Dismissal Threats as an Incentive Device

Michael T. Cooper*
University of California, Santa Barbara

November 2, 2020

Abstract
This paper explores dismissal threats as an incentive device in a laboratory experiment. Workers earn a fixed wage per period and complete real effort tasks to reduce their chance of being fired at the end of each period. Behavioral motivators are purposefully activated in addition to these monetary incentives. The design innovates on previous literature by implementing dismissal threats in a quantifiable way and by eliciting subject preferences over incentives while still imposing random variation in those incentives. The experiment produces two main results. First, dismissal threats are more motivating than comparable piece rates. Workers produce significantly more output under dismissal threats than they do under piece rates, even though the marginal benefit of output is lower. Second, the productivity gains from strengthening dismissal threats on the margin have a large self-selection component but significant heterogeneity in pure incentive effects. Workers who prefer higher pay with steeper dismissal threats appear to respond positively to this environment, but these high-pressure incentives backfire among workers who wanted to avoid them.

JEL Codes: M52, J33, D91, C91.

*E-mail: michael.cooper@ucsb.edu. Website: www.michaelthomascooper.com
1 Introduction

The threat of being fired for poor performance is a ubiquitous incentive throughout the modern economy. Workers who wish to retain their wages and benefits from employment must provide enough visible effort to avoid dismissal and a potentially costly period of involuntary unemployment. A recent poll conducted during a time of economic expansion and low unemployment found that 48% of employed Americans experienced anxiety over the possibility of losing their job, and 40% of Americans have been laid off or terminated from a job at least once (Harris Poll, 2019). Despite the prevalence of firing threats in the labor market, little is known about how strongly they motivate workers to provide effort compared to other monetary incentives.

This paper reports the results of an experiment which explores dismissal threats as a rigorously quantified incentive device in the lab. Subjects complete a real effort task for pay under various types and strengths of incentives. Under dismissal threats, subjects receive a fixed payment per period and complete tasks to reduce their chance of being fired at the end of each period. This implementation of dismissal threats in the lab as a known probabilistic relationship between output and the chance of being fired brings tractability to an incentive that is often ambiguous and not explicitly quantified in labor market contracts in the field. An important benefit of this approach is that the marginal benefit of output under dismissal threats is known to workers and observable to the researcher. This tractability allows direct comparisons of worker output under dismissal threats to output under simpler monetary incentives such as piece rates.

Few jobs are entirely devoid of firing threats, but in practice the strength of this incentive varies considerably. Some jobs, such as executive leadership of large public corporations, have high employer-initiated turnover rates – CEOs are often terminated based on poor performance (e.g., stock value). Other jobs seem to have much weaker dismissal threats, with a feature of strong job security as long as lenient baseline levels of performance are met. Jobs with unionized workforces have much higher levels of job security, as do jobs in the public sector (Freeman and Medoff, 1984; Farber, 2010). Variation in the strength of dismissal threats also occurs based on policies, including employment protection laws and court rulings (e.g., statewide “just cause” doctrines studied by Autor, Kerr, and Kugler, 2007).

An important benefit of studying dismissal threats in the lab is the ability to produce exogenous variation in the strength of the incentive. In the field, how motivating dismissal threats are to workers depends on a number of factors, including not only the relationship between output and the chance of being fired, but also considerations such as economic
conditions. Indeed, research has shown that workers increase output when unemployment would be more financially painful due to a recession or local labor market conditions (Cappelli and Chauvin, 1991; Lazear, Shaw, and Stanton, 2016). The lab environment will abstract away from these concerns, simplifying the monetary incentives of dismissal threats to merely losing income from all remaining work periods after a once-and-for-all firing has occurred. In this design, the incentives under firing threats will be characterized by three parameters: the fixed wage per period and the slope and intercept of a linear function that relates output to the probability of being fired. Exogenous variation in these parameters allows analysis of worker behavior under steeper versus flatter dismissal threats.

Surprisingly, little research has been conducted on how firing threats operate to motivate workers. Instead, the literature has frequently analyzed financial incentives like piece rates, tournaments, and teams. Although these incentives serve as important tools for employers, they seem less common in practice than fixed wages (hourly wages or salaries) combined with dismissal threats. Piece rates could be less common than dismissal threats for a variety of reasons, including the difficulty of measuring individual worker output or the fear of workers neglecting unincentivized criteria at the cost of focusing only on measured, incentivized tasks. However, one potential reason stands out as particularly interesting to behavioral and personnel economists: the threat of being fired combines neoclassical and behavioral motivators in a way that could be uniquely powerful in motivating worker effort.

