast that workers can benefit from loss-framing and therefore, if provided with autonomy and contract choice, often will favor loss-framed contracts in their own self-interest. We suggest that sophisticated workers can use loss frames as a commitment device to counter their present bias and improve their performance. We base our theory on three assumptions: (1) people have a demand for self-commitment; (2) they are aware of their loss aversion and therefore anticipate that their choice of contract will affect their performance; and (3) they are capable of making a behaviorally rational choice and will select the contract under which they expect to optimize their net-outcome.

People have a demand for self-commitment because their preferences about work effort are often time-inconsistent. Often they are aware of this inconsistency: many studies show that individuals prefer a lower-effort choice while engaged in a work task than before and after production (see, for example, Kaur et al. 2010; Cadena et al. 2011). Temporal preference inconsistencies were also replicated in lab experiments even when effort choice and performance were separated by only a few minutes (McClure et al. 2007; Brown et al. 2009).

To understand how loss-framing can serve as commitment device, consider a subject who would like to earn €15 in our study, but expects her future self – i.e., herself when

actually completing our tedious real effort task -- to have a lower preference and to perform only for €10. In this situation, the subject can choose a loss-framed *Stretching* contract as a commitment device to make it costlier for her future self to fall short of the performance goal she prefers before starting work. As negative gain-loss utility looms larger, the future self will derive a higher marginal benefit from the effort it invests below the threshold. The future self's optimal effort and monetary payoff increase. Thus, when subjects expect their future self to invest less effort than they would prefer, choosing a loss-framed contract can be a sophisticated response. By contrast, if individuals expect to have consistent preferences at t_0 when choosing the contract and at t_1 when they will perform, they are (all else equal) better off with a gain-framed contract.

We can specify the conditions under which subjects in our study should have a demand or willingness-to-pay for a commitment device. First, the current self must anticipate that the felt production costs for the future self will be higher, once it performs the work task, and therefore that the future self will have a lower productivity preference than the current self (i.e., time-inconsistent preferences). Second, the negative gain-loss utility induced by the loss frame must sufficiently incentivize the future self to raise her effort. And finally, the loss frame must improve performance sufficiently that the additional consumption utility outweighs the costs the loss frame inflicts on the current self—i.e., the that gain-loss utility is negative below the threshold and the fee subjects must pay to enter a loss-framed instead of the *Linear* contract. If these conditions are met, a subject can increase her net outcome when she selects a loss-framed contract.

Notice that the optimal level of commitment is conditioned by the actual production function and preferences of the future self. Subjects do not want to overstretch their future self`s by choosing a threshold the future self`s will not meet, as they would experience more negative gain-loss and social disutility. If the consumption value added by producing under the loss frame does not make up these costs, then the loss frame effectively lowers the subjects` net outcome. We therefore presume participants will typically optimize their net outcome by choosing a threshold contract they expect their future self to fulfill. For example, suppose a participant would like to earn €15 and can choose between the *Linear* contract and all three loss-framed contracts. The future self's expected optimal effort under the *Linear* contract may be 12 tables. Then the subject will select a *Stretching* contract if she expects her future self to reach this threshold under a loss-frame. Otherwise, if the subject expects that the loss frame will not push optimal effort far enough to make her future-self fulfill the norm, she will choose the costless *Linear* contract.

Considered in this light, we understand the previous empirical evidence not to contradict our theory. The tasks that the previous literature implement are not focused on effort and thereby make self-commitment less effective or even useless. For example, Luft (1994), implements a surprise knowledge quiz and reports (in line with standard theory) that people resent being exposed to loss-framed contracts. Given the nature of the task used, this is unsurprising: self-commitment could not help subjects to improve their performance on a task that tests knowledge rather than effort. Consequently, subjects chose a gain-framed contract. Quidt (2014) implements a task that measures accuracy and penalizes errors. The task makes it difficult for subjects to predict whether or to what extent extra effort may improve their performance and therefore whether committing their future-self would pay off. Additionally, the task does not give subjects the control needed to avoid a loss. Even Imas et al's (2017)

task is not ideally suited to analyze self-commitment: the task lasts for only 90 seconds, and therefore time-inconsistent preferences may not be the main driving force behind their results (indeed, it is plausible that subjects might have chosen the loss-frame because doing so gave them a possibility to inspect the reward (a t-shirt) before investing effort into performance).

We test our theory in the *Self-Commitment* treatment, which offers subjects a choice between the three loss-framed contracts and the *Linear* contract. If subjects choose one of the loss-framed contracts, they must pay a fee. The second cost associated with selecting a loss-framed contract is implicit: Subjects earn a lower net outcome for investing equal effort. The costs ensure that subjects do not casually choose a loss-framed contract when they are in fact indifferent between the four contracts. The design holds consumption utility from an equal performance constant across all contracts. Thereby it ensures that individuals will only choose a loss-framed contract when they expect to benefit from committing to a threshold. On the other hand, subjects who select the *Linear* contract may either expect to have time-consistent preferences and therefore have no demand for self-commitment or they may not be loss averse enough to capitalize on their bias.

We can now derive hypotheses concerning the strategy we predict subjects will employ in making their contract choice. We elicit the performance expectations subjects have for all four contract types. This data should reveal the motivation subjects have for choosing their preferred contract. We hypothesize that the participants will pick the contract they expect will optimize their net outcome; for many loss-averse subjects this should be a loss-framed contract, supposed to serve them as a commitment device. Accordingly, a significant number of subjects in our *Self-Commitment* treatment should choose one of the loss-framed contracts over the costless *Linear* agreement ($H_{1.1}$). As subjects would experience negative social and gain-loss utility when they fail the contract and realize a loss, we hypothesize subjects will choose the contract that leads their future self to the best performance, conditioned on not falling short of the threshold. We test our prediction against other possible strategies of contract choice, for instance a general preference for gain-framed contracts ($H_{1.2}$). Only loss averse subjects can benefit from loss-framing, because loss-framing can push up optimal effort only for these subjects. Therefore, the difference in the performance expectation loss-averse subjects have for working under their optimal loss-framed contract versus the *Linear* contract should be higher in comparison to types that our psychometric measure classifies as rational (or less loss averse) ($H_{1.3}$). Finally, we assume that subjects will comply with their self-commitment to avoid negative gain-loss utility and social disutility, allowing them to significantly improve their performance. As a result, loss averse subjects should outperform rational types ($H_{1.4}$).

2. Disentangling the Effect of Contract Choice and Autonomy on Performance

We now want to disentangle the separate mechanisms that we expect make a self-committing contract design that grants contract choice and autonomy effective: (a) access to private information, (b) the experience of autonomy, and (c) the elimination of norm-framing.

a. Preference Matching

All things equal, we expect an individual to perform better when her preferences and abilities are well-matched to a task.

Subjects' optimal effort absent the loss frame is given by the difference of marginal effort costs and the marginal utility they derive from performance. Whether a loss frame will increase or dampen subjects' optimal effort depends on the distance between subjects' optimal effort choice and the threshold. Choosing a threshold close to their optimal effort can elevate their future self's performance,[3] because gain-loss utility increases close to the reference point, providing a stronger performance incentive. By contrast when subjects are randomly assigned one of the three contracts, the threshold may fall well short of or may be set well above their optimal effort; in the first case, norm-framing may cause subjects to lower their effort in an attempt to conform to the norm expressed by the offered threshold contract. In the second case diminishing sensitivity dampens their effort as the reference point is too far distant.

We can now derive two experimental predictions for the *Preference Matching* treatment. The performance of subjects who received the contract they preferred should exceed the effort level of the participants who have the same contract preference but whose preferences were not matched ($H_{2.1}$). However, even when employers impose contracts on workers without considering their preferences some parties will receive their preferred contract while others will not. To demonstrate the effect against this realistic benchmark we compare the subjects of the *Preference Matching* treatment who received their preferred contract with the pooled single threshold *Low Bar*, *Stretching* and *Extreme Effort* treatments which include both subjects who received their preferred contract and subjects who did not. We expect that subjects in *Preference Matching* should perform better, given the greater instance of preference adjustment in that condition ($H_{2.2}$).

4. The Autonomy Premium

Many commentators have speculated that autonomy may have a positive impact on productivity. Existing theories and the empirical evidence provided to date mingle the effect of autonomy with the benefits of preference matching and the revelation of private information. Ryan, Deci and Frey have shown in experimental and field work that individuals have a higher *intrinsic* motivation and persistence in the performance of a task when they are granted autonomy in choosing it. However, their experimental manipulations affect not only their subjects' experience of autonomy, but they also keep the subjects from choosing according to their true preferences. For example, Moller et al. (2006) permit subjects to freely choose their preferred effort task in one condition, while in the other they impose social pressure on subjects to pick one particular task over the others. They measure higher work effort in the treatment condition, where subjects can choose free of social pressure, but they cannot identify what has caused this effect: that subjects' select the task that best matches their preferences or that they embrace autonomy. We want to provide evidence for an independent autonomy premium, caused by self-determination, not by a good match of preferences, abilities and work task.

Our theory assumes that for each unit of effort that subjects invest they derive separate self-determination utility that will linearly lift their productivity. This linear bonus on top of consumption utility should be independent of the threshold contract they choose, and it may both lift workers' initial effort preference and push their performance beyond the threshold they chose. To isolate this autonomy premium, we compare productivity in the *Autonomy*

treatment with productivity of subjects in the *Preference Matching* treatment who received their preferred contract.

We want to test this concept of an intrinsic autonomy premium against alternative theories. One group of theories suggests that making choices may cause commitment effects that raise performance. A first mechanism to drive these commitment effects is cognitive dissonance. People generally aim to avoid cognitive dissonance (Festinger, 1959). If subjects choose a contract threshold and then fail to comply with it, they may experience dissonance between their choice and their behavior. The dissonance would suggest either that they made a bad choice, or that they failed to meet a performance norm they validated as reasonable with their choice (see Leotti et al. 2010). As a result, subjects seeking to avoid dissonance may raise their actual performance to match the threshold contract they have chosen.

Guilt aversion is another mechanism that may commit workers to comply with the contract threshold they have chosen (Ederer and Stremitzer 2017). Subjects may believe that by selecting a particular contract threshold they have induced a performance expectation in the offeror. The subject anticipates experiencing guilt over disappointing this expectation and so the subject matches the threshold. Both mechanisms have no influence on performance above the threshold.

The potential effect of regret costs is different. Choice can trigger regret over the forgone gains of an alternative decision path not taken. For example, a worker chooses one job, rather than another. In trying to avoid feeling regret over not having taken the alternative job, the worker invests more effort into the option she chose to confirm its superiority over the alternative she dismissed (Loomes and Sugden 1982; Sugden 1985; Bartling et al. 2014). Regret costs are differential: the decision-maker compares the regret she feels over the benefits lost, because she did not take the alternative job, with the hypothetical regret she assumes she may feel would she not have taken her current job, but the alternative one. As a result, regret costs are positive as long as the option actually taken appears in better light than the alternative not taken, and, vice versa. If the alternative looks better, regret costs are negative. As a consequence, regret costs create the incentive to work harder, once a choice is made and an alternative dismissed. In the context of our study regret costs can both push and dampen effort. They push effort as long as the selected contract is not optimal, because productivity under a lower threshold contract not taken would have been higher. For example, a subject has chosen a *Stretching* contract, then regret costs will push her performance, until the *Stretching* contract is in fact a better choice than a *Low Bar* contract. But regret costs dampen effort when carrying on performance would render the selected *Stretching* contract inferior because the subject would have reached a higher performance under an *Extreme Effort* contract. Whenever the selected threshold is optimal for the achieved performance, regret costs do not affect productivity in any direction, neither do they push nor dampen effort. (For more detail on this point, see Appendix).

Finally, if workers value being granted autonomy and contract choice, they may also respond with positive reciprocity towards the offeror and lift their performance (see for example Hannan et al. 2005).

We will now derive the hypothesis for our theory and point out why those hypotheses are inconsistent with the predictions for the alternative theories and mechanisms: First our self-determination theory predicts that subjects in the *Autonomy* treatment will be significantly more productive compared to the participants in *Preference Matching* who did

not choose but are randomly assigned the contract they prefer ($H_{3.1}$). If individuals derive utility from self-determination the autonomy premium should linearly increase the subjects' performance irrespective of what contract the subjects effectively choose. So, we expect to find a treatment effect in comparison to *Preference Matching* for each of the contracts *Low Bar*, *Stretching* and *Extreme Effort* ($H_{3.2.}$). This evidence would be inconsistent with regret costs driving the autonomy premium: subjects who choose a *Low-Bar* contract should not raise their effort, as they cannot reduce regret costs by lifting their performance.

