I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12253

Advanced Counter-Biasing

Jeffrey Carpenter
Kevin Benscheidt

MARCH 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12253

Advanced Counter-Biasing

Jeffrey Carpenter
Middlebury College and IZA

Kevin Benscheidt
Middlebury College

MARCH 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Advanced Counter-Biasing

Viewed through the lens of the prominent two-system model of decision making, behavioral economics is seen as studying the tension between impulses (System 1) and rationality (System 2). In this context, two strategies, ”de-biasing” informing agents of their biases and “counter-biasing” using “nudges” to activate biases to positively influence choice, have improved the welfare of behavioral agents. We advance the notion of counter-biasing by demonstrating that one bias (present bias) can be pit against another (choking at high stakes) to counteract the ill effects of the second. Our results demonstrate the potential of counter-biasing as an effective policy tool.

JEL Classification: C91, D91, M52

Keywords: behavioral economics, counter-biasing, present bias, choking, experiment

Corresponding author:
Jeffrey Carpenter
Department of Economics
Middlebury College
Middlebury, VT 05753
USA

E-mail: jpc@middlebury.edu

1 Introduction

Behavioral economics was founded by economists and psychologists who studied what, at first blush, appeared to be decision-making anomalies. As these researchers dug deeper, however, they gained an appreciation for the costs associated with good decision-making and the shortcuts and rules of thumb that, no doubt, evolved to economize on cognition. What arose was a two-system model of decision making in which only a minority of choices are made purely rationally (via System 2). Indeed, we now understand that many more, perhaps biased, choices come about as the result of affective, gut reactions (i.e., System 1) or as some combination of impulse and reason. In this choice environment, what is the policy maker to do?

One option is to try to inform decision-makers or "de-bias" their choices. As demonstrated in Sunstein (2013), when agents "misfear," that is irrationally fear inconsequential risks while not fearing significant ones enough, it may help to de-bias these deliberations using cost-benefit analysis, adding structure and forcing agents to think harder about the risks and their consequences. Similarly, Frydman and Rangel (2014) show that reducing the saliency of trading prices on the financial statements of experimental traders who pay too much attention to them can attenuate the disposition effect - differentially selling assets that are doing well compared to those doing poorly.

Another option, the one closest to our focus, is to activate a System 1 bias to "nudge" people into making better choices, what has come to be known as "counter-biasing" (Thaler and Sunstein, 2008; Milkman et al, 2009). Though nascent, this literature has already yielded important lessons concerning, for instance, resetting default options as a way of leveraging status quo bias (e.g., Madrian and Shea, 2001), framing incentives as potential decrements instead of increments to leverage loss aversion (e.g., Levitt et al., 2016) and taking advantage of present bias to make opportunities to save in the future more attractive than those in the present (e.g., Thaler and Benartzi, 2004).

Our main contribution is to take the notion of counter-biasing one step further by pitting one bias against another in an attempt to cancel the ill effects of the second. The setting for our study is the recent finding that, contrary to standard theory, agent performance may not rise monotonically with financial incentives. In other words, at high stakes agents often "choke," performing worse than they did at low stakes (Baumeister, 1984; Ariely et al., 2009). In a nut shell, the problem is thought to be one of divided attention: when there is nothing at stake, individuals focus, relying on muscle memory to perform the task but at high stakes the same individuals spend too much energy analyzing the minutia of the task and on task-irrelevant thoughts and worries (Decaro et al., 2011). In the end, as the stakes continue to rise, the allocation of attention (i.e., working memory) for "chokers" becomes increasingly biased towards unproductive pondering and this harms performance.

We examine, both theoretically and experimentally, whether an intervention designed to change one's perception of the stakes can reduce the amount of unproductive second-guessing of chokers and enhance their performance. The counter-biasing intuition for the intervention is straightforward. If chokers are intimidated by the stakes, they should perform better if we can reduce their perception of the prize and, therefore, the pressure that inhibits their performance. One way to reduce the perceived stakes for a non-trivial portion of the population is to delay their payment. If individuals perform well when the stakes are V_{low} and poorly when they are V_{high}, for example, then there must be some $0 < \beta < 1$ such that, when discounted because of the delay, $\beta V_{high} = V_{low}$. In principle, delaying the payment of a prize should enhance performance, at least for a subset of the population of chokers - those that are impatient because of a high discount rate or because they are present-biased (Laibson, 1997).

It is important to note that, in addition to its anticipated efficacy, our delayed payment intervention has other attractive attributes. First, the intervention is (relatively) easy to implement for principals and simple to understand for agents. Second, the intervention will only affect one's perception of the stakes, not their actual value. In other words, for a short delay, there is no real cost imposed on decision-makers. Lastly, the intervention, as we will see below, is unlikely to have any adverse effects on the performance of other individuals. Those, for instance, who are unlikely to choke under any circumstances perform just was well regardless of their conception of the stakes. Further, patient individuals will not be bothered by the payment delay.

Our experimental results confirm what the conceptual framework in the next section predicts. We first replicate the results of earlier studies with a relatively large sample, finding a considerable amount of choking among our participants. Nearly half of our participants perform better at low stakes than when we increase the stakes by an order of magnitude. Considering our intervention in the population as a whole, those participants who were paid their winnings two months from the date of the experiment instead of immediately after were between 9 and 14 percentage points less likely to choke (depending on the definition of choking) and earned 15 percentage points more. Going one step further, using individual-level estimates of time preferences gathered from our participants allows us to calculate the welfare implications of delayed payments. Here we compute that no fewer than half, and as many as two-thirds, of our participants can expect to be better off with the delay.

When we focus on those people at whom the intervention is targeted (as opposed to the general participant population), the welfare benefits increase such that three-quarters of our participants are better off. Specifically, we find that delaying the payments for present-biased individuals helps their performance disproportionately. Comparing the difference in the effect of the intervention on present-biased and non-present-biased participants, we find that delaying payments reduces choking among the present-biased participants between 21 and 39 percentage points more than among the others participants. The difference-in-difference for earnings is also substantial. The intervention increases the earnings of the present-biased participants 30 percentage points more than it does for the non-present-biased participants. In fact, also consistent with our predictions, delaying payments had, essentially, no impact on the performance or the earnings of non-present-biased individuals.

2 Conceptual framework

The dual-processing theory of cognition posits that choice is the result of one system based on the associative heuristics developed over a lifetime of repeated exposure to coincident events (System 1) and another, more rigid, based on symbolic reasoning and explicit knowledge to structure processing (System 2). The two systems also differ in their consumption of cognitive resources. While System 2 cognition is costly and consumes a lot of working memory, System 1 is much more spontaneous, costing little working memory. Considering choices, dual-processing hypothesizes that System 1 is quick to offer input that is then mediated by System 2. Biased and incorrect choices are often the result of too little input from System 2, especially when the decision is more complex than first envisioned (Kahneman, 2011).