Consider the mixture of neoclassic and behavioral mechanisms involved when using firing threats to motivate workers. Most obviously, being fired means earning less money, causing a direct utility loss. Whether one will be fired is usually uncertain until it happens, thus bringing risk aversion into consideration. However, the widespread fear and anxiety over being fired (as mentioned in the polling data above) does not seem congruent with what would be caused under piece rates or even simple monetary incentives with risk involved. Instead, it seems likely that deeper behavioral and even evolutionary motivations are at play, such as the fear of exclusion from a social group. Fired workers may become jealous of retained coworkers who continue to earn money and perform interesting work. They may feel embarrassed about public announcements of their firing, or experience negative self-image impacts resulting from the signal that firing provides about their own value in the labor market. Loss aversion likely plays a role when keeping one’s job is the status quo and being fired feels like a large negative change compared to prior expectations.

These behavioral motivators are purposefully activated in the experiment design because they seem likely to play a large role in why firing threats are motivating. The event of being fired is framed as a loss throughout the instructions and work interface, and public
announcements of firings are made at the end of each work period.\footnote{Fired workers have their outside option tightly controlled; they must simply wait while other workers continue to earn money, further developing feelings of exclusion. Given that it would be impossible to activate these behavioral mechanisms in the lab at the full strength that they exist in the real labor market, the estimates presented in this paper of the motivating effect of dismissal threats are likely a lower bound on their effect in the field.}

The results of the experiment shed light on the use of dismissal threats to motivate workers. The main results fall into two categories: the effect of dismissal threats on the extensive margin (the existence of credible dismissal threats) and on the intensive margin (intensifying existing dismissal threats). On the extensive margin, dismissal threats are shown to be highly motivating compared to piece rates. The experiment is designed such that the marginal benefit of output under dismissal threats is always lower than it was under the high piece rate work period. Despite the lower monetary benefit, mean productivity in the first dismissal threat work period is 23.6% higher than under the piece rate, a highly statistically significant difference. Parametric analysis shows that this cannot be caused merely by the levels of risk aversion displayed in subject decisions over gambles; instead, the unique mixture of monetary and behavioral motivators invoked by the dismissal threats in the lab seems to be highly motivating in inducing worker effort.

The implications of this result are both prescriptive and descriptive. First, firms may experience large productivity gains by introducing credible dismissal threats that did not exist before (or were too weak to be salient to workers). Second, this may be an important reason why dismissal threats are so widespread in the modern labor market: they are more motivating than simpler monetary incentives like piece rates.

On the intensive margin, steeper dismissal threats combined with higher pay result in higher output than flatter dismissal threats with lower pay, as expected. However, the design allows further decomposition of this productivity increase into the pure incentive effect (steeper incentives result in higher optimal output) and the selection effect (more productive workers self-select into steeper incentives). The design elicits subject preferences over incentive schemes while retaining random assignment to those schemes, all collapsed into one round. This both saves time and avoids order contamination effects compared to previous approaches in the literature. With both subject preferences and random assignment, a clean decomposition of incentive effects from selection effects is possible.

The qualitative results of this decomposition of incentive versus selection effects are informative for both firms and policymakers. If the incentive effect dominates, then imposing

\footnote{Subjects are assigned random worker identification numbers to retain their anonymity; these worker IDs are displayed in the public announcements.}
stronger dismissal threats on an existing workforce would cause large short-term productivity gains; if the selection effect dominates, productivity would increase only gradually as the composition of workers changes over time. This decomposition is also relevant to policy debates over employment protection and productivity. Taking an extreme example, when all firms offer the same level of dismissal threats, the selection effect on productivity is entirely removed – there must be different levels of dismissal threats and pay for this effect to operate. Following this logic, widespread employment protection policies are predicted to have muted effects on productivity if the selection effect is a large component of the total effect, but would cause large declines in productivity if the incentive effect is large.

It turns out that the aggregate productivity gains on the intensive margin (strengthening existing dismissal threats) operate largely through selection effects. The workers who indicated a preference for the high-paying job with stronger dismissal threats were significantly more productive than workers with the opposite preference. Aggregate incentive effects were muted: random assignment to the steeper incentive scheme was not associated with significantly higher output.

However, these aggregate effects mask important heterogeneity in incentive effects. Among workers who preferred higher pay with stronger dismissal threats, there is suggestive evidence that they do increase output under these steeper incentives (although it is not statistically significant). On the other hand, workers who were willing to accept lower pay to avoid strong dismissal threats actually produce significantly fewer units of output when randomly assigned to the strong dismissal threats. This negative incentive effect of steeper incentives can be explained in line with Roy’s (1951) model of worker self-selection: these are precisely the workers who expect to do the worst in this high-pressure environment. Indeed, 85% of this subset of subjects would end up with greater than a 50% chance of being fired even if they sustained their level of output from the high piece rate work period. The takeaway is that imposing steeper dismissal threats may be demotivating if they are too strong compared to the ability level of the worker – an important lesson for firms considering strengthening this form of incentive on a mixed-ability workforce.