Furthermore, self-determination utility should raise subjects' effort both above and below the threshold. Consider two subjects performing under a randomly assigned *Stretching* contract; the first subject complies with the threshold, completing 15 tables, while the second falls short, reaching 10 tables. If the two subjects would now be given contract choice and selected the *Stretching* contract, the autonomy premium should increase both their performance below and above the threshold. Thus, we can derive the prediction that as self-determination utility should push the performance of some subjects above the threshold, we should see comparatively more subjects exceed the threshold in the *Autonomy* treatment versus in *Preference Matching*, also the number of tables that subjects complete above the thresholds should increase in *Autonomy* as compared with *Preference Matching* ($H_{3.3.}$). By contrast both commitment effects and regret costs should make subjects less likely to fall short of the thresholds in the *Autonomy* treatment than in *Preference Matching,* but they should not increase their performance above the threshold.

Finally, our self-determination account suggests that the autonomy premium should remain substantial when we mute positive reciprocity in the *Random Choice* treatment ($H_{3.4}$).

Identifying the mechanism by which contract choice may improve performance is crucial for determining the comparative advantage the autonomy premium and a behavioral contract design that preserves it may provide. If workers derive separate self-determination utility from their performance as we suggest, then employers can access this source of motivation and effort only by providing actual autonomy and choice. This autonomy premium would be robust: it would increase performance independently of whether thresholds are set optimally and of the contract choice workers make. Thus, it does not depend on expertise on either side of the contract, the principal's as well as the agent's side. Neither would the reason why choice is offered – whether as a benefit to workers, or as a means of saving transaction costs, or as a prod to improve performance -- affect the premium (Falk et al. 2006). By contrast if the effect of choice is limited to the improvements realized through better adjustment of contract terms and work tasks to workers preferences and abilities, then contract drafters could learn what the optimal contract terms are and implement them without offering autonomy to their employees.

c. Choice Reduces the Negative Side-Effects of Norm Framing

Choice should also interfere with the framing effect of the performance norm the contract stipulates. The provision of a choice between multiple contracts suggests that individual production functions and actual productivity vary and workers are expected to choose the threshold contract that optimizes their performance. A dominant performance norm is not communicated and therefore subjects who would otherwise exceed the threshold selected should not lower their performance to conform to the threshold. Some dampening effect

might remain only above the highest threshold. Yet this is not a severe limitation, because contract drafters can stipulate an extreme threshold in one of the contracts without risking the possibility that norm-framing will discourage workers.

Therefore, we can predict that if norm-framing is muted, significantly more subjects in the *Autonomy* treatment should exceed the threshold of the contract they elect, versus in the single threshold treatments where subjects should be more likely to conform to the threshold (H_4)[4].

4. Contract Acceptance

Finally, we test the effect of choice on rates of contract acceptance. Depending on the scarcity of qualified work supply, the efficiency of contracts may not only rest on the productivity of the parties who accept the contract. In tight labor markets, it matters also how likely workers are to enter the contract in the first place. We have seen that gain-loss utility is negative below the threshold, reducing an individual's net outcome of a loss-framed contract. Therefore, standard theory would predict that loss framing should rather reduce contract acceptance rates.

However, if sophisticated workers use the loss-frame as a commitment device, then loss-framing can increase the expected net outcome of the contract for present-biased workers. Additionally, we assume the autonomy premium will also increase subjects' net outcome and thus makes entering the contract for them more attractive. As a result, production costs are less likely to exceed the gains of performance and we expect to observe fewer contract rejections in the *Autonomy* and *Self-Commitment* treatments compared to the single threshold contracts and even in comparison to *Baseline* (H_5).

IV. Methods

Participants are either current or former students of the University of Münster in Germany with all kinds of different disciplinary backgrounds; about 30% of the sample has graduated are employed outside of the University. We recruited the participants via the University`s central email server. They received a link in the email which became inactive once used, ensuring that the same participant could not complete the study more than once. The invitation informed the participants about the time it would require them to complete the experiment, thereby making it less likely that they discontinue participation after realizing they need longer to finish the study than they expected.

We implemented the study online to impose real opportunity costs on the participants. They can discontinue the real effort task whenever their opportunity costs exceed their expected benefit from performance, without influencing the behavior of other subjects. In contrast, in a laboratory if participants were permitted to leave after finishing their performance, they would likely influence the behavior of the subjects still working on the task and considering whether or not to continue. On the other hand, if participants would have to wait until all subjects are finished, they have barely any opportunity costs for performing.

In addition, the internet platform reduces participation costs unrelated to task performance and thereby allows us to sensitively measure contract acceptance rates. In a laboratory study by contrast, subjects will likely feel constrained to enter the contract, to recover their travel and participation costs.

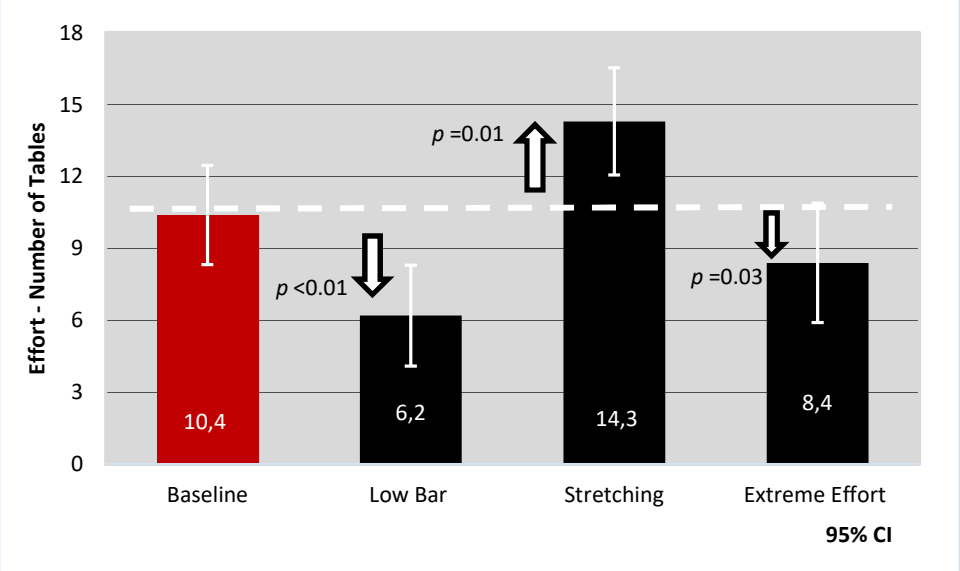
We use a set of control questions (see Appendix) to test participants' understanding of their task and incentives. Subjects are asked to calculate their payoff for different performance levels under all offered contracts. Participants are permitted to proceed with the study only if they answer the control questions correctly. All but two subjects passed this hurdle, which demonstrates that subjects understand that the alternative thresholds of the contracts they can choose between do not affect their monetary payment per table.

V. Results

1. Loss and Norm Framing

The results of the single threshold treatments that do not offer subjects a contract choice provide us with a benchmark for measuring the effectiveness of our *Autonomy* and *Self-Commitment* treatments. The results are U-shaped (see Figure 1): In the *Low Bar* treatment, subjects complete 6.2 tables; under the *Stretching* contract they invest a mean effort of 14.3 tables; while the *Extreme Effort* contract leads to a performance of only 8.4 tables. Reflecting the two risks of loss-framing, diminishing sensitivity and norm-framing, performance under both the *Low Bar* and the *Extreme Effort* contracts is significantly lower than the effort level we observe in *Baseline*, where under the *Linear* contract subjects complete 10.4 tables. Only subjects in the *Stretching* treatment exceed the productivity of subjects who perform under the *Linear* contract. The *Stretching* contract also leads to a higher performance than *Low Bar* and *Extreme Effort* (for all comparisons Mann-Whitney p-value <0.01). The results remain the same when we account for participants who reject the offered agreement. [5]

Figure 1: Hidden Costs of Loss-Framing (Brooks, Stremitzer, Tontrup JLS, 2017)



The data demonstrate that imposing loss frames on workers creates a substantial risk of dampening effort; if the loss-frame is not well adjusted, a plain *Linear* contract can lead to significantly better outcomes.

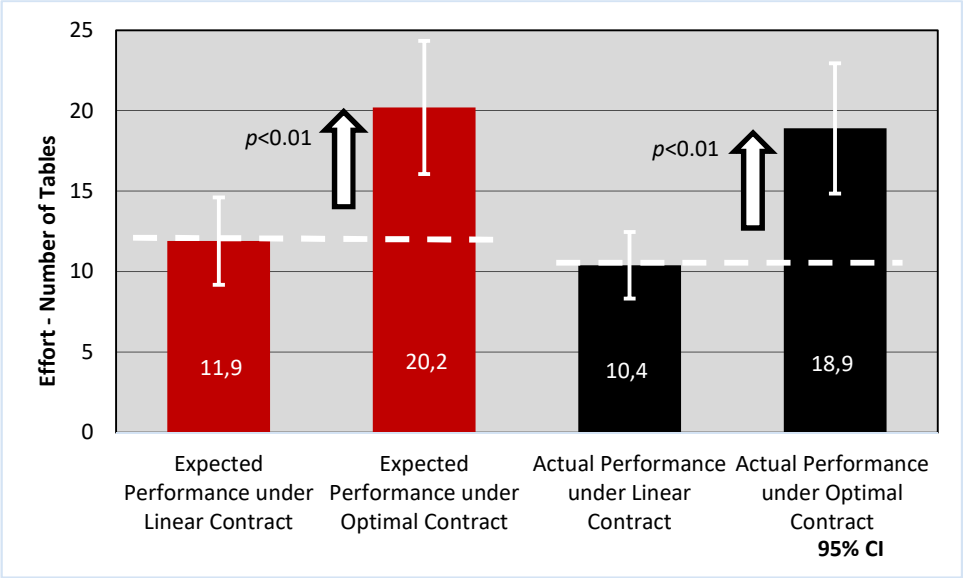
2. Self-Nudging: Evidence on Rational Self-Commitment

If given autonomy, we predict individuals can use loss-framing and their own loss aversion productively as a commitment device to nudge their future self to a better performance. In each *Choice* treatment subjects have a commitment option: By choosing a higher threshold, participants can extend the loss frame and lift their obligation, while their earnings per table remain the same. Thus, whenever subjects choose a threshold higher than the minimal frame they are offered, we can conclude that they want to self-commit to a higher performance level.

Subjects' contract choices strongly support our hypothesis that they deliberately aim to use loss-framing as a commitment device ($H_{1.1}$). In *Open Choice* 40 of 41 subjects deliberately extended the loss-frame and committed to a higher threshold (Fisher test *p*-value <0.01). On average, participants choose a threshold of 15.1 tables; this commitment significantly exceeds the mean effort of 10.4 tables that participants exert under a *Linear* contract (Mann-Whitney *p*-value 0.09). The *Autonomy* treatment provides further evidence for our hypothesis: 67 participants choose a *Stretching* contract and 34 even enter an *Extreme Effort* contract; only 54 stick to the minimum threshold and choose a *Low Bar* contract. Thus, a significant majority of subjects decides to voluntarily deepen the loss frame and to increase their obligation (65.1% (101 of 155); Fisher test *p*-value <0.01). In both treatments, subjects pay with increased negative gain-loss utility and a lower net outcome for their commitment. The *Self-Commitment* treatment however, makes their willingness-to-pay for the commitment device explicit, as subjects must pay a fee to enter a threshold contract that rises along with the threshold. Only 11 subjects avoid this fee and pick a *Linear* contract (19.7%), while the significant majority of subjects (80.3%; 45 of 56) chooses a threshold contract (Fisher test *p*-value <0.01).[6] Eleven of these subjects choose the *Low Bar* contract (19.7%), 21 subjects pick a *Stretching* contract (37.5%) and 13 choose the even more costly *Extreme Effort* contract (23.2%).

To reveal the motivation subjects have for choosing a loss-framed contract, we required subjects to indicate the effort they expect their future self to exert under all three threshold contracts as well as the *Linear* contract. With this data we calculate the mean performance each participant expects to reach across the three threshold contracts. When we compare this performance expectation for the threshold contracts with the effort the same subject expects her future self to invest under the *Linear* contract, the threshold contracts improve expected productivity significantly but the effect size is relatively small: 13.7 versus 11.8 tables (Wilcoxon *p*-value 0.02). The effect increases sharply, however, when we consider only the one contract under which a subject expects she will perform best. Under this optimal contract subjects expect to reach a mean productivity of 22.3 tables, almost doubling the performance compared to the 11.8 tables the same subjects expect their future self will complete under the *Linear* contract (Wilcoxon *p*-value <0.01). Apparently, subjects anticipate that the level of effort their future self is going to invest will largely depend on the contract they enter; they anticipate their own loss aversion and want to capitalize on it.