In the context of choking when the stakes are high, this dual-processing theory predicts an increase in competition for working memory that reduces the influence of System 2 on the outcome (Beilock and DeCaro, 2007). Specifically, distraction-based theories propose that the stakes can divert attention toward thoughts that are irrelevant to the task, like the social consequences of failure (Beilock et al., 2004; Markman et al.,

2006) and self-monitoring theories suggest that increased stakes and pressure can cause performers to think even harder about their technique and skill execution. The bottom line, according to DeCaro et al. (2011), is that in many high stake situations both these pressures may be at work, "simultaneously disrupting working memory availability and directing what attention that remains in ways that are counterproductive." As an example, and foreshadowing the experiment described below, consider checking to see if two numbers add up to 10. When the stakes are low and working memory is readily available, the problem $6.34 + 4.56$ seems easy but when the stakes increase enough to be distracting and you are redirecting working memory by worrying about your speed instead of your technique, you might be more apt to focus quickly on the last two digits and conclude the sum is 10, not 11.

To dovetail with DeCaro et al. (2011)'s dual-processing theory of choking and derive explicit hypotheses about the effect of payment timing on this propensity, we extend the model developed by Sanders and Walia (2012). In the resulting setup, the psychological pressure of high stakes consumes working memory and affects the probability of achieving a performance goal both directly (through self-monitoring) and indirectly via the cost of providing effort (because individuals become distracted). Beginning with this probability, suppose an individual's likelihood of successfully completing a task depends on how hard she works, the stakes involved and the psychological pressure experienced, which can be thought of as a proxy for the amount of working memory consumed. A natural formulation for this probability, $0 \leq s_i \leq 1$, is

$$s_i = \frac{e_i}{e_i + \mu_i + p_i V}$$

where e_i is the productive effort exerted by individual i, $\mu_i > 0$ is a random component of the production process (i.e., luck), $0 \leq p_i \leq 1$ is the psychological pressure felt by the individual and $V > 0$ is the value of the monetary reward given to individuals who complete the task. For simplicity, we characterize individuals for whom $p_i = 0$ as "non-chokers," those individuals whose performance is unaffected by pressure. For the other individuals, as p_i increases, they become more sensitive to the stakes because they attend too closely to the minutia of the task which disrupts their efforts and are, therefore, less likely to complete the task at any stakes (i.e., they "choke"). As a result, we see that the probability of success decreases in the stakes, but only for chokers.

Psychological pressure is also hypothesized to affect the cost of providing productive effort. In this case the rationale is that one's attention is divided and instead of spending more of it on executing the task, chokers become preoccupied by other, unproductive thoughts. Specifically, where $c_i(e_i, p_i, V)$ is this cost, in general, we assume that the marginal cost of effort increases for those feeling pressure (i.e., $\frac{d^2 c_i}{de_i dp_i} > 0$) because they can become distracted by the size of the prize (as in Beilock and Carr, 2001) such that increasing the stakes will also increase the marginal cost of effort (i.e., $\frac{d^2 c_i}{de_i dV} > 0$). Putting this altogether, the expected benefit of working is just $s_i V$and so the individual's expected utility can be written as,

$$EU_i = \frac{e_i V}{e_i + \mu_i + p_i V} - c_i(e_i, p_i, V)$$

and $e_i^*(\mu_i, p_i, V)$, the optimally chosen level of effort, will typically be positive given the marginal benefit is decreasing but always positive and the marginal cost increases.

Because it is often easier to observe who succeeds in accomplishing a task than how hard someone is trying, and with a nod to what we measure in our experiment, the more important derivation might be how likely are the different types of individuals to complete the task at $e_i^*(\mu_i, p_i, V)$ and does this rate of success vary with the stakes? To examine this prediction, we simply substitute $e_i^*(\mu_i, p_i, V)$ back into s_i and evaluate the comparative static $\frac{ds^*}{dV}$. As Sanders and Walia (2012) note (in Proposition 1), this derivative is ambiguous and if the effects of psychological pressure are large enough (i.e., p_i is large enough), increasing the incentives can harm performance.

Our conjecture, however, is that by delaying the receipt of the prize, we can affect the perception of the stakes in a way that will target a nontrivial subset of chokers, helping them, without harming non-chokers. Further, the intervention is on the "lite" side of paternalism because we affect just the perception of the stakes, not the actual monetary reward. With this in mind, suppose individuals discount future payments by a factor $0 \leq \beta_i \leq 1$, then expected utility becomes

$$EU_i = \frac{e_i \beta_i V}{e_i + \mu_i + p_i \beta_i V} - c_i(e_i, p_i, \beta_i, V).$$

As one can immediately discern, the effect of delaying the payment of the prize to winners will depend not only on how patient the individual is, but also on whether or not the stakes inflict pressure on the individual. For those unaffected by the stakes (i.e.,

$p_i = 0$), delaying the payment of the prize a short while can only affect those individuals that are impatient. For this group, delaying the payment will reduce the incentive to work but because effort appears in both the numerator and the denominator of the probability of success, this affect will be muted if one only observes whether individuals succeed or not. By contrast, for potential chokers, delaying payments has three effects: the one just mentioned, an effect on the chances of succeeding and an effect on the cost of effort, the later two will help chokers, but mostly if they are impatient.

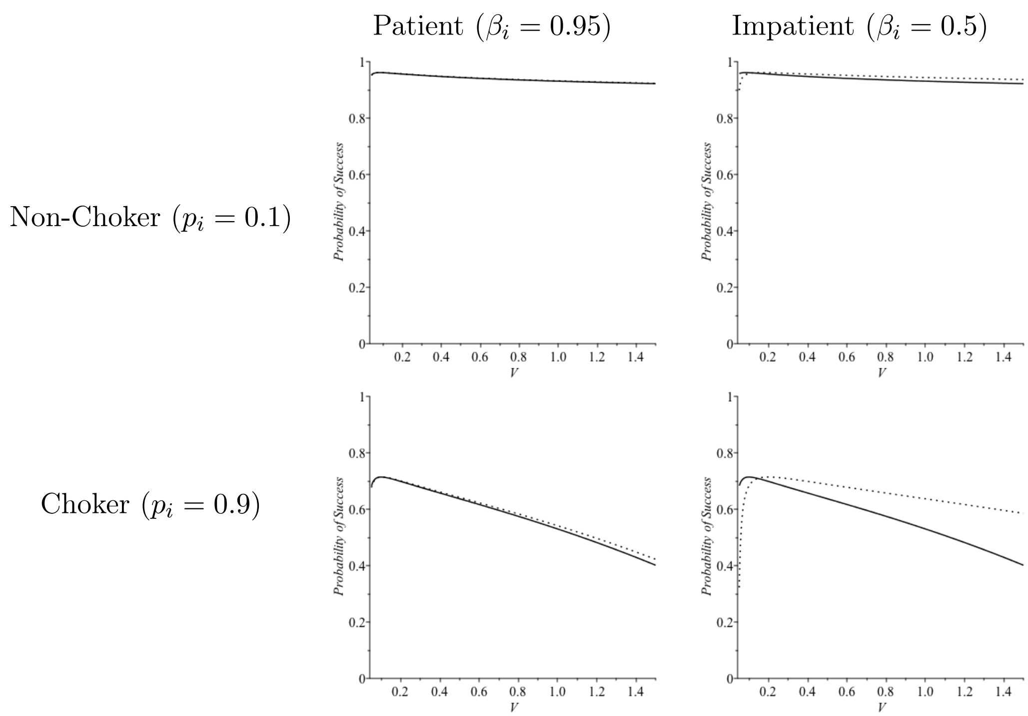

Figure 1: Probabilities of success for four individuals (the combinations of $p_i \in \{0.1, 0.9\}$ and $\beta_i \in \{0.5, 0.95\}$) and their predicted reaction to a payment delay (dotted lines).