The rest of the paper will be organized as follows. Section 2 will discuss the related literature and the unique contributions of this paper compared to previous work. Section 3 will provide details about the experimental design. Section 4 will present the results, and Section 5 will conclude.
2 Related Literature

This paper contributes to several different strands of economics literature, but it most directly relates to the personnel economics literature on motivating workers to provide effort. Early contributions to this literature established the principal-agent model and showed how providing piece rates can overcome the moral hazard issue facing employers trying to motivate their employees to provide costly effort (Ross, 1973; Stiglitz, 1975). Much of the subsequent experimental literature has focused on comparing piece rates to other payment schemes such as tournaments or team-based revenue sharing (Bull, Schotter, and Weigelt, 1987; Niederle and Vesterlund, 2007; Dohmen and Falk, 2011).

Applied work has also made many important contributions to understanding how motivating different types of incentives are for workers. In a seminal paper, Lazear (2000) studies the shift from hourly wages to piece rates in a large windshield repair company. Making this study even more relevant to the current design, Lazear also uses econometric methods (worker fixed effects) to try to decompose the productivity gains into pure incentive effects and selection effects. He finds that roughly half of the productivity gains come from the selection effect of hiring and retaining more productive workers. However, he cannot rule out that changes in management coinciding with the incentive changes may confound some of these measurements. Indeed, this is a perennial issue in applied work: identifying policy changes in a firm or government that are truly exogenous. It can also be difficult to find changes to employment contracts which only impact one part of the incentive structure.

Jacob (2013) produces the applied work most relevant to this experiment by studying a policy change in a Chicago public school district that made dismissing certain teachers with limited years of experience easier. He measures productivity in the number of days of teacher absences. Similar to Lazear (2000), Jacob uses worker fixed effects to try to isolate the selection effect on productivity. He finds that changes to the composition of teachers (i.e., the selection effect) is the predominant cause of the post-policy reductions in teacher absences, and finds much more limited evidence of pure incentive effects. In other work, Jacob (2011) shows that after this policy, principals are more likely to fire teachers with more absences and lower value-added scores. Although the policy he analyzes in Chicago provides a plausibly exogenous reduction in employment protection, it is impossible to cleanly quantify and manipulate the strength of dismissal threats in the field in the way one can in a controlled laboratory experiment. Other advantages of an experiment include the ability to observe the same workers under a variety of incentive schemes (e.g., both piece rates and dismissal threats), and the possibility of measuring many worker traits with incentivized tasks.
Kwon (2005) answers a similar but distinct question: When a company decides to fire a worker, was it done to incentivize other workers or because the fired worker was a poor match for the job? He develops a model that shows each explanation predicts an opposite relationship between tenure and dismissal probability. Using proprietary data from a large U.S. insurance company, he tests his theoretical predictions and finds evidence that workers are dismissed as a way of providing incentives for effort, rather than because they were a bad match for the job. This conclusion of course depends on various assumptions of the model, and differs from the primary questions of the present design: how strongly motivating dismissal threats are as an incentive device, and whether the productivity gains from exogenous changes in dismissal threats are primarily due to selection or incentive effects.

Other researchers have conducted important applied work on the effects of employment protection and productivity. Autor, Kerr, and Kugler (2007) exploit the variation in state court cases that established just cause doctrines and other exceptions to at-will employment. They find that this trend in employment law is associated with capital deepening and declines in total factor productivity in manufacturing plants, but caution that the findings are “suggestive but tentative” due to concurrent increases in employment. Acemoglu, Daron, and Angrist (2001) find a link between a major employment protection law for disabled people and subsequent reductions in their employment rate.

A number of experiments have alternatively allowed subjects to select into their preferred type of incentives with certainty, or randomly assigned subjects an incentive scheme. For example, Dohmen and Falk (2011) run an important experiment on multidimensional sorting of workers into fixed wages versus variable pay (i.e., payments that depend on output: piece rates, tournaments, or teams). However, because subjects are guaranteed the incentive scheme they choose, the authors are unable to precisely decompose the productivity gains under variable pay into incentive and selection effects. Instead, they are merely able to assert that selection on productivity exists, since subjects who selected the variable pay schemes were more productive on average in an earlier piece rate work period.