Figure 2: Expected and Actual Performance under Linear vs. Optimal Loss Framed Contract



As predicted the anticipation of their loss aversion allows individuals to employ a sophisticated strategy when choosing a contract. We assumed they would pick the loss-framed contract that (1) optimizes their future self's productivity without (2) making it fall short of the threshold and realizing a loss ($H_{1.2}$). Supporting the first element of this strategy, we find a vast majority of 92.8% (52 of 56) of the participants elect the contract under which they expect to be the most productive. For a significant majority of subjects (75.9%; N=44; Fisher test p-value <0.01), the contract they believe will make them most productive is loss-framed.[7] Thirty-nine of these 44 subjects indeed choose a loss framed contract (88.6%). Fourteen of the 56 subjects assigned to the treatment (24.1%) expect their performance to be equally good under a *Linear* contract, and 11 of these (78.5%) accordingly choose a *Linear* contract. This result is consistent with our expectation that many but not all subjects are loss averse and have time-inconsistent preferences; for these subjects the *Linear* contract leads to a higher net-outcome.

Now we turn to the second element: Loss aversion should discourage over-commitment and we therefore hypothesized that subjects would elect a contract they expect to comply with to avoid experiencing negative social and gain-loss utility ($H_{1.2}$). Indeed, participants' performance expectations and the contract they actually choose reveal that 92.9% of the subjects (52 of 56) select a contract they expect to fulfill (Fisher test p-value <0.01). Only 4 subjects (7.1%) choose a threshold contract they presume they will fail.

In sum, we find strong support for our theory that participants are capable of using a behaviorally sophisticated strategy for their contract choice: a strong majority of participants selects the contract they expect will optimize their future self's performance, while avoiding the experience of a loss from falling short of the threshold.

Only loss averse subjects can use the loss frame as a commitment device. We therefore assumed that those subjects our loss aversion measure can clearly classify as loss averse should expect to benefit more from choosing a threshold contract than participants who are either rational or only marginally loss averse ($H_{1.3}$). Accordingly, we predicted the difference in the performance expectations participants have for their optimal contract versus the *Linear* contract to be larger for the clearly loss averse subjects. Indeed, loss averse subjects benefit significantly more from binding to their optimal loss-framed contract: they expect to raise

their performance by 10.6 tables compared to the expectation they have for the *Linear* contract. By contrast, the subjects that are rational or so marginally loss averse that our measure could not detect their bias, expected to improve only by 6.2 tables. The difference is strongly significant (Mann-Whitney p-value <0.01). The effect holds when we include the subjects who reject entering an agreement (10.5 versus 6.2; Mann-Whitney p-value <0.01).

The fact that the subjects our loss aversion measure did not identify as loss averse also benefit from self-commitment likely suggests that our test was not sensitive enough to detect their bias. The expected gain that the measure's lottery offers exceeds the losses by a margin of €3. Some subjects who participated in the offered lotteries and who we therefore did not identify as being loss averse, may not have entered if the difference between gains and losses had been only €1 rather than €3. In that case a more sensitive test would have revealed their less pronounced loss aversion.

Finally, we move our analysis from expectations to actual performance to understand how effective the self-commitment strategy of the participants is. We have hypothesized that subjects should fulfill the contract they have chosen to avoid negative gain-loss and social utility and therefore should significantly improve their performance ($H_{1.4}$). The subjects who preferred to enter one of the three loss-framed contracts selected a mean threshold of 22.3 tables which is substantially above the performance under the *Linear* contract. A significant majority of 82.2% of the participants comply with the threshold contract they voluntarily committed to (Fisher test p-value <0.01).[8] Accordingly, the actual productivity in the self-commitment treatment (18.9 tables) is also significantly higher compared to the performance that we observe in the *Baseline* treatment under a *Linear* contract (10.4 tables; Mann-Whitney p-value <0.01).[9]

Importantly, in line both with our theory and subjects' expectations, loss averse participants benefit more from committing to a loss-framed contract: the loss-averse subjects who selected a loss-framed contract (all but one did so) reached a significantly higher productivity (22.6 tables) versus the rational (or relatively less biased) types (14.5 tables; Mann-Whitney p-value 0.02).[10] As an alternative measure for the effectiveness of self-commitment we compare the productivity subjects expect to reach under the contract they have chosen with their actual performance under that contract. Their actual mean productivity of 18.9 tables almost matches the 20.2 tables they expected to reach under the contract they did choose (Wilcoxon p-value 0.59). Thus, supporting $H_{1.4,}$ most subjects comply with the threshold they committed to, leading them to a significantly higher performance compared to subjects in the *Baseline* treatment (Mann-Whitney p-value <0.01).

In summary, the *Self-Commitment* treatment shows that loss-framing and contract choice leverage individual autonomy to improve performance by providing loss averse subjects with a commitment device they employ rationally in their own self-interest and to the contract parties' mutual benefit.

3. Disentangling the Effect of Choice and Autonomy on Performance

a. Preference Matching Improves Contract Performance

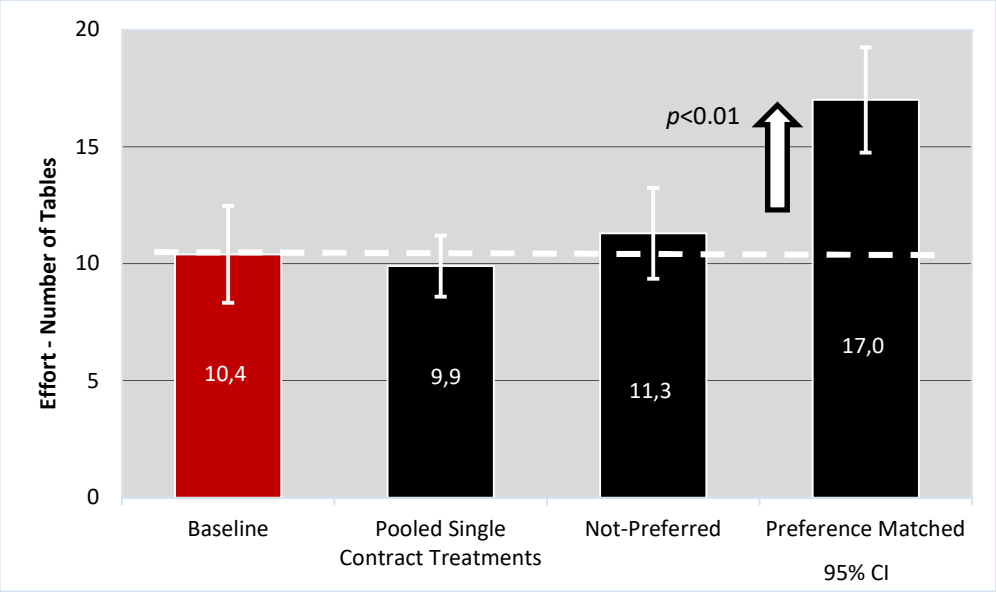
We have seen that loss-framing and self-commitment have almost doubled subjects' performance compared to the *Linear* contract that rests on monetary incentives alone. Now we will disentangle the mechanisms that may drive this positive impact on productivity. We

focus our analysis first on preference matching. We predicted subjects should improve their performance because choice allows them to select a contract that matches their production function ($H_{2.1}$).

Three-hundred subjects participated in the *Preference Matching* treatment, of which 275 accepted a contract. We elicited the subjects' contract preferences: 105 indicated that they prefer a *Low Bar* contract, 118 favor a *Stretching* and 77 an *Extreme Effort* contract. To isolate the effect of preference matching on performance we compare the effort level of the participants who received the threshold contract they indicated they prefer with those subjects who did not. The data support our hypothesis: Across the three threshold contracts, subjects who accepted the contract and whose preferences were matched (N=143) reach a mean productivity of 17.0 tables, while the participants who were randomly assigned a contract they would not have chosen (N=132) complete only a mean 11.3 tables (Mann-Whitney p-value <0.01). Results hold when we account for the 25 subjects who rejected the agreement (16.2 versus 10.3 tables).

To confirm the result we conduct an OLS regression with dummy variables for contract types and contract acceptance; another dummy variable that distinguishes whether subjects received their preferred contract or not estimates the main effect. The result holds and supports $H_{2.1}$: preference matching substantially improves performance (reg beta=-7.618 p-value <0.01).[11]

Figure 3: The Effect of Preference Matching on Performance



We can also measure the performance advantage preference matching generates by comparing the effort subjects invest in the *Preference Matching* treatment with the productivity we observe when we pool the single contract treatments *Low Bar*, *Stretching*, and *Extreme Effort*.[12] This comparison accounts for that whenever employers impose loss-framed contracts on workers without considering their actual preferences, some workers will nevertheless receive their preferred threshold. As expected ($H_{2.2}$), the participants (N=142) in the single contract treatments reach a significantly lower productivity (9.9 vs. 17.0 tables; Mann-Whitney, p-value <0.01).[13] An OLS regression including the same variables as above

(reg beta=5.220 *p*-value <0.01) confirms that preference matching substantially raised productivity.

b. The Autonomy Premium

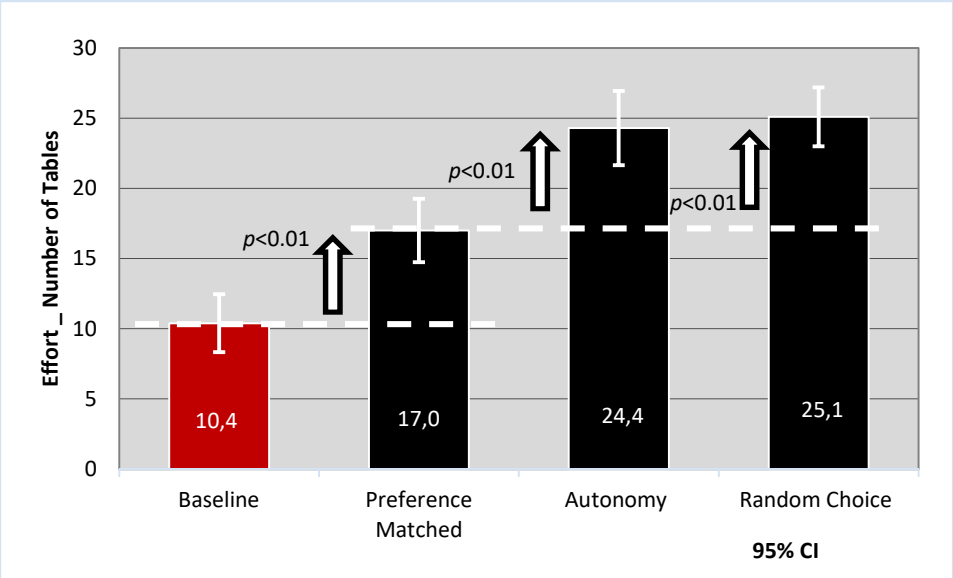
To isolate the autonomy premium and distinguish it from any influence preference matching has on productivity, we compare participants in *Autonomy* who are free to choose one of the contracts with the subjects in *Preference Matching* who are randomly assigned their favored contract. We theorized that the experience of self-determination utility will improve subjects' effort levels linearly, irrespective of the contracts they choose ($H_{3.1}$).

As expected, mean performance of subjects in *Autonomy* (24.4 tables *N*=153) is higher than subjects' productivity in the *Preference Matching* treatment (*N*=143; 17.0 tables; Mann-Whitney *p*-value <0.01). The effect holds when we include the 27 subjects (*N*=25 in *Preference Matching*) who rejected the contract (24.1 tables vs. 16.3; Mann-Whitney *p*-value <0.01).

We perform an OLS regression with dummy variables for contract type and acceptance; another dummy variable measures whether subjects received their preferred contract or not and an interaction term Matched Preference*Treatment estimates the main effect. The regression confirms a significant autonomy premium (reg beta=-12.645 *p*-value <0.01).

After establishing the autonomy premium, we want to identify the mechanism that is driving it. Our account of self-determination utility suggests we should find an increase of effort independently of which contract subjects choose ($H_{3.2}$).

Figure 4: The Autonomy Premium



As predicted, the autonomy premium is present irrespective of the contract type subjects decide to enter: participants who select a *Low Bar* contract (*N*=54; 16.4 tables) significantly exceed the performance of subjects in *Preference Matching* (*N*=47; 10.1 tables) whose preference for the *Low Bar* contract was matched randomly (Mann-Whitney *p*-value <0.01). Participants who elect a *Stretching* contract (21.3 tables; *N*=67) reach a higher mean

productivity than the participants in *Preference Matching* who complete only 16.5 tables (*N*=53) under the same contract (Mann-Whitney *p*-value <0.01). Finally, participants who choose the *Extreme Effort* contract (*N*=32) achieve a mean effort level of 44.3 tables, while subjects in *Preference Matching* who work under the same contract complete a mere 29.3 tables (*N*=35; Mann-Whitney *p*-value <0.01).