Figure 1 graphically summarizes the potential affects of delaying prize payments. As the top row suggests, the effect of the intervention on non-chokers is almost imperceptible.[1] Patient non-chokers continue to succeed at high rates and behave, more or less, the same regardless of when the stakes are paid. Impatient non-chokers are also mostly unaffected but, as one can see, at very low stakes there is a slight reduction in performance because the perceived expected benefit is lower. However, because even these individuals have a small inclination to choke, as the stakes increase they too accrue a little benefit from the delay. Patient chokers in the second row also experience

[1]To present optimal behavior graphically, we assume the simplest formulation that satisfies the conditions above, $c_i(e_i, p_i, V) = cp_iVe_i$ where $c > 0$ is a scaling parameter.

an advantage from the delay but the effect is again small, this time because they are patient. As hoped, however, impatient chokers are helped considerably by the intervention. Though the effect of the delay does reduce the expected perceived benefit from trying hard, as it does for all individuals, this decrement to motivation is more than outweighed by the reduction of self-monitoring and distractions.

In sum, this logic suggests that if a significant proportion of chokers are impatient in the right way, deferring the payment of prizes should significantly improve their performance and reduce the overall incidence of choking substantially and, importantly, this happens without adversely affecting the other individuals.

3 Methods

Our experimental design is both within-subject and between-subject. To assess the propensity to choke, each individual attempts a real effort task at both low and high stakes. To assess the potency of our intervention, half the individuals are paid immediately for succeeding at the high stakes and for the others, this payment is delayed. Linking our study to the related literature, we use the same task implemented in Ariely et al. (2009). In addition to replicating earlier results, one of the benefits of the chosen task is that it allows us to focus on task completion and the probability of success while also allowing us to examine measures of effort and pay, more generally.

This task, known as the "Math Matrix" game, presents participants with a sequence of 4-by-3 matrices of 12 numbers between 0 and 10. For each matrix presented, the participant is asked to identify the two numbers that sum to exactly 10. After a five-matrix practice period is completed, participants are given 4 minutes to correctly solve at least 10 of the up to 20 matrices they are shown.

Our implementation of the Math Matrix game also replicates that of Ariely et al. (2009) in that the monetary prize offered in the high stakes instance is ten times greater than the low stakes game. There are a few ways, however, in which our implementation differed. To begin, Ariely et al. (2009) gathered a modest sample in this game (24 participants). We decided, based on power calculations using their treatment effects, to increase the sample substantially. We gathered data from 199 participants. Instead of a low stakes prize of $15, we chose to use $5. Our high stakes prize was, therefore, $50, however, given the entire experiment lasted just half an hour, the high stakes prize was, substantial. Ariely et al. (2009) displayed a counter to let the participants know how many matrices they had currently solved correctly and offered a piece rate for matrices solved beyond the winning threshold of 10 to incentivize them to continue to work past the target. Our design is simpler. No piece rate was offered but we also did not let participants know how many matrices they had solved until the game had finished. Hence, the resulting uncertainty should have kept them motivated throughout each work period. Lastly, while Ariely et al. (2009) forced all their participants through the same ordering of the tasks - the low stakes instance first then the high stakes one - we randomized the order of the experiment at the session level to make sure the base rate of choking was not driven by fatigue, learning or by becoming accustomed to the task.

To implement our intervention, all our participants were randomized, within sessions, to be paid their high stakes earnings either immediately after the session had concluded or in two months. In both cases, the participants were paid in cash. However, those for whom their earnings were delayed by two months collected their payments from their campus mailboxes.

Because our intervention is designed to target impatient potential chokers, in particular, we also conducted a survey after the experiment to gather information about our participant's time preferences. We implemented the 12-question DEEP protocol developed in Toubia et al. (2012) which generates measures of both a participant's discount rate and present bias. While our implementation of DEEP was unincentivized, the incentives in Toubia et al, (2012) were low (\$2 plus a $\frac{1}{N}$ chance of playing one of the 12 gambles) and our mean responses are very similar to the behavioral parameters they elicited. In addition, the fact that Toubia et al. (2012) tested the external validity of their low incentive results gave us confidence that our measures captured the heterogeneity in our participants time preferences.

Our post-experiment survey also included demographic questions and a few controls that we felt might affect performance. We asked participants to reveal their gender, class and whether or not they worked on campus during the school year. As for the other controls, we asked participants to disclose how competitive they were (in general and in sports) and they completed a 13-question version of the Rotter (1966) scale to measure their locus of control. Lastly, to test that any affects of our intervention are due to time preferences and not to the fact that some participants distrusted that we would actually pay them in two months, we asked a standard question about trust: "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"

Considering our protocol, we acquired a random sample of email addresses from the registrar for recruitment. Potential participants were invited to the behavioral lab on campus for one of fourteen sessions that took place during the winter of 2018. The experiment was programed as a Qualtrics survey and in each session (lasting about half an hour each), participants first read instructions (available in the appendix), gave consent (the protocol was vetted by the Middlebury College IRB), went through a 5-matrix practice round, then two 20-matrix, 4-minute rounds (the order of the stakes being randomized) and then completed the survey. At the end, participants came to a separate room, one at a time, for payment.

4 Results

The observed characteristics of our participants are summarized in the appendix (Table 1A). Overall, about as many participants experienced the low stakes challenge then the high stakes one as the reverse order, despite it being hard to balance this feature because sessions were of different sizes. However, 59% of the participants paid immediately experienced low stakes then high ones and only 46% of those paid later did, a difference that is marginally significant. Fewer women than men participated overall and the women disproportionately participated in the pay later treatment ($p = 0.02$). Though underclassmen were better represented than seniors, we mostly achieved balance across the treatments. The one exception is that more sophomores participated in the pay later treatment ($p = 0.01$). Lastly, only half of our participants worked on campus, a proportion that does not vary by treatment.

Considering the behavioral data we collected from our participants, we constructed an indicator to measure competitiveness based on our two 10-point likert-scored questions, one on general competitiveness and the other on sports. We categorized participants as competitive if they recorded a response greater than 7 on both questions. Our measure of Locus of Control comes from (Rotter, 1966) and we categorized our participants as having an "internal" locus if more than half of their responses were consistent with this perspective. Overall, 43% of our participants are distrusting, a level that does not vary by treatment. Lastly, the results of our implementation of the DEEP time preference module resulted in a mean daily discount rate (Delta) of 0.005 and a mean estimate of present bias, Beta, equal to 1.01, indicating that the average participant was not present-biased. For our analysis we turn these measures into indicators too.[2] We consider participants to have a high discount rate if it is larger than the median and, by definition, participants are "present-biased" if they register a Beta less than 1. Based on this definition, 41% of our participants are present-biased, a fraction that does not vary with treatment (despite the mean value of Beta being a bit lower in the pay later treatment).