Other examples of experiments with guaranteed selection into incentive schemes include Cadsby, Song, and Tapon (2007) who examine piece rate contracts, and Leuven, Oosterbeek, Sonnemans, and van der Klaauw (2011) who examine selection into different tiers of tournaments. Fehrenbacher, Kaplan, and Pedell (2017) have two treatments: one with random assignment of incentive scheme, and one with guaranteed assignment to the selected incentive scheme. However, because they do not elicit the preferences of subjects who are randomly assigned, they cannot compare how well subjects with different job preferences would have performed under alternative incentive schemes.

The design reported in this paper significantly innovates on this strand of literature on
selection effects by eliciting subject preferences over incentives in an incentive-compatible way while retaining random assignment to incentive schemes. This allows the observation of many subjects in the incentive scheme they did not select, thus providing a clean decomposition of selection versus sorting effects. Additionally, the focus on dismissal threats in a tangible, quantifiable manner which can be exogenously strengthened or weakened is unique in the literature and easily exported to other experiments on dismissal threats as a motivating device.

Some experimental work has directly implemented dismissal threats in the lab. Corgnet, Hernan-Gonzalez, and Rassenti (2015) create groups of subjects who are randomly assigned to be bosses or employees, allowing the boss to dismiss one employee at the end of each work period. They do not examine selection effects, but find that employees increase their productivity and engage in “impression management” when dismissal threats are introduced (i.e., they spend less time on real leisure and more time on real effort tasks when both are visible to the boss). Their design differs from the one in this proposal because with human bosses and employee counterparts, the strength of the firing threat cannot be known or quantified by employees. Thus perceptions (or misperceptions) of the strength of dismissal threats may drive differences in worker behavior. Additionally, when examining the question of how motivating dismissal threats are in increasing worker output, they only compare dismissal threats to a situation with no monetary incentives at all, rather than to other kinds of monetary incentives such as piece rates.

Other experimental work seems to uncover similar behavioral forces to those analyzed here: the discontinuous effect of introducing a fear of exclusion or financial survival into an economic decision-making problem. For example, Kopányi-Peuker, Offerman, and Sloof (2018) show that contributions greatly increase in a weakest link game when a designated manager is able to permanently exclude subjects with low contributions from the team. Oprea (2014) finds strong evidence that subjects suboptimally hoard excess cash to improve their odds of avoiding bankruptcy. It seems likely that similar behavioral forces are activated in the fear of being fired.

Finally, this paper is tangentially related to a few other literatures in labor economics. First, it is related to the literature on efficiency wages (Shapiro and Stiglitz, 1984). It is more costly to lose a job with higher pay, and this effect operates in this experiment as well. Even holding constant the probability of dismissal, increasing pay would be predicted to increase optimal output for this reason: there is more to lose by becoming unemployed. Second, it is related to the literature on deferred compensation (Lazear, 1979; Kuhn, 1986). The idea is that workers are paid below their marginal product when they are young, and then they collect rents above their marginal product when they are old. A similar force operates in the
experiment because subjects can earn full job security in the final period and shirk without consequence – this is a motivating factor to provide higher effort in earlier periods.

Third, this paper is somewhat related to the literature on relational contracting in terms of the distinction between contracts which are locked in and repeated versus contracts which must be actively renewed after each transaction. The design in this paper has actively chosen a loss framing for dismissal (i.e., subjects are told they will “lose” their job and will not be able to earn more money). This stands in contrast to the alternative of temporary job contracts which are expected to be terminated unless actively renewed (more common in certain European countries with strong employment protection; see Cahuc and Postel-Vinay, 2002). Cromwell, Goerg, and Leszczynska (2018) show that this distinction can have an important impact on behavior in the lab even when it provides no tangible difference in flexibility or incentives. Therefore, the decision to frame dismissal as a loss of a contract compared to the status quo could have important implications for the power of dismissal threats as an incentive device in the present design.

3 Experiment Design

This section will explain the experimental design and the motivation behind various design decisions. The primary purpose of the experiment was to expose subjects to various types and strengths of incentives and observe their resulting productivity in a real effort task. The subject pool consisted primarily of students at the University of California, Santa Barbara. The experiment was run entirely online, with instructions read over a Zoom video chat conference and payments made through the mobile application Venmo. The experiment software was coded in oTree (Chen, Schonger, and Wickens, 2016).

The real effort task involved counting shapes that appear in randomly generated images. This shape-counting task is largely the same real effort task used by Caplin, Csaba, Leahy, Nov (2020). A primary strength of this particular geometric reasoning task is that Caplin et al. show that subject performance is responsive to different levels of monetary incentives. Additionally, it is culturally neutral and relatively simple to explain to subjects. The task was slightly modified from the implementation of Caplin et al. to disincentivize guessing and to provide more granular data on worker output. The difficulty level of the task did not vary throughout the experiment. Additionally, to increase variation in subject productivity, real leisure was included in the work interface in the form of interesting facts (e.g., on animals or

---

Subjects initially presented identification in a private breakout room, but were muted with their cameras off for the remainder of the experiment. Thus the anonymity of subjects from one another was maintained. Venmo is ubiquitous among the subject pool; all subjects agreed to accept payment through this app before participating in the experiment.
Astronomy). A more detailed description of this task along with a screenshot of the work interface are provided in the appendix.