The results hold in regression analysis with dummy variables for contract type and acceptance, another dummy variable to distinguish whether subjects received their preferred contract or not, and the interaction term Matched Preference*Treatment which estimates the difference between *Autonomy* and *Preference Matching* treatments. The autonomy premium on subjects' performance under the *Low Bar* contract is measured by the two-way interaction of Matched Preference*Low Bar*Treatment (reg beta=-10.637 *p*-value <0.01). The effect is also significant for the *Stretching* (reg beta=-15.764 *p*-value <0.01) and the *Extreme Effort* contracts (reg beta=-8.573 *p*-value 0.04).

While this evidence supports our account of self-determination theory, it is inconsistent with either regret costs or commitment effects causing the autonomy premium. Regret costs cannot improve effort under the *Low Bar* contract; performance under that contract should have remained the same. Moreover, both regret costs and commitment effects can only lift subjects' performance up to the chosen threshold but cannot (unlike self-determination utility) increase performance above the threshold. Thus, if significantly more subjects raise their performance above the threshold in the *Autonomy* than in the *Preference Matching* treatment, the evidence supports self-determination utility and allows us to reject commitment effects and regret costs as the cause of the autonomy premium ($H_{3.3.}$). In support of self-determination utility as the causal mechanism, we find significantly more subjects in the *Autonomy* treatment exceed the contract threshold (68.6%; 105 of 153), than in *Preference Matching* (34%; 46 of 132; Fisher test *p*-value <0.01). A logistic regression, with dummy variables for contract types and acceptance rate, confirms the treatment effect (reg beta=-0.446 *p*-value <0.01). In particular, total performance above the threshold (all tables completed) is more than three times higher under the *Autonomy* treatment versus in *Preference Matching* (1056 tables vs. 325 tables).[14]

Finally, we analyze whether positive reciprocity might have added to the autonomy premium we observe. If positive reciprocity had an impact than the subjects in *Autonomy* should be more productive than the participants in the *Random Choice* treatment ($H_{3.4}$). However, performance in the *Random Choice* treatment (25.1 tables; *N*=43) does not fall short compared to the *Autonomy* treatment (24.4 tables; *p*-value 0.13 Mann-Whitney).

In summary, our results support that individuals derive substantial self-determination utility from their performance when they are granted autonomy. Importantly we can demonstrate that this autonomy premium is not an information effect driven by better adjustment of the work task to the subjects' preferences and abilities. The autonomy premium will only increase work performance if contract drafters offer workers actual choice.

c. Choice Eliminates Negative Side-Effects of Norm-Framing

Norm-framing causes workers to conform their effort choice to the stipulated performance norm. Offering a menu of threshold contracts suggests however that effort is approved within the range of the lowest and highest contract thresholds that the worker can choose between.

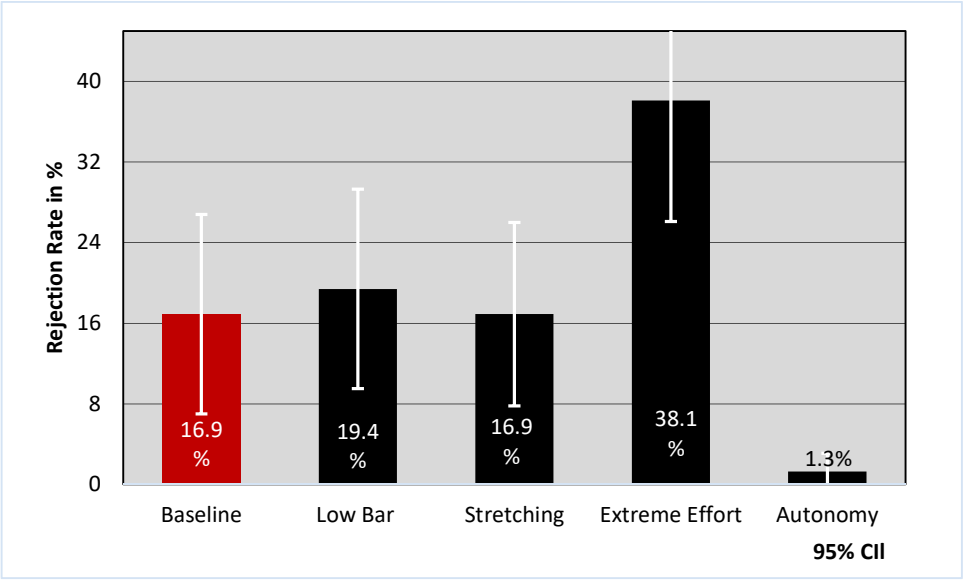
With norm-framing muted, we expect more subjects to exceed the contract threshold in *Autonomy* compared to the participants in the single threshold treatments (H_4). We pool the *Low Bar*, *Stretching* and *Extreme Effort* treatments and find that in these treatments only 34.2% (49 of 143) of the subjects exceed the prescribed threshold. By contrast, in the *Autonomy* treatment the percentage of participants who push their performance beyond the threshold is 68.6% (105 of 153) -- more than twice as high (Fisher test *p*-value <0.01).[15] We conduct a logistic regression with dummy variables for contract type and acceptance rate, and the treatment effect we want to estimate. The regression confirms that subjects are more likely to exceed the threshold suggesting that they experience no or less social disutility when they do (reg beta=2.04 *p*-value <0.01). We conclude that, in support of H_4, contract choice largely reduces subjects' conformity behavior – behavior that, in the single threshold treatments, drags down mean effort to the social norm.

4. Self-Commitment and Autonomy Increase Contract Acceptance

We predicted that self-commitment and the autonomy to choose between different gain and loss-framed contracts should reduce rejection rates (H_5). Indeed, only 2.7% (7 of 254) of the participants across all *Choice* treatments (*Open Choice N*=3, *Autonomy N*=2, and *Self-Commitment N*=2) dismissed the contract offer. This rejection rate is considerably lower compared to the pooled single contract treatments including *Baseline* (21.2%; 54 of 255; Fisher test *p*-value <0.01). Additionally, we perform a logistic regression with dummy variables for contract type and acceptance and a dummy that estimates the difference between the *Autonomy* and the pooled single threshold treatments we are interested in. The effect of choice on contract acceptance remains significant (reg beta=3.16 *p*-value <0.01).

The data demonstrate also that subjects are not simply more likely to accept an agreement because the choice treatments allow them to opt out of a higher-threshold into a lower-threshold contract. The rejection rate in the *Autonomy* as well as the *Self-Commitment* treatments is lower than under both the *Low Bar* and *Baseline* treatments (for all comparisons Fisher test *p*-value <0.01).

Figure 5: Contract Rejection Rates across Treatments



The results support our hypothesis H_5: with self-commitment and the autonomy premium elevating productivity, acceptance rates rise along with the outcomes the subjects expect to reach.

5. Economic Value of Work Contract Offering Self-Commitment and Autonomy

We have demonstrated the comparative advantage of a behavioral work contract design that does not impose loss framing on employees but offers it as an optional self-commitment device. We have also disentangled the impact of each mechanism that we assume contributes to the effectiveness of this contract design: i.e., preference matching, the autonomy premium, self-commitment, and eliminating negative side-effects of norm-framing.

Now we want to determine the economic value of this autonomy preserving contract design by estimating the costs of bringing about an equivalent enhancement of productivity by raising monetary performance incentives alone. For this estimation we doubled the performance incentive of the *Linear* contract, paying subjects €2 per table instead of €1. Doubling the incentives elevated mean productivity from 10.4 to 21.6 tables, suggesting that performance in our effort task may almost be a linear function of payment. Strikingly, productivity in this *Double-Incentive Baseline* treatment still falls short of performance under the *Autonomy* treatment even though subjects are paid $2 instead of $1 per table (Mann-Whitney *p*-value 0.04). Assuming a linear relationship between performance incentives and productivity, we can approximate the monetary value of the behavioral contract design we propose. In order to increase subjects' mean productivity by one table, the payment per table would have to be raised by 9.2 cents. Thus, in order to reach the productivity of subjects who perform under the *Autonomy* contract, incentives would have to be set at €2.29 per table. For example, in order to produce 25 tables a worker would have to be paid €25 under the *Autonomy* contract, while under a *Linear* contract she would have to be incentivized with €57.25 to produce the same 25 tables. To equal the performance that we measure under the *Self-Commitment* treatment, workers would still have to be incentivized with €1.79 per table under Baseline's *Linear* contract. The results show the enormous economic value of sophisticated behavioral contract design and expresses the extensive intrinsic value people place on their (work) autonomy in monetary terms.

VI. Policy Implications and Conclusions

Our findings suggest a variety of policy applications and directions for future research, both within contract law and design, but also for many regulatory tasks at the intersection of private and public law.

1. Internal and External Validity

To make the work task as realistic as possible, the payment matches what students would earn in typical part-time jobs at the University or elsewhere to finance their studies and living expenses.

But our study differs also in some respects from real-world contracting environments. We eliminate communication and context outside of the contract to cleanly isolate the effects the changes of the contractual terms have on work performance and to identify the

mechanisms that cause these effects. However, extra-contract communications may establish a work culture that is neither explicitly expressed nor even suggested by the contract itself.

Communication and work culture may weaken the effects of contract choice if an employee realizes that co-workers follow informal rules different from the norms the contract suggests, or when the principal signals standards that negate what the contract appears to offer. But work culture may also reinforce our findings, if the employer encourages workers to exercise the choice options their work contract offers and if co-workers act upon the freedom the contract provides. Employers exercise substantial influence over both the contract and work culture. To make behavioral contracting successful, employers must ensure that the choice and autonomy the contract grants are not undermined by informal social rules and practices of the workplace.

In real world workplaces financial incentives can be so high-powered that workers may view them as controlling, crowding out the effects on internal work motivation that we report (Frey 1997). However, often compensation is not the sole factor to determine work effort; intrinsic motivation is crucial whenever workers have key qualifications or job security, or when performance is difficult to observe and workers can shirk. We expect compensation to mediate the strength of the effects we report; higher payment for example with the independence it brings may also strengthen the effects we find.[16]

For the relevance of our results it is also important, how stable the effects of self-commitment and autonomy may be in repeated work situations. Benartzi and Thaler (1995) analyzed trading behavior with bonds and stocks; over hundreds of repetitions their subjects preferred the safer bonds, even though they realized that the returns of stocks had been systematically higher. Camerer et al. (1997) observed NYC cabdrivers over a period of multiple years, suggesting that they formed a reference point of expected earnings for a given work-day. They drove longer hours on bad days (i.e., on days with nice weather, where taxi usage tends to be lower) to avoid a loss relative to that reference, while they would have maximized their hourly income working longer on good (cold and rainy) days. In neither study did experiencing losses weaken workers' loss-aversion. Learning might even increase the effectiveness of bias-self-management: subjects might learn how productively they can use their own loss aversion as a commitment-device. There are also indications that autonomy's effects on work productivity may be persistent: survey data shows that long-time job satisfaction and identification increase with the autonomy employees have (for example Wheatley 2017; Bradely et al. 2003).

Our number-counting task keeps intrinsic motivation purposefully low, motivating subjects only via the monetary incentives that we offer. We had four methodological reasons for this design. First, our study is designed to identify an effect of intrinsically-valued autonomy on contract performance. Extending an employee's autonomy will often change the work process, making it more creative or engaging, and as a result more enjoyable. Implementing a simple and tedious task ensures that our manipulations do not alter the characteristics of the task. This allows us to conclude that subjects lift their effort because they place intrinsic value on being granted autonomy, and not because they derive more satisfaction from completing a task they experience as more interesting because of our manipulation.

Second, we wanted to design a formal and minimal manipulation of autonomy. People are more likely to experience self-determination utility when they make more meaningful life

choices -- where to live, with whom to interact or do business, and such basic elements of individuals' control over their lives such as whether one plans her workday or must adhere to someone else's plan (see the qualitative study by Eaton-Walley and Low 2017). That we find a strong autonomy premium, even though subjects are exercising decision-making power over something as trivial as how many numbers they want to count, makes a very strong case for the autonomy effects we report and the value that people place on their work autonomy.

Third, the task focuses on effort rather than knowledge or skills and is therefore well suited to allow subjects to improve their performance by self-commitment.

Finally, the task, allows us to eliminate any skill-related variance between subjects, which could only have added noise to our data.