Pooling across all 199 of them, our participants solved an average of 7.48 matrices correctly in the 4-minute low stake condition and an average of 8.16 matrices in the high stakes condition. As this indicates, our participants do, as traditional theory would predict, perform better, on average, when the stakes are high (t = 2.52, p = 0.01). However, this difference in average performance does not mean that there isn't a substantial incidence of choking. We consider two definitions of choking, one more strict than the other. In our strictest definition, participants "win" the low stakes challenge (by solving at least 10 matrices) but "lose" the high stakes challenge. Overall, 16% of our participants choked according to this definition. A less restrictive definition that we will also consider is simply performing worse at high stakes (i.e., solving fewer matrices at high stakes than at low stakes). Using this definition, the instances of choking are more prevalent - 42% of participants choke using this definition. While we think that performance is the correct domain in which to measure choking, to compare our results to those of Ariely et al. (2009), we also report the fraction of the total possible pay (i.e., $60) that our participants achieved. Although a third of our participants won the high stakes challenge, only 12% won both challenges, so the average fraction of the total possible amount paid was 0.38, which is close to the high stakes average of 0.43 recorded in Ariely et al. (2009).

4.1 Treatment effects

The framework in Section 2 predicts that delaying payments should reduce the incidence of choking, does it? Figure 2 depicts the treatment effects of delaying payments on the three outcome measures we constructed from our data. Starting on the left of Figure 2 with our most strict measure of choking - winning the low stakes challenge but failing at the high stakes one - we see that, as mentioned above, fewer than 20%

[2]Compared to the data collected by Toubia et al. (2012), our DEEP participants appear a bit more patient and slightly less present-biased. Toubia report a mean discount rate of 0.0121 and a mean Beta of 0.925. One factor that we replicate, however, is the relatively strong correlation between Beta and Delta. In our implementation the correlation is -0.605 compared to -0.654 found by Toubia et al.

of participants choke this way but that the rate of choking in this manner is, indeed, halved when we delay payments. A simple equality of proportions test suggests that this difference is marginally significant ($z = 1.79$, $p = 0.07$). In other words, delaying payments seems to have the intended effect.

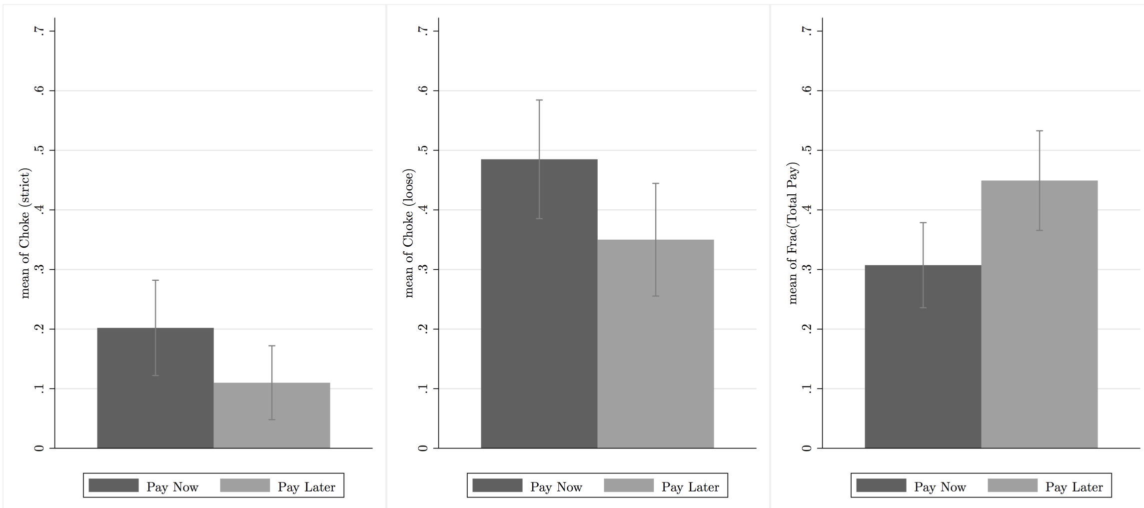

Figure 2: The treatment effect of paying high stakes winners later on the incidence of choking (choking in the left panel is wining at low stakes and failing at high; choking in the middle is solving fewer matrices at high stakes than at low; right panel depicts the fraction of the total pay achieved).

In the center panel of Figure 2, we consider the less restrictive definition of choking: performing worse when the stakes are high, regardless of winning or losing. Being less restrictive, we see an increase in the rate of choking according to this definition. That said, the pattern is the same: delaying the payment of winnings reduces choking. This time the (raw) treatment effect is larger, closer to 14 percentage points and the difference is now significant at conventional levels ($z = 1.93$, $p = 0.05$). In the last panel of Figure 2, we see the implications of reduced choking with delayed payments. Those participants who are paid later earn an average of 45% of the total on offer while those who aren't earn only 31%, on average. This is a considerable difference in earnings, one that is significant at the 1% level ($t = 2.54$, $p = 0.01$).

As a more complete examination of the treatment effects of paying one's winnings later, we evaluate the regression results reported in Table 1. The regressions are linear probability models that control for the observables we collected in the survey and include robust standard errors clustered on the experimental session level. For the interested reader, full tables of all of our regression results, including the point estimates for the controls, are included in the appendix.

Table 1: The treatment effects of delayed payments.

	Choke (strict)	Choke (loose)	Frac(Total Pay)
Pay Later (I)	-0.094	-0.138**	0.154***
	(0.065)	(0.063)	(0.046)
Intercept	0.178	0.532***	0.324***
	(0.103)	(0.150)	(0.140)
Controls included?	Yes	Yes	Yes
Observations	199	199	199
Adj. R^2	0.133	0.033	0.058

Notes: Linear probability estimates; (robust standard errors) clustered on the session; unreported controls include experiment order, gender, graduation year, employment, competitiveness, locus of control, distrust; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Although the experiment is not completely balanced (see appendix Table A1), the observable controls have little effect on the magnitudes of the delayed payment treatment effects. After the inclusion of the controls, the base rate of choking, using the strict standard, is just shy of 18%. This does fall by a bit more than half (9.4 percentage points) for those participants whose payments were delayed. However, this rather large reduction is only significant at the 17% level once we include the controls. Using the looser definition of choking, we continue to find a nearly 14 percentage point reduction in the incidence of choking after controlling for observables, an effect that remains significant at the 5% level. The precision of the point estimate of the effect of payment delays on earnings is sharper in Table 1 ($p = 0.005$), but like with the other two outcome measures, the magnitude remains unchanged after controlling for the observables. In other words, our treatment effects are sizable and robust. Considering the controls themselves, there are only three significant results: sophomores are 10 percentage points less likely ($p = 0.096$), competitive participants are 18 percentage points more likely ($p = 0.013$) to choke under the strict definition and women tend to earn 10 percentage points less of the total possible than men ($p = 0.006$).

4.2 Heterogeneous effects

As the model in Section 2 suggests, the magnitude of the treatment effects in the previous section should depend on the number of "impatient" participants. The more prevalence is impatience, the larger the expected effect. This implies that our estimates (in the previous section) of the raw treatment effects are likely to be restricted by the fact that the estimates are effectively averaged across the rows of Figure 1. In this section, we analyze the effects of delaying payment separately for patient and impatient participants, an analysis more consistent with the specific predictions of the model.

Before we begin, we acknowledge that there are now believed to be two aspects of patience: one's discount rate and whether or not one is present-biased (Labison, 1997). We purposely collected both parameters from our participants so that we can allow the data to tell us if either (or both) are important for the success of our intervention. Considering Beta, Figure 3 suggests heterogeneous treatment effects that are consistent with what we expected at the end of Section 2. On the left, we see that delayed payments do, differentially help present-biased participants. In fact, very few present-biased players choke in the extreme when their pay is delayed. In the middle panel of Figure 3, where we consider the second definition of choking, the result is the same. Paying participants their winnings later doesn't have much of an effect on the "normal" (i.e., non-present-biased) participants, but it seems to help the present-biased ones considerably. On the right, because present-biased players choke so infrequently when paid later, they are among the most successful earners in the experiment and they seem to do significantly better than the present-biased individuals who are paid now.