After the initial instructions, subjects proceeded through a series of work periods, each carefully designed to answer questions about dismissal threats as an incentive device. First, subjects completed a two-minute practice period to familiarize themselves with the task. Next, subjects were told they would be paid per task completed in two successive five-minute work periods: the first period with a piece rate of $0.01 per task, and the second period with a $0.20 payment per task. These two pay rates were displayed saliently before either piece rate period began so subjects would know that their effort was significantly more valuable in one period than another.

Following the piece rate work periods, there were three work periods under dismissal threats, each lasting 10 minutes. Dismissal threats were implemented as a linear relationship between output and the probability of being fired at the end of each remaining work period. Subjects were paid a fixed wage up front at the start of each period. Given that there was no fourth dismissal threat period, the third period acted as a simulation of full job security; trivially, optimal output in the final period is zero. The monetary incentives of dismissal threats thus can be fully characterized by three parameters: the payment per period along with the intercept and slope of the linear firing function. Because these parameters are known to the experimenter and the subjects, the marginal benefit of output can be calculated in the first and second dismissal work periods, making direct comparisons with piece rates possible.

Workers are assigned to one of two jobs, displayed in Figure 1. In the Safe Job, the pay per period is $2.50, the firing-chance intercept is 50%, and the slope is a 1% lower chance of being fired per task completed. In the Risky Job, these parameters are $5.00, 100%, and 2%, respectively. Crucially, the two piece rates of $0.01 and $0.20 were chosen as the lower and upper bounds on the marginal benefit of output under dismissal threats. This allows the design to answer the following research question:

**Question 1: Extensive Margin.** Are dismissal threats more motivating than piece rates with comparable monetary value?

---

3A silver lining of running the experiment online is that subjects could not be prevented from using their phones during the work periods – another form of real leisure that increases variation in output.

4Subjects were also shown a series of tutorial screens with screenshots of the work interface. The tutorial and unpaid practice period are important to reduce the potential effects of learning to do the task more efficiently.

5Although the two parametrizations will be called the Safe Job and the Risky Job throughout the text, they were only referred to with neutral language (“Job 1” and “Job 2”) in the experiment instructions and software.

9
Given that the marginal benefit of output under dismissal threats is always lower than $0.20, one would expect that worker output under dismissal threats is lower than output under the piece rate of $0.20. If output under dismissal threats is higher than output under this piece rate, it would be strong evidence that dismissal threats are a more motivating type of incentive than piece rates.

The marginal benefit of output in each period and job is calculated in Table 1 by taking the derivative of expected pay with respect to output in periods 1 and 2. Of course, the marginal benefit of output in period 1 depends on the level of output in period 2; but across the whole range of potential period 2 outputs (from 1 to 50), the marginal benefit of output in period 1 remains between $0.01 and $0.20. The table confirms that under any job-period combination, and while incorporating the option value of being retained after period 1, the marginal benefit of output under dismissal threats is always lower than $0.20.

Of course, dismissal threats constitute more than simply risky payments; behavioral motivators are purposefully activated by the design as well. The instructions framed retention as the status quo and being fired as an event which caused a subject to miss out on payments. Being retained was described in green text while being fired was always displayed in bold red text. To additionally activate social and self-image concerns, public announcements of firing decisions are made at the end of each work period. These announcements emphasized that

---

Note that the marginal benefit of output in period 2 is calculated from the perspective of period 2 (i.e., for a worker who has already been retained) Trivially, output in period 3 never impacts total pay, so that period is excluded from the analysis.
Table 1: Marginal Benefit of Output during Dismissal Periods

<table>
<thead>
<tr>
<th></th>
<th>Marginal Benefit of $q_1$</th>
<th>Marginal Benefit of $q_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Risky Job</strong></td>
<td>$\frac{\partial E[\text{Pay}]}{\partial q_1} = 0.01 + 0.002q_2$</td>
<td>$\frac{\partial E[\text{Pay}]}{\partial q_2} = $0.10$</td>
</tr>
<tr>
<td></td>
<td>if $q_2 = 0 \rightarrow $0.01$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>if $q_2 = 50 \rightarrow $0.11$</td>
<td></td>
</tr>
<tr>
<td><strong>Safe Job</strong></td>
<td>$\frac{\partial E[\text{Pay}]}{\partial q_1} = 0.0375 + 0.00025q_2$</td>
<td>$\frac{\partial E[\text{Pay}]}{\partial q_2} = $0.025$</td>
</tr>
<tr>
<td></td>
<td>if $q_2 = 0 \rightarrow $0.0375$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>if $q_2 = 50 \rightarrow $0.05$</td>
<td></td>
</tr>
</tbody>
</table>

retained workers would continue to earn money. Subjects were kept anonymous through the use of randomly assigned worker identification numbers, so examples of these notifications would be “Worker 3 was fired!” in bold red text, or “Worker 7 was not fired! They will continue to earn money next period.” in green text.