2. Policy Implications and Directions for Future Research

Our findings have implications across many different areas of law and policy. In the most closely relevant private law context, contracting, our results suggest that offering workers a choice of loss and gain-framed contracts can be mutually beneficial for both sides of the contract and across a range of different types of contractual relationships. Business contracts have rarely been loss-framed in the past (Baker et al., 1988; Lazear, 1991). Contractors may have believed that trying to take advantage of loss aversion will raise the disapproval of counterparties and deter potential partners from entering an agreement (see Luft 1994, Imas et al. 2017). Or offerors may have expected that they would have to pay a premium to entice workers to accept a loss-framed contract. In their eyes the productivity advantage of loss framing may not have outweighed these extra costs.

Recently however, as the Øresund example illustrates, loss-framing seems to have become a more common method of addressing problems typical for service or construction contracts in employer-agent relationships of independent players – i.e., where little authority can be exercised over the partner, with an asymmetry of information and often also expertise between the parties exacerbating the monitoring problem. Incentives can attenuate the lack of control, when the party has sufficient information and expertise to set sophisticated thresholds. The Øresund construction contract tried to circumvent the information problem by specifying a deep loss-frame with a very high threshold to push construction firms to work harder no matter what level of suspension they could possibly realize; a detrimental solution that accounts neither for diminishing sensitivity nor for norm-framing. Offering a menu of loss-and gain framed contracts avoids these informational risks and allows workers to turn their own loss aversion into a commitment device.

The autonomy premium we show evidence for has implications beyond loss-framing and has already entered the design of work contracts and work culture. Firms offer employees flexible work contracts that allow them to choose where and when they want to work (see Brewster et al. 1997 and Martinez-Sanchez et al. 2007). Google, Dropbox, and Netflix leave it to their employees to decide how much vacation they want to take (see Patty McCord 2014; McKinsey 2017). Most of these firms have their business in creative industries where quality and effort are less easily measurable limiting the direct control over the employees and their work. A recent field study has shown a positive performance impact of such policies (Bloom et al. 2014; with survey evidence de Menezes Kelliher 2011). Their study's results align with our findings: autonomy seems to improve productivity and work quality (our results on quality effort are reported in the Appendix). Employees appear to self-select into demanding

work schemes.[17] These policies also attract talent as workers prefer employers that implement them (see McKinsey 2017).

Our findings also provide suggestive evidence for applications to public law and regulation. In many areas people are governed by a mix of rules at the intersection of public and private law. Privately-sponsored retirement savings plans supposed to raise the savings rate of present-biased workers are an example where loss-framed contracts can be used to enable individuals' self-nudging. Many American workers appear to have a present bias and tend to save less for their retirement than is optimal. A survey study by Choi et al. (2006) suggests that many of these workers are aware of their present bias: two-thirds of the respondents in the Choi et al survey study indicate that rationally they should save more than they currently do. Policy interventions have attempted to increase savings by providing tax exemptions that incentivize employers to offer beneficial retirement saving plans for their employees. The positive effect on savings however, appears to be small (see Chetty et al. 2013; Madrian and Shea, 2001) for two reasons: Retirement plans under the current policy offer present-biased workers an employer contribution that matches ex post the money they have actually saved. However, as this plan design promises rewards only for the future, present-biased workers respond to these incentives only weakly. Ironically, the workers' present bias is both the reason why the regulation was created and the reason why it fails. A second cause for failure is that in the current regulatory framework most employers are not interested in increasing actual savings. They offer retirement plans to attract employees with what appears to be a higher overall compensation combining wage and plan benefits, but once the employees are hired, companies want to economize on the matching contributions they have to pay. Consequently, they offer matching contribution plans with low default contributions (see Bubb et al. 2015 and Batchelder 2018) that most workers conform to, including those who would have saved more if the default was absent (Madrian and Shea 2001).

Our findings suggest that a more effective plan design should be loss framed to provide myopic workers with a commitment device allowing them to nudge their future self to save more. Workers may choose between different saving targets adjusted to their income; their choice entitles them to the employer's contribution that matches the savings goal they select. In contrast to current plan designs, employees should receive this contribution in advance (i.e., ex-ante instead of ex-post matching contributions); the advance is theirs, however, it can only be used for retirement. If workers fail to meet their chosen savings goals, they would then lose these employer contributions. Because workers get the matching contributions in advance rather than in a (distant) future, this plan design overcomes their present bias. Additionally, loss aversion pushes the employees to in fact meet their savings goals.

Regulation may either directly mandate a retirement contract design with advance contributions and socially-optimal defaults or incentivize employers to offer such plans voluntarily, for example by conditioning an employer's tax exemptions on reaching a defined minimum rate in actual savings across all their employees. Of course, regulators could also mandate non-elective contribution plans,[18] but our results suggest that preserving autonomy with smart defaults may have comparative advantages for increasing actual savings rates: employees may select and comply with more ambitious savings targets, compared to the compulsory savings rates that may realistically be politically achievable.[19]

Another interesting path for future research is the choice between rules and standards in structuring legal policies. The debate focuses on the relative costs of designing rules versus standards, and costs of enforcement, and compliance in both settings (see Kaplow 1992). To create an effective and fully-specified rule the law-maker must obtain more detailed knowledge of the regulatory target, expertise which may be difficult and costly to acquire or not available at all, compared to an open standard that can be filled in ex post, often by courts. On the other hand, for those subject to the law, self-enforcement can be cheaper under a regime of rules, because the well-specified rule assures them of a way to comply with the law, while a standard often requires interpretation and therefore imposes risk. To avoid liability, people may be more likely to overinvest in compliance when they are provided with a mere standard.

Our study may add a new perspective to this debate. Legal commands structured as open-ended standards prescribing "reasonable" behavior present people with some autonomy in conforming their conduct to the law. For example, a standard directing drivers to proceed at a "reasonable speed under the circumstances" empowers them to make a responsible and autonomous decision taking factors like the density of traffic or the weather conditions into account. Our findings encourage exploring whether individuals are more motivated to comply with a legal standard because the standard grants them a more autonomous decision, relative to legal rules that tightly prescribe expected conduct (see Waldron 2017). Tight monitoring and enforcement are often not only expensive but for an open society not desirable. Could the use of standards, in comparison to rules, improve people's intrinsic inclination to comply with the objectives of the law?

A similar question can be asked in contract law. Workers tend to strategically reduce quality effort to increase quantity when limited monitoring allows for it. Heywood et al. (2013) find that this effect can be mediated by strengthening intrinsic work motivation. The autonomy premium may have the same effect and reduce strategic shirking behavior considerably.

3. Theoretical Implications: Bias-Self-Management

Our study demonstrates that individuals can strategically employ their loss aversion to their advantage. The subjects nudge their future self to work harder and earn a higher payoff. They do so by voluntarily self-selecting into a choice architecture that elicits their loss aversion and commits their future self to the productivity goal their current self prefers. This phenomenon of *self-nudging* should change our perspective on biases. Biases can cause cognitive error, but our study shows that they can also be a valuable tool for individual decision making.

Most behaviorally-informed regulation seeks to identify potential biases and to design interventions to overcome them. The intervention typically takes one of three forms: 1) it insulates biases from affecting outcomes (with many examples Hanson and Kysar 1999a, b), 2) it aims to debias individuals (Jolls and Sunstein 2006) or, 3) it nudges them towards a better outcome (Taylor and Sunstein 2008). An insulating intervention is designed to prevent a bias from becoming consequential. For example, when biased consumers tend to overestimate their expertise in handling a dangerous product, a prohibition to use the particular good does not remove the consumer's bias, but prevents it from causing any harm.

In contrast, a debiasing intervention seeks to remove the bias and enable individuals to make an unbiased decision. An example is using natural frequencies instead of conditional probabilities when informing patients of the risks and benefits of a potential treatment, an intervention which prevents both patients and doctors from falling prey to the base rate fallacy (for a numerical illustration see the footnote and Gigerenzer et al. 2007).[20]

Nudging aims to change the choice architecture an individual is presented with to trigger particular heuristics and decision-making routines that likely lead the chooser to a favorable outcome, as determined either by her personal welfare or by a social standard. An example of a nudge that aims at a public good is implementing a default for becoming an organ donor; if the legal default treats individuals as donors, it focuses them on the social benefits of being a donor, i.e. the possibility to save multiple lives with their donations (Johnson et al., 2003). The opposite default leads individuals to more easily construct the advantages of not being a donor, for example they do not face the fear they might be declared dead too easily once they are a donor, and their body remains "intact" after their death.

All of these interventions tend to assume that decision-makers are either not aware of their bias or at least are unlikely to overcome the bias without the regulation. Here we propose an alternative approach. We refer to people's capacity to avoid, overcome, or to even exploit their biases to their advantage as *Bias-Self-Management* (BSM). BSM rests on the assumption that individuals are sometimes aware of their biases, and can act to improve their decision-making outcomes by deliberately employing the same three strategies that otherwise law-makers tend to use: 1) they self-debias; 2) they insulate their own biases; and 3) they engage in self-nudging. To do so, individuals may access and use some institution or mechanism that facilitates the BSM strategy. We have provided an example of one such institution in our experiment – a loss-framed contract serving the decision-maker as a commitment device.

While this article analyses the BSM equivalent of nudging (self-nudging), Arlen and Tontrup (2015a) focus on the equivalent to debiasing. Their study shows that property owners often are aware that ownership creates a bias and deliberately involve agents when selling their property. The responsibility that the agent assumes for making the trade allows the owners to give the agent unbiased instructions for the sale by reducing their personal regret costs over losing their property. There is also an analogy to an insulating intervention: Arlen and Tontrup (2015b) show that instead of removing their bias, individuals often try to deprive their bias of its effect, such that it does not prevent them from making their absent the bias preferred decision. In their study, owners, when anticipated regret over selling their property burdens an otherwise beneficial trade, strategically access market information to learn whether other participants have traded. The information can shift their focus of attention from the status quo to the choice of the others, and thus from their endowment to the gains from trade. By accessing the market information, owners do not de-bias, but by shifting their focus towards the benefits of the trade they strategically relocate their bias.

BSM identifies and analyzes the decision-making strategies that people in fact use to manage their biases. This property makes BSM policy-relevant. Many studies in psychology and economics by contrast try to construct optimal debiasing strategies and test their performance relative to rational benchmarks. For example, Roese et al. (2012) suggest a strategy to suppress hindsight bias; they ask subjects to think of a chain of events that would have led to an alternative outcome. Or Herzog et al. (2009) propose a technique to increase the accuracy of predictions under uncertainty, which are often biased by saliency effects.

After an initial best guess of a quantity, the decision-maker assumes that the first estimate was wrong, guesses again, and then averages the two estimates. These studies do not analyze whether people are in fact aware of being biased without an intervention, nor whether they are already using some (effective) debiasing strategy. For legal policy this is a crucial elision. Legal policy is interested in the gap between actual and optimal behavior; an intervention becomes sensible if the gap between the two reveals sufficient room for improvement. If people use BSM strategies effectively, a standard behavioral economics intervention may not improve decision quality so substantially that it exceeds the regulatory costs. This may be true even if the BSM strategy is in itself not optimal.

This study shows that BSM can have cost advantages over external regulatory interventions. BSM imposes costs only on those parties who are in fact biased and choose to overcome their bias or take advantage of it. In our study, only the subjects with time-inconsistent preferences can benefit from choosing a loss-framed contracts; they have to sustain the costs of loss-framing, if they want to use their loss aversion as a commitment device. All other subjects however, can choose and are better off with a gain-framed *Linear* contract. In contrast, a contract that does not allow subjects to choose between a loss frame and a gain frame imposes costs on *all* subjects, whether they have time-inconsistent preferences or not. Indeed, external policy interventions often cannot isolate individuals who are biased and address them separately from those who are not; in that case the intervention treats everyone the same imposing the regulatory costs on all people subject to the regulation whether they benefit from the intervention or not (see Arlen and Tontrup 2015a; Korobkin and Ulen 2000; Rizzo and Whitman 2009).

BSM might also help save indirect costs of regulation. For example, interventions that restrict people`s choices to push them in a particular direction have been reported to trigger rejection and reactance (see Arad and Rubinstein 2015). The relatively low productivity under the single-threshold contracts in our study, combined with the much higher rates of contract rejection, might be an example of such dampened motivation. BSM by contrast relies on individual choice and autonomy. We have shown that they strongly support intrinsic motivation. Sunstein (2017) presents new evidence pointing in the same direction: that preserving choice increases acceptance among those subject to a policy intervention.