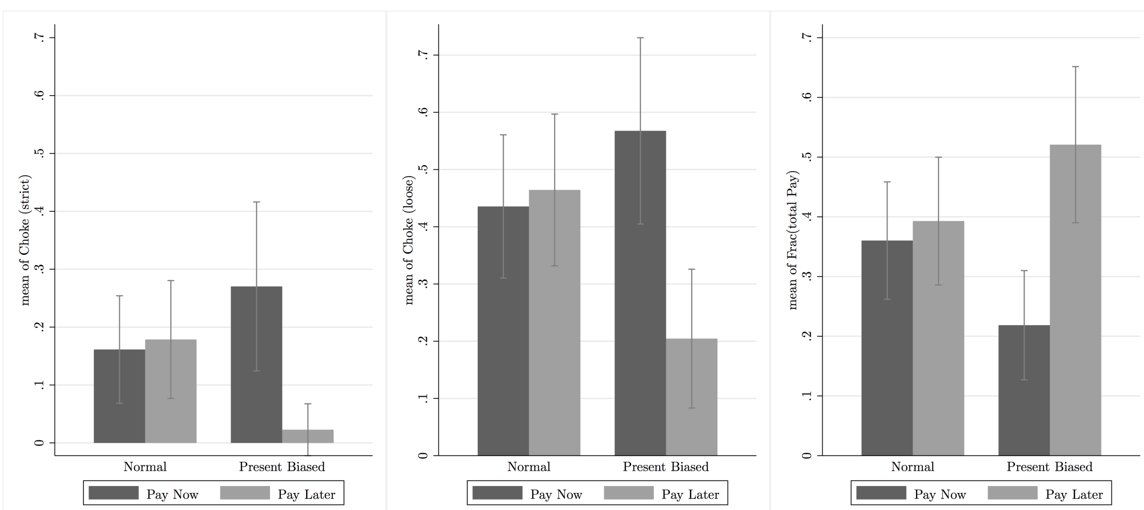

Figure 3: Heterogeneous treatment effects of paying later on present-biased and non-present-biased participants (choking in the left panel is wining at low stakes and failing at high; choking in the middle panel is solving fewer matrices at high stakes than at low; right panel depicts the fraction of the total possible pay achieved).

To get point estimates for these heterogeneous effects, we included a present bias indicator and an interaction with being paid later in the standard regression equation. The results are listed in Table 2. Indeed, paying normal participants later has very little effect on their likelihood of choking, though present-biased players are a bit more likely to choke, overall. However, the difference-in-difference is large and significant at better than the 5% level. Delaying payments reduces the incidence of choking by 21 percentage points more for present-biased individuals than for non-present-biased ones ($p = 0.03$). Our estimate of the difference-in-difference grows when we move to the looser definition of choking in the second column of Table 2. Here present-biased individuals who are paid later are now 39 percentage points less likely to choke than normal individuals who are also paid later ($p = 0.02$). The observed differential reduction of choking for present-biased participants that are paid later translates into a sizable earnings "bump." In the last column of Table 2, we see that these participants earn 30 percentage points more than their colleagues who are not present-biased but are also paid in two months. Taking stock, the results described in Figure 3 and Table 2 closely resemble the predictions from Section 2. Our delayed payment intervention differentially helps present-biased individuals but also notice that, as predicted, delaying payments does not appear to harm the other participants.

Table 2: Heterogeneous treatment effects based on present bias.

	Choke (strict)	Choke (loose)	Frac(Total Pay)
Pay Later (I)	-0.010	0.017	0.037
	(0.055)	(0.075)	(0.061)
Present-Biased (I)	0.073	0.132	-0.149**
	(0.094)	(0.125)	(0.064)
Pay Later × Present-Biased	-0.210**	-0.386**	0.297**
	(0.084)	(0.146)	(0.108)
Intercept	0.139*	0.460***	0.388**
	(0.078)	(0.131)	(0.121)
Controls included?	Yes	Yes	Yes
Observations	199	199	199
Adj. R^2	0.154	0.072	0.089

Notes: Linear probability estimates; (robust standard errors) clustered on the session; unreported controls include experiment order, gender, graduation year, employment, competitiveness, locus of control, distrust; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.3 Robustness

We examine the robustness of our results in three ways ways. First, we test whether the differential effect of the payment delay is better explained by participants' high discount rates than present-biased. Second, we allow the discount rate and present bias to go "head-to-head" to test if present bias continues to account for our heterogenous treatment effects when both aspects of patience are allowed to compete for the available variation at the same time. Third, because we may worry that present bias might just be a proxy for not trusting that we would actually pay participant in two months, we also check to see if our heterogeneous effects can be explained by our surveyed measure of trust. In other words, is our present bias indicator just a proxy for trust?

When we substitute (in appendix Table A4) an indicator for having a relatively high daily discount rate (our cutoff was the median, 0.00403) for being present-biased in our analysis, we see that both high and low discount rate participants choke less often when their pay is delayed and, in fact, the gap seems a bit larger for the low discount rate participants. In other words, the difference-in-difference we are looking for is not sizable and seems to go in the wrong direction. When we include both present bias and the discount rate (appendix Table A5), in some cases controlling for the discount rate actually increases the point estimates on the present-biased difference-in-difference. At the same time, the results for the discount rate are largely unchanged. These results strengthen the case for our intervention working on the present bias aspect of patience. Lastly, when (in appendix Table A6) we substitute our distrust indicator and an interaction with being paid later for the corresponding present bias variables and re-run our regressions, we find that distrusting does a poor job of explaining our results. Hence, we conclude, again, that our intervention improves outcomes mostly because it helps present-biased individuals.

4.4 Welfare

Although we already discussed the fact that, on average, participants earn more when their payments are delayed, we can also examine something closer to the welfare implications of delayed payments by utilizing the individual-level preference estimates of Beta and Delta we gathered. Specifically, to determine whether the delay will make them better off or not, suppose individuals calculate the expected costs and benefits, understanding the benefits will depend on how much more likely one will be to achieve the performance goal if payments are delayed and the costs are proportional to how patient individuals are and how long they have to wait. The expected value of the delay can be written as $EV(delay) = \overline{p}\beta\delta^{60}V + (1-\overline{p})0$, where $\overline{p}$ is the probability of achieving the goal when payments are delayed two months and $\delta = \frac{1}{1+Delta}$. Likewise, the expected value of being paid now is $EV(now) = \underline{p}V + (1-\underline{p})0$. Here $\underline{p}$ is the probability of achieving the goal when payments are not delayed. Individuals expect the delay to be good for them when the difference, $(\overline{p}\beta\delta^{60} - \underline{p})V$, is positive.