The outside option of workers was tightly controlled to prevent uncontrolled variation in the cost of being fired and to further induce feelings of exclusion among fired workers. Fired workers had to complete a boring filler task for any remaining work periods: clicking a large button when it occasionally turned red to avoid losing a small amount of money. This control over the outside option increases the internal validity of the design by preventing workers from attempting to get fired on purpose in order to engage in other work or leisure.

Further elements of the experiment were designed to analyze the intensive margin effects of strengthening dismissal threats:

**Question 2.1: Intensive Margin.** Do stronger dismissal threats result in higher output than weaker dismissal threats?

The parameters of the Risky Job are predicted to be significantly more motivating than the parameters of the Safe Job. Crucially, the slope of the firing function in the Risky Job is twice as steep, with each unit of output reducing the chance of being fired by 2% rather than 1% in the Safe Job. Also, since the pay per period is higher in the Risky Job, subjects have more to lose by being fired.

---

7Subjects were shown a screenshot of an example firing announcement ahead of time so that social/self-image concerns would be activated in the first dismissal work period.

8The button-clicking task was intentionally designed to be extremely easy and merely required subjects to stay attentive at their computer. Exceedingly few dismissed subjects ever missed a button click, and most of those subjects missed only a single click.
The answer to the above question is largely predictable, but the design seeks to further analyze these intensive margin productivity gains:

**Question 2.2: Incentive and Selection Effects.** Do stronger dismissal threats increase output primarily through the incentive effect or the selection effect?

The incentive effect is the straightforward impact of steeper incentives increasing the optimal level of output. This means that exogenous assignment to the Risky Job should result in higher levels of output. The selection effect is that more productive workers should prefer the Risky Job while less productive workers should prefer the Safe Job. Although the quantitative decomposition of these effects will not be representative of the labor market, the qualitative results (i.e., which effect is larger) have important implications for firms and policymakers, as discussed in the introduction.

The design allows for a clean decomposition of the incentive and selection effects by eliciting worker preferences over jobs while still retaining a random element to job assignment. Specifically, subjects have a 60% chance of receiving the job they prefer. Because they are more likely to receive their preferred level of pay and dismissal threats, this mechanism retains incentive compatibility – subjects should honestly report their preference. The mechanism also imposes exogenous variation in the strength of dismissal threats, allowing the isolation of the effect of random assignment to a steeper or flatter firing function. This mechanism thus provides significant benefits over implementations in previous experiments, which have opted for either guaranteed assignment to one’s preferred incentive scheme or for fully random assignment. Simple regression analysis can decompose the intensive margin productivity gains of steeper dismissal threats into the effect of preferring the Risky Job and the effect of being randomly assigned the Risky Job.

In the remainder of the experiment after the work periods, subjects completed a series of incentivized tasks and unincentivized survey questions designed to measure a wide variety of worker traits. Full descriptions of these measurement procedures are provided in the appendix. For the body of the paper, it will suffice to note that standard methods from

---

9Note that under the implemented parameters, the expected total pay in the Risky Job will be higher than in the Safe Job for any level of worker output. In the pilot session, the pay was $3.00 in the Safe Job to ensure that expected pay was higher for very low productivity subjects in the Safe Job. However, the majority of subjects chose the Safe Job under these parameters, leading to the final adjustment to $2.50 which caused a much more even split in job preference.

10The experiment was not designed to decompose the extensive margin productivity gains into selection and incentive effects for two main reasons. First, it was unclear ex ante whether dismissal threats would actually result in productivity gains compared to higher-powered piece rates. Second, the mechanism requires subjects to be assigned to only one of the two incentive types (piece rates or dismissal threats). This would have prevented important within-subject analysis of whether the productivity gains are driven by individual risk preferences.
the modern experimental literature are used to elicit the following traits: risk aversion, loss aversion, patience, confidence, altruism, trust, and reciprocity.

4 Results

This section will present the main results of the experiment. A total of 8 sessions were run, and 101 subjects participated. Each subject participated in only one session. The average payment (including the show-up payment) was $16.36, with a minimum of $5.91 and a maximum of $30.01. The typical session lasted about 1 hour and 20 minutes, and no session lasted longer than 1 hour and 30 minutes.