Further, unlike an external nudge, BSM avoids the difficulty of having to identify the true preferences of the biased individuals the intervention is supposed to make better off (Goldin 2017). If people are responsive to nudges then their choices are altered by elements of the decision-making context that are unrelated to their actual preferences (for example, by defaults). Therefore, once a nudge is implemented, the choices of those affected by the nudge may no longer reveal their true preferences to the policy maker.[21] Before the intervention the individual's bias prevents the policy maker from observing their true preferences in their actions. BSM avoids this problem, because individuals themselves identify and manage their biases according to their true preferences.

Our study also shows that policy makers, or in our case contract drafters, can support and encourage BSM by providing decision makers with institutions or decision-making infrastructure. Behavioral contract design that allows workers to use their own loss aversion productively as a commitment device is the example we give in this paper. Arlen and Tontrup (2015a, b) have analyzed agency and market information serving decision makers to overcome regret costs and endowment bias.

We propose BSM as a complementary instrument in the behavioral toolbox of regulators and contract drafters. While BSM can have significant advantages, it also has limits and should not be misunderstood as a replacement for legal interventions. BSM requires sophistication: decision makers must be aware of their bias and they need to be able apply a countervailing strategy like the subjects in our study who commit their myopic future self to their preferred productivity goal. But some individuals are less likely than others to be aware of their cognitive and motivational biases. And some biases are more likely to elude notice by even the most self-aware individuals. Awareness of the hindsight bias, for example, is rare, as it effectively changes the decision-maker's memory (Hoffrage et al., 2000). To fill out the full potential of BSM, future research must identify the legal and social contexts in which people can effectively recognize their biases, manage them or employ them to their advantage. Present bias, lack of self-control, and risk aversion are other obvious candidates of motivational biases that decision makers may be able to deliberately overcome by delegating choices or imposing self-commitments.

References

Arlen, J. and S. Tontrup (2015 a): *Does the Endowment Effect Justify Legal Intervention? The Debiasing Effect of Institutions*. The Journal of Legal Studies 44 (1), 143-182.

Arlen, J. and S. Tontrup (2015 b): *Strategic Bias Shifting: Herding as a Behaviorally Rational Response to Regret Aversion.* Journal of Legal Analysis 7(2), 517–560.

Abeler. J.; Falk, A.; Goette, L., and D. Huffman (2011): *Reference Points and Effort Provision.* American Economic Review 101(2), 470-92.

Altmann, S., and A. Falk (2009): *The Impact of Cooperation Defaults on Voluntary Contributions to Public Goods.* mimeo.

Arad, A., and Rubenstein, A. (2017): *The People's Perspective on Libertarian-Paternalistic Policies*. Forthcoming.

Babcock, P., Bedard, K., Charness, G., Hartman, J., and H. Royer (2015): *Letting down the Team? Social Effects of Team Incentives*. Journal of the European Economic Association 13(5), 841-70.

Bartling, B., Fehr, E., and H. Herz (2014): *The Intrinsic Value of Decision Rights*. Econometrica 82(6), 2005-39.

Batchelder, L. (2018): *Improving Retirement Savings Choices Through Smart Defaults*. Unpublished manuscripts.

Baujard, A. (2007): *Conceptions of Freedom and Ranking Opportunity Sets: A Typology*. Homo Oeconomicus 24(2), 231–54.

Bazerman, M.H. (2018). Judgment and decision making. In R. Biswas-Diener & E. Diener (Eds), *Noba textbook series: Psychology*. Champaign, IL: DEF publishers.

Bell, D. (1982): *Regret in Decision Making under Uncertainty*. Operations Research 30(5), 961-81.

Bénabou, R., and J. Tirole, (2002): *Self-Confidence and Personal Motivation*. Quarterly Journal of Economics 117(3), 871-915.

Shlomo Benartzi and Richard Thaler (1995): *Myopic Loss Aversion and the Equity Premium Puzzle*. The Quarterly Journal of Economics, 110(1), 73-92.

Bloom, N., Liang, J., Roberts, J., and J.Y. Zhichun (2015): *Does Working from Home Work? Evidence from a Chinese Experiment*. The Quarterly Journal of Economics 130(1), 165–218.

Brooks, R.W.; Stremitzer, A., and S. Tontrup (2017): *Stretch it don't Break it –The Hidden Costs of Contract Framing*. Forthcoming Journal of Legal Studies 46(2), 399-426.

Brooks, R.W.; Stremitzer, A., and S. Tontrup (2011): *Contracts - Why Loss Framing Increases Effort*. Journal of Institutional and Theoretical Economics 168(1), 62-82.

Brunnermeier, M.K., and J.A. Parker (2005): *Optimal Expectations*. American Economic Review 95(4), 1092-1118.

Bubb, R., Corrigan, P., and P.L. Warren (2015): *A Behavioral Contract Theory Perspective on Retirement Savings*. Connecticut Law Review 47, 1317-66.

Bubb, R., and R.H. Pildes (2014): *How Behavioral Economics Trims its Sails and Why*. Harvard Law Review 127, 1594–678.

Camerer, C.F., Babcock, L., Loewenstein, G., and R.H. Thaler (1997): *Labor Supply of New York City Cab Drivers: One Day at a Time.* Quarterly Journal of Economics 112(2), 407-41.

Casscells, W., Schoenberger, A., T.B. Grayboys (1978): *Interpretation by physicians of clinial laboratory results*. New England Journal of Medicine 299(18), 999–10001.

Charness, G. (2000): *Responsibility and Effort in an Experimental Labor Market*. Journal of Economic Behavior and Organization 42(3), 375–84.

Charness, G., and U. Gneezy (2010): *Portfolio Choice and Risk Attitudes: An Experiment.* Economic Inquiry 48(1), 133-46.

Chlaß, N., Güth, W., and T. Miettinen (2014): *Purely Procedural Preferences: Beyond Procedural Equity and Reciprocity*. Max Planck Unpublished Working Paper.

Chaudhry, S., and D. Klinowski (2016): *Enhancing Autonomy to Motivate Effort: An Experiment on the Delegation of Contract Choice*. Unpublished Working Paper.

Chetty, R., Friedman, J., Leth-Petersen, S., Nielsen, S., and T. Olsen (2014): *Active vs. Passive Decisions and Crowd-out in Retirement Savings Accounts: Evidence from Denmark.* Quarterly Journal of Economics 129(3), 1141-219.

Choi, J., Laibson D., Madrian B., and A. Metrick (2006): Saving for Retirement on the Path of Least Resistance. In: McCaffrey E. and J. Slemrod, *Behavioral Public Finance: Toward a New Agenda*. New York: Russell Sage Foundation; 304-51.

Christman, J. (2011): *Autonomy in Moral and Political Philosophy*. The Stanford Encyclopedia of Philosophy, ed. by N. Zalta. Available at http://plato.stanford.edu/archives/spr2011/entries/autonomy-moral/

Dal, B.P., Foster, A., and L. Putterman (2010): *Institutions and Behavior: Experimental Evidence on the Effects of Democracy.* American Economic Review 100(5), 2205-29.

Deci, E.L., Koestner, R., and R.M. Ryan (1999): *A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation*. Psychological Bulletin 125(6), 627-68.

Deci, E.L. and R.M. Ryan (2000): *The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior*. Psychological Inquiry 11(4), 227–68.

DeCharms, R. (1968): *Personal Causation: The Internal Affective Determinants of Behavior*. New York: Academic Press.

Dickinson, D. and M.C. Villeval (2008): *Does Monitoring Decrease Work Effort? The Complementarity between Agency and Crowding-out Theories*. Games and Economic Behavior 63(1), 56–76.

Dominguez-Martinez, S., Sloof, R. and F. von Siemens (2014): *Monitored by Your Friends, not Your Foes: Strategic Ignorance and the Delegation of Real Authority*. Games and Economic Behavior 85(c), 289-305.

Dowding, K. and M. van Hees (2009): *Freedom of Choice*. In Handbook of Rational and Social Choice. Oxford University Press, Oxford, UK.

Eaton-Walley, T. and M. Lowe (2017): *The Importance of Work and Job Autonomy and Independence to Professional Staff employed in Local Government and Different Career Stage*. International Journal of Applied HRM 2(3), 1–11.

Engel, C. (2002). *Abfallrecht und Abfallpolitik*. Baden-Baden: Nomos.

Falk, A., and M. Kosfeld (2006): *The Hidden Costs of Control*. American Economic Review 96(5), 1611-30.

Falk, A., and U. Fischbacher (2206): *A Theory of Reciprocity*. Games and Economic Behavior 54(2), 293-315.

Fehr, E., Herz, H., and T. Wilkening (2013): *The Lure of Authority: Motivation and Incentive Effects of Power.* American Economic Review 103(4), 1325-59.

Fehr, E.; Klein, A., and K. M. Schmidt (2007): *Fairness and Contract Design.* Econometrica 75(1), 121-54.

Frederick, S. (2005): *Cognitive Reflection and Decision Making.* Journal of Economic Perspectives 19(4), 25-42.

Frenz, W. (1996): *Die Verwirklichung des Verusacherprinzips im Abfallrecht*. Berlin: Dunker & Hombolt

Frey; B. S. (1997): Not for the Money – An Economics theory of Edgar Elgar Publishing.

Fryer, R.G., Levitt, S.D., List, J. and S. Sadoff (2012): *Enhancing the Efficacy of Teacher Incentives Through Loss Aversion: A Field Experiment*. NBER Working Paper No. 18237.

Gill, D. and V. Prowse (2012): *A Structural Analysis of Disappointment Aversion in a Real Effort Competition*. American Economic Review 102(1), 469-503.

Gigerenzer, G. (2014): *Risk Savvy: How to Make Good Decisions*. London: Allen Lane.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L.M., and S. Woloshin (2008): *Helping Doctors and Patients Make Sense of Health Statistics*. Psychological Science in the Public Interest 8(2), 53–96.

Goldin, J. (2015): *Which Way To Nudge? Uncovering Preferences in the Behavioral Age*. Yale Law Journal 125, 226-270.

Goette, L; Huffman, D., and E. Fehr (2004): *Loss Aversion and Labor Supply*. Journal of the European Economic Association 2(2-3), 216-28.

Hanson, J.D., and D.A. Kysar (1999). *Taking Behavioralism Seriously: Some Evidence of Market Manipulation*. Harvard Law Review 112(7), 1420–572.

Hanson, J.D., and D.A. Kysar (1999). *Taking Behavioralism Seriously: The Problem of Market Manipulation*. New York University Law Review 74, 630–749.

Hayek, F.A. (1945): *The Use of Knowledge in Society*. American Economic Review 35(4), 519-30.

Hennig-Schmidt, H., Sadrieh, A., and B. Rockenbach (2010): *In Search of Workers Real Effort Reciprocity a Field and a Laboratory Experiment.* Journal of the European Economic Association 8(4), 817-37.

Heywood, J.S., Siebert, S., and W. Xiangdong (2013): *The Consequences of a Piece Rate on Quantity and Quality: Evidence from a Field Experiment.* Unpublished Discussion Paper, Institute for the Study of Labor, No. 7660.

Hoffrage, U., Lindsey, S., Hertwig, R., and G. Gigerenzer (2000): *Communicating statistical information*. Science 290, 2261–62.

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.

Holt, C. A., and S. K. Laury (2002): *Risk Aversion and Incentive Effects*. American Economic Review 92(5), 1644-55.

Hoppe, E. I., and P. W. Schmitz (2011): *Can Contracts Solve the Hold-Up Problem? Experimental Evidence.* Games and Economic Behavior 73(1), 1861-99.

Hossain, T.; and J. List (2012): *The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations.* Management Science 58(12), 2151-67.

Imas, A., Sadoff, S. and A. Samek (2015): *Do People Anticipate Loss Aversion?*. Management Science 63(5), 1271-84.

Jolls, C., and Sunstein, C. (2006): *Debiasing Through Law*. Journal of Legal Studies 35(1), 199–242.

Johnson, E.J., and D.G. Goldstein (2003). *Do Defaults Save Lives?* Science 302, 1338–39.

Kahneman, D., J. L. Knetsch and R. Thaler (1990): *Experimental Tests of the Endowment Effect and the Coase Theorem.* Journal of Political Economy 98(6), 1325-48.

Kahneman, D., J. L. Knetsch and R. Thaler (1991): *Anomalies: The Endowment Effect, Loss Aversion, and Status-Quo Bias.* Journal of Economic Perspectives 5, 193-206.

Kahneman, D., and A. Tversky (1979): *Prospect Theory: An Analysis of Decisions under Risk. Econometrica* 47(2), 263-92.

Kaur, S.; Kremer, M., and S. Mullainathan (2010): *Self-Control and the Development of Work Arrangements*. American Economic Review 100(2), 624-28.