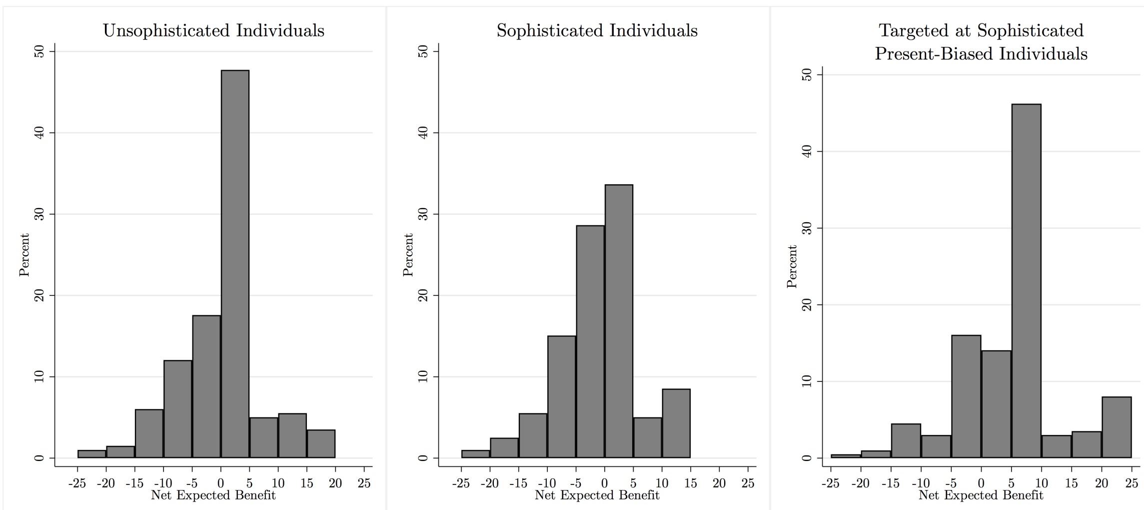

Figure 4: The welfare effects of delayed payments based on individual preference and the chances of performing better at high stakes (left panel shows estimates of unsophisticated individuals who fail to realize that doing better correlates with present bias; middle panel depicts net benefits for sophisticated individuals who account for the effect of present bias on the delay's effect; right panel shows the effect of targeting the intervention at present-biased individuals).

Suppose we assume one of two things about the individuals: either they are unsophisticated in that they ignore the fact that Beta is likely to be correlated with the increased likelihood of achieving the goal when payments are delayed or they are sophisticated and try to control for this fact when calculating the benefits and costs of the delay. Estimating first the net expected benefit of the delay for each participant using their preferences and the average treatment effect of the delay on the probability of doing better when the states are high, we calculate the benefit as we assume unsophisticated individuals will and see, in the left panel of Figure 4, that most people benefit a small amount (in expectation) from the delay. These calculations indicate 62% of participants have positive expected benefits. In the middle of Figure 4, we see that after accounting for the fact that the benefit of the delay will be larger for the more present-biased individuals, the estimates of the net benefit of the delay fall, on average. Now approximately half the individuals (47%) anticipate benefitting from the payment intervention. Lastly, on the right of Figure 4 we graph the distribution of benefits when when the intervention is targeted at just the present-biased individuals and we assume they are sophisticated. Now the benefits of the intervention are positive for 75% of people.

5 Discussion

The purpose of our study is three-fold. Our first, most modest goal is to replicate previous work indicating that, contrary to standard theory, performance need not rise monotonically with incentives. That is, individuals may choke because the stakes begin to loom large, distracting them and hindering their concentration. Our second goal is to devise and test a simple method to reduce the incidence of choking in the population. Our last, and ultimate, goal is to offer one of the first tests of the behavioral idea of advanced counter-biasing. Here we ask if one common bias (present bias, in this case) can help attenuate another (choking at high stakes) and improve the outcomes of, at least some, individuals.

Considering the replicability of previous experimental work on stakes and the incidence of choking, we find that, although our design is not identical to that of Ariely et al. (2009), our results are roughly similar. Specifically, in their study of the Math Matrix game, in which everyone was paid immediately at the end of the experiment, the majority of participants (i.e., 71%) choked, in that they performed worse (i.e., earned a smaller share of the total earnings offered) when the stakes were high. By comparison, considering only those of our participants that were also paid immediately, we find less choking, 49%, but it is still very common. In fact, given the modest sample collected in the Ariely et al. study the difference in the proportions of choking are only just significant at the 5% level. Hence, we are confident that changes in our subject pool and protocol have not fundamentally changed the experiment - ours does seem to be an "apples-to-apples" comparison and we mostly replicate this earlier work.

Other interventions have been suggested in the psychological literature to try to attenuate the anxiety felt by individuals who are apt to choke but these interventions can be complicated and time-consuming. Reeves et al. (2007), for example, show that self-consciousness training can reduce the incidence of missing penalty kicks in soccer. By comparison, our intervention is simple to understand and implement. While the optimal structuring of the payment delays is likely to be an interesting topic for future study, it should in most setting, be simple to delay payments. Indeed, the payoff to the intervention is substantial. Among all our participants, our estimates indicate that delaying payments increases the fraction of the stakes achieved by 15 percentage points. Considering our stakes, this translates into a $9 increase in pay for a half-hour time commitment. Further, restricting attention to our target population, present-biased participants, this pay increment is closer to 33 percentage points or almost $20. In other words, individuals would have to be both very impatient and very control averse to not see the benefit of the intervention.

Although the expected benefits from our intervention are substantial, the results of our study generalize beyond the design of contracts to prevent performance decrements due to strong incentives. In particular, our study is among the first to demonstrate the power of advanced counter-biasing - the use of choice architecture to activate one behavioral bias to counteract another.[3] In this instance, the psychological pressure experienced when the stakes are high is irrational, especially considering that this pressure (and the related self-consciousness) can only distract attention and hinder performance. For once, however, being impatient can be seen as beneficial because it can cause the perception of payments received in the future to fall below the threshold leading to unproductive arousal and poor performance.

Our data suggest that it is the present bias aspect of impatience (not simply having a high discount rate) that interacts best with this counter-biasing manipulation and the implications of this are also important for welfare. The importance of present bias suggests that payment delays need not be long to have a substantial effect on performance, for instance. Not only will shorter delays make implementation less cumbersome and less difficult to account for, the shorter the delay, the less the potential cost of any paternalism because, in principle, any welfare loss from receiving payments later can be minimized and should be outweighed by the increased rate of success.

[3]Other related studies include Kahneman and Lovallo (1993), Besharov (2004) and Bansal and Rosokha (2018).

6 References

Ariely, D., Gneezy, U., Loewenstein, G., Mazar, N., 2009. Large stakes and big mistakes. Review of Economic Studies 76, 451–469.

Bansal, S., Rosokha, Y., 2018. Impact of Compound and Reduced Specifications on Valuation of Projects with Multiple Risks. Decision Analysis 15, 27–46.

Beilock, S., Carr, T., 2001. On the fragility of skilled performance: what governs choking under pressure? Journal of Experimental Psychology 130, 701–725.

Beilock, S., DeCaro, M., 2007. From poor performance to success under stress: working memory, strategy selection, and mathematical problem solving under pressure. Journal of Experimental Psychology: Learning, Memory, and Cognition 33, 983-998.

Beilock, S., Kulp, C., Holt, L., Carr, T., 2004. More on the fragility of performance: choking under pressure in mathematical problem solving. Journal of Experimental Psychology: General 133, 584-600.