This section will be divided into two subsections analyzing the extensive and intensive margin effects of dismissal threats on productivity, respectively. First, worker output under piece rates and dismissal threats will be compared. The results show a strong extensive margin effect of dismissal threats: this form of incentive is significantly more motivating than comparable piece rate incentives. The large difference in output cannot be explained by individual risk preferences.

Second, the intensive margin effects of stronger dismissal threats will be analyzed. As expected, stronger dismissal threats result in higher output than weaker ones. This aggregate productivity gain is caused almost entirely by selection effects; random assignment to steeper dismissal incentives appears to cause only a small and statistically insignificant productivity gain. However, these aggregate results hide important heterogeneity in incentive effects. Among workers who prefer to sacrifice pay for weaker dismissal threats, random assignment to strong dismissal threats appears to backfire and reduce output. Workers who wanted higher pay along with steeper dismissal threats do produce more output under these steeper incentives, although the difference is not statistically significant.

4.1 Extensive Margin: Dismissal Threats versus Piece Rates

The real effort task successfully induced large differences in output between subjects. During the 5-minute work period with the high $0.20 piece rate, there was significant dispersion in output levels. The mean level of output was 10.6 tasks with a standard deviation of 4.9. Figure 2 displays a histogram of subject output during this work period, visually confirming large differences in productivity between subjects facing the same level of incentives. This dispersion in output under identical incentives is evidence that the real effort task was sensitive to pre-existing differences in subject ability – a necessary prerequisite for detecting selection on worker productivity into different incentive schemes.
Additionally, there is strong evidence that output changes in response to financial incentives. When the piece rate increased from $0.01 to $0.20, 80.2% of subjects increased output. The mean increase in tasks completed was 2.89 tasks (with a median of 3), and the mean percentage increase in tasks completed was 82.4% (although this metric is inflated by certain subjects who heavily shirked under the low piece rate – the median percentage increase was 44.2%). Of the 20 subjects who did not increase output in response to the piece rate increase, 6 subjects produced the same amount in both periods and 16 subjects showed a decline in output of 2 units or fewer. A Wilcoxon signed-rank test rejects the null hypothesis that the distribution of output was equivalent in both piece rate periods (p < 0.0001). Taken together, the between-subject dispersion in output while facing the same incentives along with the within-subject dispersion in output while facing different incentives implies that the real effort task was sensitive to both pre-existing ability and financial incentives.

Having established that subjects change output in response to variation in simple piece rate incentives, the analysis will now turn to output under dismissal threats. Differences in output between the two jobs with different pay and strength of dismissal threats will be primarily reserved for the following subsection which analyzes the incentive and selection effects; the remainder of this subsection will focus on comparing piece rates to dismissal threats in general.

The following analysis will present compelling evidence that the unique mixture of monetary and behavioral motivators invoked by dismissal threats were highly motivating – indeed,
much more motivating than would be expected under piece rates combined with uncertainty and risk aversion. First, simple nonparametric analysis will be presented, showing that worker productivity during the first two dismissal work periods is significantly higher than productivity during the high piece rate work period. This occurs even though the monetary marginal benefit of output is lower during the dismissal work periods than during the high piece rate period (this comparison includes the option value of working in the second period after being retained). Second, parametric analysis is conducted in which a simple quadratic functional form for the cost of effort function is assumed. The cost of effort function is calibrated using output from both the low and high piece rate work periods. The resulting utility function is used to predict each worker’s output during the dismissal periods. Observed output is significantly higher than predicted output. This result holds true even when incorporating risk preferences into the output predictions, using workers’ incentivized decisions over gambles to recover their CRRA parameters. Finally, analysis of output over time within the same work period under the same incentives is presented to show that learning to perform the task more efficiently is not driving this conclusion.

Recall that the marginal benefit of output under all job-period combinations under dismissal threats is below the high piece rate of $0.20 (as shown in Table 1 above). This calculation takes into account the option value of being retained. This implies that under reasonable theoretical assumptions, the optimal level of output for workers under dismissal threats should be below the level of output under the high piece rate. This is not the pattern of behavior that is observed.

Result 1.1: Extensive Margin. Dismissal threats are more motivating than piece rates: Subjects are more productive under dismissal threats even though the monetary benefit of output is lower.

The vast majority of subjects in both dismissal periods 1 and 2 increase their productivity compared to the high piece rate period. Doing a simple within-subject comparison, 76.2% of subjects increase productivity between the high piece rate period and the first dismissal period, while 79.7% of retained subjects do the same in the second dismissal period. In terms of magnitude, the mean change in productivity between the high piece rate period and first dismissal period is a 23.6% increase in tasks per minute (the equivalent increase in productivity is 32.9% in the second dismissal period among retained workers).