Koehler, J.J. (1996): *On conveying the probative value of DNA evidence: frequencies, likelihood ratios, and error rates*. University of Colorado Law Review 67, 859–86.

Korobkin, R.B., and T.S. Ulen (2000): *Low and Behavioral Science: Removing the Rationality Assumption from Law and Economics*. California Law Review, 88(4), 1051-44.

Koszegi, B., and M. Rabin (2006): *A Model of Reference-Dependent Preferences.* Quarterly Journal of Economics 121(4), 1133-65.

Laibson, D. (1997): *Golden Eggs and Hyperbolic Discounting.* Quarterly Journal of Economics 112(2), 443-78.

Laury, S.K. (2005): *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*. Handbook of Experimental Economics, ed. by C. A. Plott, and V. Smith, pp. 1047-53.

Lazear, E.P., Malmendier, U. and R.A. Weber (2012): *Sorting in Experiments with Application to Social Preferences*. American Economic Journal: Applied Economics 4(1), 136-63.

Loewenstein, G., O'Donoghue, T. and M. Rabin (2003): *Projection Bias in Predicting Future Utility*. Quarterly Journal of Economics 118(4), 1209-48.

Loewenstein, G., and D. Adler (1995): *A Bias in the Prediction of Tastes.* The Economic Journal 105(431), 929-37.

Loewenstein, G., and P.A. Ubel (2008): *Hedonic Adaptation and the Role of Decision and Experience Utility in Public Policy*. Journal of Public Economics 92, 1795–1810.

Loomes, G., and R. Sugden (1982): *Regret theory: An Alternative Theory of Rational Choice under Uncertainty.* Economic Journal 92(368), 805-24.

Luft, J. (1994): *Bonus and Penalty Incentives Contract Choice by Employees.* Journal of Accounting and Economics 18(2), 181-206.

Madrian, B.C., and D.F. Shea (2001). *The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior*. Quarterly Journal of Economics 116(4), 1149–87.

Mellers, B., Schwartz, A., and I. Ritov (1999): *Emotion-Based Choice.* Journal of Experimental Psychology: General 128(3), 332-45.

McCord, P. (2014): *How Netflix Reinvented HR*. Harvard Business Review.

Musch, J.; and U.-D. Reips (2000): *A Brief History of Web Experimenting, In Psychological Experiments on the Internet*. Ed. by M. Birnbaum, chap. 3, pp. 61-88. Academic Press.

Neri, C. and H. Rommeswinkel (2016): *Freedom, Power and Interference: An Experiment on Decision Rights*. Available at SSRN 2485107.

Nosenzo, D. (2016): *Employee Incentives: Bonuses or Penalties?.* IZA World of Labor.

Odean, T. (1998). *Are Investors Reluctant to Realize Their Losses?* Journal of Finance 53(5), 1775-1798.

O'Donoghue, T. and M. Rabin (1999): *Doing it Now or Later*. American Economic Review 89(1), 103-24.

Owens, D., Grossman, Z. and R. Fackler (2014): *The Control Premium: A Preference for Payoff Autonomy.* American Economic Journal: Microeconomics 6(4), 138–61.

Pope, D. G., and M. E. Schweitzer (2011): *Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition and High Stakes*. American Economic Review 101(1), 129-57.

Rizzo, M.J., and D.G. Whitman (2009): *The Knowledge Problem of New Paternalism*. 2009 Brigham Young Law Review 905–68.

Rommeswinkel, H. (2016): *Measuring Freedom in Games*. Working Paper.

Samuelson, W. and R. Zeckhauser (1988): *Status Quo Bias in Decision Making.* Journal of Risk and Uncertainty 1(1), 7-59.

Schnedler, W. and R. Vadovic (2011): *Legitimacy of Control.* Journal of Economics and Management Strategy 20(4), 985-1009.

Sen, A. (1985): *Well-Being, Agency and Freedom - The Dewey Lectures 1984.* Journal of Philosophy 82(4), 169-221.

Shefrin, H. and M. Statman (1985): *The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence*. Journal of Finance 40(3), 777- 90.

Sunstein, C. (2017). *Human Agency and Behavioral Economics: Nudging Fast and Slow*. Palgrave Advances in Behavioral Economics.

Thaler, R.H., and S. Benartzi (2004): *Save More Tomorrow—Using Behavioral Economics to Increase Employee Savings*. Journal of Political Economy 112(1), S164–S187.

Waldron, J. (2016): Self-Application. Unpublished Working Paper.

Weber, M. (1946): *From Max Weber: Essays in Sociology*. Oxford University Press, New York.

Zeelenberg, M., Beattie, J., Van der Pligt, J., and N. K. de Vries (1996): *Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making*. Organizational Behavior and Human Decision Processes 65(2), 148-58.

Zeelenberg, M. (1999): *Anticipated Regret, Expected Feedback and Behavioral Decision Making*. Journal of Behavioral Decision Making 12(2), 93-106.

Zuckerman, M.; Porac, J.; Lathin, D.; Smith, R. and E.L. Deci (1978): *On the Importance of Self-Determination for Intrinsically Motivated Behaviour*. Personality and Social Psychology Bulletin 4(3), 443–46.

APPENDIX

I. In the following we provide additional information on the design and methods, some more data and applications that can add to the understanding of our project, but that we did not want to distract focus and flow of the main paper. We provide a full set of experimental instructions in a separate appendix XX. **Methods**

We conducted our study online using the Limeservice platform and the open-source software Limesurvey (see http://www.limesurvey.org/). Implementing the experiment online allows us to impose real opportunity costs on subjects and reduce participation costs that are not task-related. This is particularly important for measuring acceptance rates (see Section IV, supra). We invited the subjects via email to the experiment, drawing from a large subject pool (approx. 1,500 subjects) that we established at the University of Münster. A link in the email leads the subjects directly to the study. After participants log in, the link becomes inactive, ensuring that the same participant can complete the study only once. The invitation informs the participants about the time a subject typically requires to complete the experiment, thereby ensuring that subjects do not discontinue participation because it is taking them longer to finish the study than they expected. The e-mail message did not describe the purpose of the study in order to avoid participants self-selecting into an experiment in which they were interested.

To ensure that subjects focus on the experimental task we implement a time limit. If participants fail to make an input after 180 seconds or if they log out, we exclude them from the experiment without payment. The conditional payment and the easy-to-handle real effort task result in a low drop-out rate of ~10% of participants who began the study.

For calibrating the payoffs, we wanted to assure that incentives are comparable to a normal working environment in a regular student job in and outside of the university. A student job in Münster would offer approximately €8 per hour. In our real-effort task, participants needed less than 1 minute to complete a table. Reading the instructions took around 5 minutes, such that the subjects earned an hourly wage of €55. However, we randomly selected every fifth participant for payment. In expectation subjects therefore earned €11 an hour, slightly more than in a student job. For the random selection we presented all subjects with pictures of 5 buildings with entry-way numbers obscured; the subjects had to pick one building and if that building had the house-number 5, the subject was selected for payment. Stochastic payouts like the one we implemented are routinely used in experimental economics. Evidence suggests that larger amounts paid at a lower probability can simulate higher stakes decisions: that is, while actual payoffs are equivalent the incentives are more effective than if smaller amounts are paid with higher probability (Laury 2006; Laury and Holt 2008).

In the *Preference Matching* treatment, we compare the subjects who received their preferred contract with subjects who did not. Since we have three contracts, the likelihood that subjects receive the contract they prefer would only be 1/3. To generate a similar number of observations for both comparison groups, we implement a two-step protocol: We first present the subjects with pictures of the two buildings whose numbers do not match their preference and subjects must select one. Once they made their decision, we replace the photo the subjects choose with the building in the second set that has the same number. We show this picture together with the building that does have the number that matches the threshold the choosing subject prefers. The participants pick again. We reveal the street numbers and the participants are assigned the contract threshold corresponding with the number of the building they selected in the second round. For example, a subject may have a preference for a Stretching contract. Then she is presented with the buildings that have the numbers 5 and 50 in the first round. If she selects the building with the number 50, she is presented with a new building that has the number 50 and a second building that has her preferred number 15 in the second round. When she now selects the building with the 15 she is assigned her preferred Stretching contract, otherwise she is assigned the Extreme Effort contract.

We did not induce contract acceptance, in order to measure whether the use of loss-framing as a commitment device and granting autonomy would increase subjects' willingness to enter a loss-framed contract. In the regression analysis of our data we used a dummy variable to control for contract acceptance. For the non-parametric tests we analyzed the data with the contract rejections included and excluded. When including the rejections, we assumed that subjects who did not enter a contract exerted zero effort. Note that both sets of results – i.e., the one in which we account for rejections, and the other in which we exclude them from the sample -- reflect a situation principals care about. If the supply of workers is unlimited, a principal may consider only how agents perform under the offered contract and is indifferent to the possibility that the contract terms may deter workers from acceptance. Considering only subjects in our analysis who accept the terms reflects this situation. On the other hand, the supply of qualified workers might be scarce, such that unattractive terms may severely harm production. We capture this scenario by including all participants in our analysis.

II. Results

A. Evidence for Norm-framing

We designed a *Random Threshold* treatment to cleanly disentangle the effects of norm and loss-framing and to analyze how they cause a U-shaped effort function. A threshold can only communicate what effort choice the principal expects the workers to make, if workers perceive the threshold as being set with purpose. Therefore, offering a randomly selected threshold should mute the influence of norm-framing and make subjects less likely to conform to the threshold. We therefore expected subjects to invest more effort in the *Random Threshold* treatment in comparison to the *Low Bar* treatment.

As we expected, with the norm-effect switched off subjects invest significantly more effort when they are offered a randomly selected *Low Bar* contract, compared to the regular

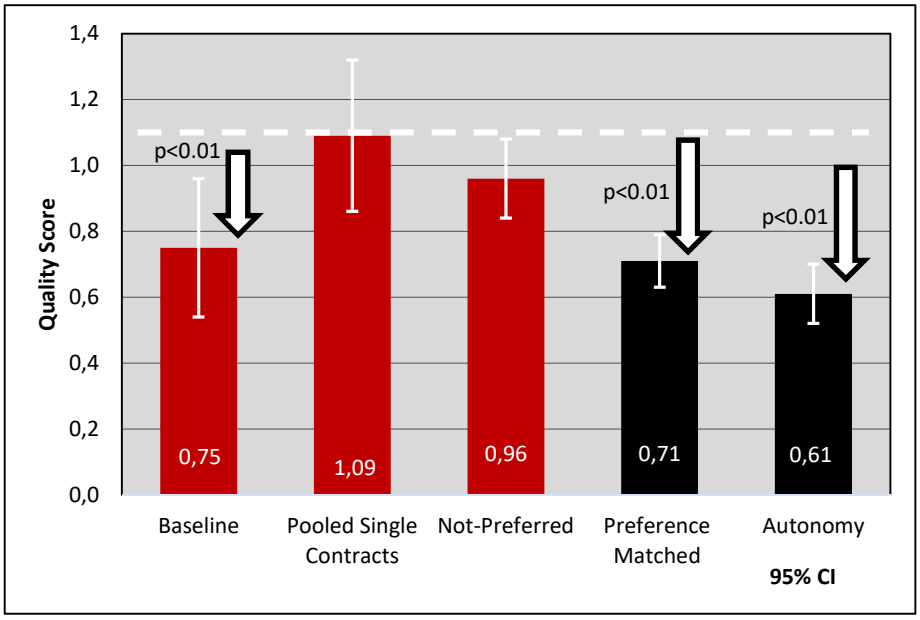
Low Bar treatment (11.1 vs. 6.2 tables; Mann-Whitney *p*-value <0.01). The effect holds when we include the subjects who do not enter the contract. For additional evidence that *the Random Threshold* treatment mutes norm framing we compare how often subjects conform to the threshold under the randomly assigned *Low Bar* contract versus the regular *Low Bar* treatment. As expected, in the regular *Low Bar* treatment subjects conform significantly more often to the threshold: 16 of 50 subjects exactly match the norm of 5 tables, while in the *Random Threshold* treatment only 2 out of 44 subjects match the low threshold (Fisher test *p*-value <0.01). Finally, consistent with choice muting norm-framing, but not loss-framing, loss aversion gains stronger influence on subjects' productivity in the *Random Threshold* treatment. We compare the performance of loss averse types (i.e., those subjects who reject at least one of the two lotteries) with the productivity of rational or only weakly-loss-averse types (i.e., those subjects who participate in both lotteries). Loss averse types complete a mean 5.3 tables more under the *Random Threshold* versus under the regular *Low Bar* treatment; rational types in contrast increase their effort only by 2.2 tables. To estimate the difference, we conduct an OLS regression with a dummy variable distinguishing between the two treatments and the interaction term Loss-Aversion*Treatment estimating how the impact of loss aversion on effort varies by treatment. The interaction is significant (reg beta=-3.37 *p*-value 0.05), supporting our hypothesis that loss aversion should exert a stronger influence on subjects' performance once the norm-effect is muted.