Besharov, G., 2004. Second-Best Considerations in Correcting Cognitive Biases. Southern Economic Journal 71, 12–20.

DeCaro, M., Albert, N., Thomas, R., Beilock, S., 2011. Choking under pressure: multiple routes to skill failure. Journal of Experimental Psychology 140, 390–406.

Frydman, C., Rangel, A., 2014. Debiasing the disposition effect by educing the saliency of information about a stock's purchase price. Journal of Economic Behavior and Organization 107, 541–552.

Kahneman, D., 2011. Thinking, Fast and Slow. Farrar, Straus and Giroux, New York.

Kahneman, D., Lovallo, D., 1993. Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking. Management Science 39, 17–31.

Laibson, D., 1997. Golden Eggs and Hyperbolic Discounting. Quarterly Journal of Economics 62, 443–478.

Levitt, S., List, J., Neckermann, S., Sadoff, S., 2016. The behavioralist goes to school: leveraging behavioral economics to improve educational performance. American Economic Journal: Economic Policy 8, 183–219.

Madrian, B., Shea, D., 2001. The power of suggestion: inertia in 401(k) participation and savings behavior. Quarterly Journal of Economics 116, 1149–1187.

Markman, A., Maddox, W., Worthly, D. 2006. Choking and excelling under pressure. Psychological Science 17, 944-948.

Milkman, K., Chugh, D., Bazerman, M., 2009. How can decision making be improved? Perspectives on Psychological Science 4, 379–383.

Reeves, J., Tenenbaum, G., Lidor, R., 2007. Choking in front of the goal: the effects of self-consciousness training. International Journal of Sport and Exercise Psychology 5, 240–254.

Rotter, J.B., 1966. Generalized Expectancies for Internal Versus External Control of Reinforcement. Psychological Monographs; General and Applied 80, 1–28.

Sanders, S., Walia, B., 2012. Shirking and "choking" under incentive-based pressure: a behavioral economic theory of performance production. Economics Letters 116, 363–366.

Sunstein, C., 2013. If misfearing is the problem, is cost-benefit analysis the solution?, in: The Behavioral Foundations of Public Policy. Princeton University Press, Princeton and Oxford, pp. 231–242.

Thaler, R., Benartzi, S., 2004. Save More Tomorrow: Using behavioral economics to increase employee savings. Journal of Political Economy 112, 164–187.

Thaler, R., Sunstein, C.R., 2008. Nudge: improving decisions about health, wealth, and happiness. Yale University Press, New Haven, CT.

Toubia, O., Johnson, E., Evgeniou, T., Delquie, P., 2012. Dynamic experiments for estimating preferences: an adaptive method of eliciting time and risk parameters. Management Science 59, 613–640.

7 Appendix

7.1 Experimental instructions

1. Thank you for participating today. This experimental session should take about fifteen minutes to complete and it will consist of two parts: a task, for which you can earn money, and a survey. The experiment will consist of a few repetitions of the task and this will be followed by the survey. Any money you earn during the experiment will be paid in cash. If you click the button below we will get started.

2. Middlebury College Economics Experiment Informed Consent Form: I freely and voluntarily consent to be a participant in this research project conducted through March of 2018. I understand I will be one of many people participating in this

research and that this study will last approximately 15 minutes. I understand that the purpose of this research is to examine adults' behavior and performance in market settings. To this aim, I will be asked to fill out a survey and make decisions with monetary consequences. I understand that I am free to discontinue my participation at any time without forfeiting my show-up payment. I understand that all my responses will be anonymous and confidential. Only Professor Jeffrey Carpenter will have access to this data. I understand that any personal identifiers will be removed from the data. I understand that information from all the participants will be grouped together to provide general information about human behavior. I understand that if I would like more information about this research, I can contact Jeffrey Carpenter at 802-443-3241. I understand that questions about my rights as a research subject should be directed toward the Institutional Review Board at irb@middlebury.edu. I have read and I understand the above. I have been offered a copy of this informed consent form. By signing this form, I acknowledge that I am at least 18 years old. I understand that by clicking on the 'next' button I am confirming that I have read and understood the above instructions and that I consent to participate in the study.

3. Please enter your Middlebury College mailbox number (this will be used for payment purposes only).

4. Instructions for the task: In each of the boxes on the next page, you will find a matrix filled with twelve numbers, all of which are less than 10. To solve each matrix, you must select the two numbers among all the options that sum to exactly 10. Performance on the assessed portions of this game will be provided at the end of the experiment. Click on the boxes to select your two answers. You may click any box again to un-select that choice. Click the next arrow to proceed to 5 practice questions.

5. Task 1 (for payment): You will now have 4 minutes to solve as many matrices as you can. There are 20 matrices in total. If you solve 10 or more, you will win an additional $5. This $5 will be paid to you in cash immediately after the experiment. Click next to begin Task 1.

6. [Control] Task 2 (for payment): You will now have another 4 minutes to solve as many matrices as you can. There are 20 matrices total. If you solve 10 or more,

you will win an additional $50. This $50 will be paid to you in cash immediately after the experiment.

7. [Treatment] Task 2 (for payment):You will now have another 4 minutes to solve as many matrices as you can. There are 20 matrices total. If you solve 10 or more, you will win an additional $50. This $50 will be paid to you in cash two months after the experiment.

7.2 Survey questions

1. What is your Middlebury mailbox number? (Used for payment purposes only)
2. What year are you at Middlebury?
3. Sex/gender: Do you identify as Female, Male, Neither?
4. What are your average weekly earnings from a job during the school year?
5. 13-question Rotter Scale.
6. Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?
7. For the following 2 questions, please rate your competitiveness level on a scale of 0-10.
   (a) In general, how competitive do you think you are?
   (b) Concerning just sports and leisure activities, how competitive do you think you are?
8. 12-question DEEP time module.

7.3 Additional results

Table A1: Participant characteristics (by treatment).

	Overall	Pay Now	Pay Later	p-value
Low then High (I)	0.53	0.59	0.46	0.05
Female (I)	0.42	0.33	0.50	0.02
Class of 2021 (I)	0.30	0.36	0.24	0.06
Class of 2020 (I)	0.31	0.22	0.40	0.01
Class of 2019 (I)	0.24	0.27	0.20	0.23
Class of 2018 (I)	0.15	0.14	0.16	0.71
Not Work (I)	0.48	0.48	0.48	0.95
Competitive (I)	0.32	0.39	0.25	0.03
Internal Locus (I)	0.51	0.53	0.49	0.62
Distrust (I)	0.43	0.43	0.42	0.84
Delta	0.0049	0.0046	0.0052	0.17
Beta	1.01	1.04	0.98	0.04

Note: (I) denotes indicator; characteristic means reported.

Table A2: The treatment effects of delayed payments.