B. **Choice and Autonomy Increase Performance Quality**

Our study design allows us also to measure performance quality. We estimate quality by how much subjects' entries deviate from the true value (i.e., the true number of appearances of the specified digit in the given table). We consider failed and successful attempts; successful attempts can still vary by quality, as an input within a range of +/-2 of the true value moves subjects forward to the next table. We calculate deviations per table completed successfully and refer to the outcome as a "quality score". The software did not enable us to record the exact entry for failed inputs, so we assume in our analysis the minimal deviation of 3 from the true value for all failed trials. Our design makes it very difficult for subjects to deliberately trade off quantity and quality. As the quality threshold required for moving to the next table permits only for a small margin of error subjects are discouraged from attempting to shirk on quality to reach a higher quantity result. This allows us to attribute changes in performance quality to subjects' motivation and attention to the real effort task.

The single contract treatments *Low Bar, Stretching and Extreme Effort,* are less likely to match the subjects` individual preferences and production functions with the thresholds they offer. A poor matchas in the the low and the extremely high threshold treatments , leads to a low quality performance. In both the *Low Bar* (1.23) and *Extreme Effort* (1.45) treatments, performance quality is significantly worse than under the *Stretching* (0.68) and the *Linear* contract (0.75) in *Baseline* (for all four comparisons Mann-Whitney *p*-value <0.01).

Figure 6: Mean Quality Score Across Treatments



Choice and autonomy effectively increase performance quality: In *the Autonomy* treatment we find with a mean score of 0.61 a significantly better-quality performance than in the pooled single-threshold treatments (1.08; Mann-Whitney *p*-value <0.01). We perform an OLS regression s with dummy variables for contract type and acceptance , another dummy variable estimates the difference between the *Autonomy* and the single threshold treatments we are interested in. The difference is clearly significant (reg beta=-4.68 *p*-value <0.01).

, Performance quality is higher for all three contract types in the *Autonomy* treatment and it barely varies across the three contract types -- consistent with a better match of preferences with contract thresholds when subjects can choose their preferred contract. We find for subjects who chose the *Low Bar* contract a quality score of 0.70, for those who picked a *Stretching* contract 0.53, for those undertaking the *Extreme Effort* contract 0.59. The deviation between the three threshold contracts is clearly insignificant (Kruskal-Wallis *p*-value 0.91). Even though the treatment is different in some respects, *Self-Commitment* allows us also to approximate a quality score for the *Linear* contract, when subjects are granted choice. The score is very similar to the threshold contracts: 0.70.

Finally, the *Preference Matching* treatment provides direct evidence that a poor mapping of production functions and contract thresholds demotivates subjects and lowers performance quality: participants who are assigned the contract they prefer, reach a significantly higher quality score (0.51), compared to the subjects who perform under a contract they would not have chosen (0.98; Mann-Whitney *p*-value <0.01). We confirm the result with an OLS regression with dummy variables for contract type, and one that distinguishes whether subjects received their preferred contract or not estimating the effect of preference matching we are interested in. The effect is highly significant (reg beta=-0.46 *p*-value <0.01).

Importantly, we can conclude from our data that the strong increase in productivity that choice and autonomy drive in our study is not achieved at the expense of performance quality. Indeed, the data suggests the opposite effect: the provision of choice and autonomy appear to motivate both higher-quantity and quality effort.

III. Implications

A. Internal and External Validity

Our study differs in some respects from real-world contracting environments. In particular, in our study the contract is the only source that informs the subjects about their work; they have no further contact with their employer or with other workers. We made this experimental design choice for methodological reasons: Eliminating communication and context outside of the contract allows us to cleanly isolate the effects that the changes of the contractual terms have on work performance and to identify the mechanisms that cause these effects.

In real work settings the factors that drive work motivation will vary across types of jobs, departments or branches, and individuals. Some individuals may only work for the money, some will be more intrinsically motivated by the enjoyment of their tasks, some may have strong internal standards or will want to adhere to norms of their profession. Our task does not induce intrinsic motivation for the reason to cleanly separate the autonomy premium. It seems plausible that the demand for self-commitment declines the more intrinsic a worker's motivation. Thus, in environments where intrinsic motivation plays a dominant role like in arts or sciences self-commitment might be less effective to impove performance (may be though for finishing on time). On the other hand, the autonomy premium is likely more readily experienced with tasks that are driven by intrinsic motivation as we explained in the main text.

B. Policy Implications and Directions for Future Research

An additional example for a public policy application is waste management. Many countries try to reduce waste production, but regulation that relies solely on monetary incentives faces a tradeoff: on the one hand increasing prices for waste disposal would incentivize industry and households to reduce their waste production, on the other hand, when prices are pushed too high, the rate of illegal waste disposal increases dramatically (Engel 2002; Frenz 1996). Therefore, regulation often compromises, providing only moderate financial incentives for waste reduction.

A behavioral design that uses our findings could help attenuate this conflict of objectives at no extra cost. Communities can design a loss-framed schedule for the fees they charge. Households would pay an initial flat-tax. They would then choose their preferred disposal plan. The plan entitles them to back-payments conditioned on the amount of waste they want to be disposed. If a household does not comply with the chosen plan, they lose this payment. The back-payment at stake elicits loss aversion and increases the cost of not meeting the waste management goal. The design has two main advantages: it should generate an autonomy effect, elevating people`s intrinsic motivation to meet the target (or at least reduce resistance against the regulation), and second, communities might be able to offer more ambitious plans without risking the side-effect of increasing illegal waste disposure, as people are more willing to accept demanding thresholds when they are given a choice.

We could further extend this list of applications. Car companies must meet emission and fuel consumption goals across the car fleet they sell (see EU regulation No 443/2009 and the US CAFÉ standards). Instead of paying fines when they fail to reach the norm, the companies could pay an initial tax and choose an emission target within limits prescribed.

Depending on the target they choose, the companies would receive a tax bonus. If they fall short of their self-selected goal, they lose (percentages of) the bonus.

The last example raises the interesting question whether our findings may apply to corporate actors. Brokers acting as agents for large investment firms can be motivated by loss aversion (Odean 1998; Shefrin and Statment 1985); they often receive bonuses tied to their company's profit or they may be compensated with shares. A slightly different take is that corporate agents might might engage in defensive decision making and therefore act in their jobs as if they were loss-averse (see Gigerenzer 2014): realized losses may be more easily perceived as a management error than potential gains that were not pursued.

Would agents experience discretion that regulators or contractual partners grant their company as "autonomy," comparable with what an individual acting on her own behalf might experience as self-determination? Perhaps corporate actors experience autonomy if they have an affective relationship to their corporation. Consider traditional firms in Asia, where successive generations of families often work for the same employer and enter the company at a young age with the company also paying for their education.

It could also be interesting to analyze why some firms like Google, Netflix and Dropbox introduced policies of strong autonomy, while other firms did not. Traditional law firms for example often impose extreme thresholds of billable hours on their attorneys.

---

[1] The decision times we measured confirm that subjects are not guessing or finding a solution by trial and error.

[2] In the Koszegi-Rabin framework individuals' take their own expectations as reference points. The Koszegi-Rabin framework can explain our results equally well assuming that subjects refer to the thresholds to form their expectations.

[3] Choosing the threshold closest to their optimal effort absent the threshold must not optimize subjects` performance; as gain-loss utility is asymmetric, choosing a deeper loss frame can further improve net outcome, even when the reference point is more distant to optimal effort absent the threshold.

[4] Selection effects work against our hypothesis: While in the *Autonomy* treatment subjects who aim for a lower performance can be expected to choose the *Low Bar* contract, a random sample of low as well as high performers is assigned to the *Low Bar* treatment.

[5] When we assume that subjects who reject the agreement exert zero effort, we find the same U-shaped result with a performance of 5.0 tables for *Low Bar*, 12.1 for *Stretching*, and 5.2 tables for *Extreme Effort*. All test results remain significant (Mann-Whitney p-value <0.01).

[6] We can also compare the number of subjects who chose a threshold contract to a rational choice strategy which would strictly prefer a *Linear* contract. This comparison is also significant (Fisher test *p*-value <0.01).

[7] As the price for choosing a threshold contract is lower than the 1 Euro that subjects earn for completing an extra table, subjects should always select the threshold they expect will lead them to the best outcome.

[8] Subjects are equally compliant with their commitment in the other *Choice* treatments: In *Open Choice,* 89.5% comply with their commitment and in *Autonomy* 89.9% of the subjects who either choose a *Stretching* or an *Extreme Effort* over a *Low Bar* contract comply (for both comparisons Fisher test *p*-value <0.01).

[9] We find a very similar result when we include the subjects who rejected the contract: 18.2 vs. 8.4 tables (Mann-Whitney *p*-value <0.01).

[10] We find a similar result in the *Autonomy* treatment: Loss averse participants reach with 26.8 tables a significantly higher performance (the rational or marginally loss averse types reach 21.9 tables; Mann-Whitney *p*-value 0.04). The effect holds in a regression that controls for contract types.

[11] When we compare subjects who have a preference for the same threshold contract we find expectedly strong effects for *Stretching* and *Extreme Effort* contracts: The subjects (*N*=40) who prefer and received an *Extreme Effort* contract complete 29.4 tables, while subjects (*N*=52) with the same preference, but who were assigned a *Stretching* or a *Low Bar* contract reach only 12.7 tables. Subjects who favored and received the *Stretching* contract they preferred (*N*=53) complete 15.9 tables, while the other subjects (*N*=49) reach only 9.9 tables. For

the *Low Bar* contract, we find a value of 7.9 tables for the subjects who received the *Low Bar* contract they had favored (*N*=60) vs. 7.9 tables for the subjects (*N*=46) who did not.

[12] A principal who imposes thresholds may of course try to adjust them to what she thinks the agents' production function is. The autonomy premium however, is tied to offering contract choice.

[13] The comparisons by contract type show a tendency in the predicted direction: For the *Extreme Effort* we find 30.9 tables in the *Preference Matching* treatment vs. 8.4 tables in the single threshold treatment, for the *Stretching* contract we find 18.5 tables vs. 14.3 tables, and for *Low Bar* 11.5 vs. 6.2 tables.

[14] Mean performance above the threshold is similar (Autonomy: 9.9 tables versus Preference Matching: 8.2 tables; Mann-Whitney *p*-value 0.54).

[15] The difference is even larger when we consider the subjects who reject the contract: *Autonomy* 67.7% vs. 24.2% in the pooled single threshold treatments (Fisher test *p*-value <0.01).

[16] In our study, we adjusted the payment to typical student jobs. If we would have paid either one cent or $10 per table completed, subjects would likely not have been sensible to our manipulation of the contract terms.

[17] Note this survey data is based on companies' self-report and may thus include self-promotion.

[18] Bubb et al. seem to suggest that preserving choice is a philosophical pre-commitment of Sunstein and Thaler that does not itself have a behavioral effect.

[19] What regulation is optimal remains an empirical question; i.e., whether loss aversion is a sufficiently strong commitment device to counter consumption temptations or whether only mandated savings plans can reach higher savings rates (see also Bubb and Pildes 2015; Thaler and Benartzi 2004).

[20] An example is reported by Casscells et al. (1978): In their study, medical professionals from Harvard Medical School were given a fairly simple diagnostic problem presented in terms of conditional probabilities: "If a test to detect a disease whose prevalence is 0.1% has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease…?" A vast majority of the participating doctors neglected the base rate – that is the low prevalence of the disease in the overall population - their most common answer being 95%. Only 18% of the professionals inferred the correct value. By contrast a presentation of the same problem in the format of natural frequencies preserves the base rate information and makes the correct inference obvious: 1 in 1000 persons have the disease. Of these 1000 persons, 50 get a positive test even though they are not ill. Now the answer is salient: the chance that the positively tested patient has the disease is a mere 1 in 50. Apparently, treatment decisions will be hugely affected when doctors and patients assume the probability to be 95% instead. An application to legal trials is DNA testing. Consider a jury is presented DNA evidence either in the format of conditional probabilities or based on natural frequencies (Hoffrage et al.2000; Koehler 1996).

[21] The suggestion of a majoring rule which directs the nudge such that it pushes in the direction preferred by most people - as suggested by Sunstein and Thaler (2008) - overlooks the epistemic problem, that the bias may mask what the majority actually wants.