	Choke (strict)	Choke (loose)	Frac(Total Pay)
Pay Later (I)	-0.094	-0.138**	0.154***
	(0.065)	(0.063)	(0.046)
Low then High (I)	0.005	-0.070	0.025
	(0.062)	(0.079)	(0.050)
Class of 2020 (I)	0.123	0.009	0.051
	(0.112)	(0.108)	(0.092)
Class of 2019 (I)	-0.100*	-0.026	0.019
	(0.056)	(0.098)	(0.105)
Class of 2018 (I)	-0.088	0.021	0.029
	(0.080)	(0.132)	(0.093)
Not Employed (I)	-0.081	-0.062	0.027
	(0.061)	(0.101)	(0.061)
Female (I)	0.007	-0.027	-0.105***
	(0.041)	(0.073)	(0.032)
Competitive (I)	0.179**	0.031	0.007
	(0.062)	(0.066)	(0.052)
Internal LoC (I)	0.023	0.049	-0.019
	(0.028)	(0.063)	(0.080)
Distrust (I)	-0.028	-0.005	-0.054
	(0.054)	(0.049)	(0.060)
Intercept	0.178	0.532***	0.324***
	(0.103)	(0.150)	(0.140)
Observations	199	199	199
Adj. R^2	0.133	0.033	0.058

Notes: Linear probability estimates; (robust standard errors) clustered on the session;

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Heterogeneous treatment effects based on present bias.

	Choke (strict)	Choke (loose)	Frac(Total Pay)
Pay Later (I)	-0.010	0.017	0.037
	(0.055)	(0.075)	(0.061)
Present-Biased (I)	0.073	0.132	-0.149**
	(0.094)	(0.125)	(0.064)
Pay Later × Present-Biased	-0.210**	-0.386**	0.297**
	(0.084)	(0.146)	(0.108)
Low then High (I)	0.013	-0.055	0.013
	(0.060)	(0.078)	(0.054)
Class of 2020 (I)	0.130	0.023	0.045
	(0.109)	(0.101)	(0.089)
Class of 2019 (I)	-0.082	0.007	-0.001
	(0.059)	(0.095)	(0.099)
Class of 2018 (I)	-0.076	0.043	0.013
	(0.075)	(0.120)	(0.091)
Not Employed (I)	-0.078	-0.056	0.026
	(0.053)	(0.084)	(0.056)
Female (I)	0.020	-0.003	-0.119**
	(0.037)	(0.056)	(0.040)
Competitive (I)	0.166**	0.008	0.028
	(0.055)	(0.060)	(0.050)
Internal LoC (I)	0.023	0.049	-0.018
	(0.026)	(0.063)	(0.079)
Distrust (I)	-0.032	-0.012	-0.045
	(0.048)	(0.044)	(0.067)
Intercept	0.139*	0.460***	0.388**
	(0.078)	(0.131)	(0.121)
Observations	199	199	199
Adj. R^2	0.154	0.072	0.089

Notes: Linear probability estimates; (robust standard errors) clustered on the session;

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Heterogeneous treatment effects based on discount rate.

	Choke (strict)	Choke (loose)	Frac(Total Pay)
Pay Later (I)	-0.136*	-0.163**	0.076
	(0.068)	(0.067)	(0.069)
High Discount Rate (I)	0.046	-0.061	-0.090
	(0.043)	(0.118)	(0.072)
Pay Later × High Discount Rate	0.086	0.052	0.161*
	(0.058)	(0.142)	(0.084)
Low then High (I)	0.013	-0.069	0.032
	(0.062)	(0.080)	(0.046)
Class of 2020 (I)	0.113	0.011	0.049
	(0.118)	(0.108)	(0.091)
Class of 2019 (I)	-0.105	-0.025	0.016
	(0.064)	(0.097)	(0.103)
Class of 2018 (I)	-0.101	0.021	0.020
	(0.082)	(0.130)	(0.092)
Not Employed (I)	-0.084	-0.064	0.021
	(0.066)	(0.103)	(0.063)
Female (I)	0.008	-0.029	-0.109***
	(0.044)	(0.074)	(0.030)
Competitive (I)	0.192***	0.028	0.010
	(0.064)	(0.065)	(0.047)
Internal LoC (I)	0.022	0.052	-0.012
	(0.032)	(0.062)	(0.075)
Distrust (I)	-0.041	-0.003	-0.058
	(0.051)	(0.051)	(0.059)
Intercept	0.159	0.560***	0.367**
	(0.107)	(0.172)	(0.153)
Observations	199	199	199
Adj. R^2	0.151	0.035	0.068

Notes: Linear probability estimates; (robust standard errors) clustered on the session;

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Heterogeneous treatment effects comparison.

	Choke (strict)	Choke (loose)	Frac(Total Pay)
Pay Later (I)	-0.059	-0.045	-0.008
	(0.056)	(0.080)	(0.060)
High Discount Rate (I)	0.044	-0.065	-0.085
	(0.042)	(0.111)	(0.071)
Pay Later × High Discount Rate	0.167**	0.163	0.103
	(0.074)	(0.141)	(0.085)
Present-Biased (I)	0.071	0.135	-0.145**
	(0.096)	(0.127)	(0.062)
Pay Later × Present-Biased	-0.286***	-0.425**	0.286**
	(0.076)	(0.150)	(0.114)
Low then High (I)	0.028	-0.046	0.016
	(0.062)	(0.078)	(0.050)
Class of 2020 (I)	0.122	0.021	0.046
	(0.114)	(0.102)	(0.089)
Class of 2019 (I)	-0.081	0.008	-0.001
	(0.069)	(0.099)	(0.099)
Class of 2018 (I)	-0.094	0.033	0.010
	(0.076)	(0.122)	(0.090)
Not Employed (I)	-0.080	-0.060	0.023
	(0.055)	(0.085)	(0.057)
Female (I)	0.027	-0.003	-0.122***
	(0.038)	(0.060)	(0.037)
Competitive (I)	0.183***	0.012	0.026
	(0.057)	(0.059)	(0.048)
Internal LoC (I)	0.024	0.055	-0.012
	(0.032)	(0.065)	(0.076)
Distrust (I)	-0.051	-0.019	-0.044
	(0.041)	(0.042)	(0.065)
Intercept	0.113	0.486***	0.426***
	(0.081)	(0.158)	(0.130)
Observations	199	199	199
Adj. R^2	0.191	0.078	0.095

Notes: Linear probability estimates; (robust standard errors) clustered on the session;

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Heterogeneous treatment effects based on distrust.

	Choke (strict)	Choke (loose)	Frac(Total Pay)
Pay Later (I)	-0.124	-0.123	0.120
	(0.106)	(0.097)	(0.077)
Distrust (I)	-0.064	0.013	-0.095
	(0.099)	(0.084)	(0.078)
Pay Later × Distrust	0.071	-0.035	0.082
	(0.136)	(0.168)	(0.105)
Low then High (I)	0.005	-0.070	0.026
	(0.063)	(0.080)	(0.050)
Class of 2020 (I)	0.121	0.009	0.049
	(0.110)	(0.108)	(0.092)
Class of 2019 (I)	-0.102*	-0.025	0.016
	(0.053)	(0.100)	(0.104)
Class of 2018 (I)	-0.084	0.019	0.034
	(0.085)	(0.133)	(0.092)
Not Employed (I)	-0.081	-0.062	0.027
	(0.061)	(0.101)	(0.062)
Female (I)	0.007	-0.027	-0.105***
	(0.040)	(0.073)	(0.032)
Competitive (I)	0.184**	0.029	0.013
	(0.064)	(0.068)	(0.050)
Internal LoC (I)	0.020	0.050	-0.022
	(0.028)	(0.064)	(0.082)
Intercept	0.193	0.524***	0.341**
	(0.110)	(0.164)	(0.148)
Observations	199	199	199
Adj. R^2	0.135	0.033	0.060

Notes: Linear probability estimates; (robust standard errors) clustered on the session;

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.