Figures 3A and 3B display CDFs of output showing this comparison. Figure 3A displays the CDFs of high piece rate output side-by-side with the first 10-minute dismissal period;
Figure 3B shows the same comparison with the second dismissal period. In both cases, the CDF of output in the dismissal periods is clearly to the right of the CDF of piece rate output. Hypothesis tests confirm this visual evidence. Signed-rank tests and t-tests reject the null hypotheses that output in the first and second dismissal periods are equal to output in the high piece rate period (all with \( p < 0.0001 \)).

Keep in mind another consideration working against this finding: economists typically assume that cost of effort functions are increasing and convex. This implies that subjects would reach a higher marginal cost of effort during a 10-minute period without breaks than during a 5-minute period. Thus even under constant marginal incentives, output during a 10-minute work period would be predicted to be lower than double output during a 5-minute work period.

Parametric analysis was also performed to predict output during the dismissal periods, with the primary purpose being to show that risk aversion alone cannot explain the increased output. This requires making assumptions about the functional form of the cost of effort. The cost of effort function was assumed to be increasing and convex, as is the convention in theoretical literature on principal-agent models. It was also assumed that the cost of producing zero units of effort is zero. (Although effort is unobservable, one unit of effort is normalized as the amount required to create one unit of output, without loss of generality.) To satisfy these assumptions with the simplest functional form possible, a quadratic form is assumed: \( C(e) = ae^2 + be \). Additionally, utility is assumed to be additively separable in money and the cost of effort: \( U = y - C(e) \), where \( y \) is total pay from that work period.

The parameters of the cost of effort function are calibrated using observed output from both piece rate work periods along with the known marginal benefit of output in each period. This results in a system of two equations with two unknown effort cost parameters for each subject: \( a \) and \( b \). This setup is only rationalizable for subjects who increase output in response to a higher piece rate, so only the 80.2% of subjects who do so are included in the following analysis.

To account for the difference in the length of the periods (5 minutes under piece rates versus 10 minutes under dismissal threats), the amount of output observed in the piece rate periods was doubled before calibrating the cost of effort function, giving the best possible chance to the null hypothesis that predicted dismissal output is greater than or equal to observed output.\(^{12}\)

\(^{12}\)Again, piece rate output is doubled for this hypothesis test to make a fair comparison between a 5-minute period and a 10-minute period.

\(^{13}\)Doubling piece rate output before calibrating the cost of effort function results in larger dismissal output predictions than the alternative of doubling the output predictions after they are made. This is due to the assumed convexity of the cost of effort function. However, both versions of this prediction exercise were run,
Figure 3A

CDF of Output, Piece Rate vs. 1st Dismissal

Figure 3B

CDF of Output, Piece Rate vs. 2nd Dismissal
The calibrated cost of effort function was then used to predict the output of each worker in the first and second dismissal work periods, given the job they were assigned. This exercise was first conducted assuming risk neutrality, and the results are displayed in Figures 4A and 4B. Also note that only subjects who were retained after period 1 are included in Figure 4B, further restricting the sample to 47.5% of subjects.

Incorporating risk aversion in the model predicts higher effort than assuming risk neutrality. The data exercise incorporates risk aversion by using the decisions over gambles made by subjects in the second part of the experiment. These decisions elicited a certainty equivalent for a 50/50 gamble between $0 and $6. This certainty equivalent is used to calibrate a CRRA parameter for each subject. These risk preferences are incorporated into the expected utility function of each subject, and the optimal levels of output in dismissal periods 1 and 2 are recalculated. To give the null hypothesis that predicted output equals observed output the best possible chance, the following analysis only includes subjects who exhibited risk averse preferences over the gamble options (i.e., risk-loving subjects are excluded). Combined with the other sample restrictions, 51.5% of subjects are included in Figure 5A and 25.7% of subjects are included in Figure 5B.

Whether one assumes risk neutrality or incorporates the measured risk aversion of subjects, the results of the parametric analysis are the same: subjects produce much more under dismissal threats than any rational neoclassical model would predict given their behavior throughout the rest of the experiment. Hypothesis tests strongly echo the visual evidence displayed in Figures 4 and 5. Sign-rank tests and t-tests reject the null hypotheses that first- or second-period dismissal outputs are equal to predicted output from either the risk neutral or the risk averse models (all p-values < 0.001).

**Result 1.2: Risk Aversion.** Parametric analysis suggests that risk aversion does not explain the high productivity under dismissal threats.

An important potential criticism to address is whether the above results are caused merely by subjects learning to complete the task more efficiently over time. For example, the first dismissal period comes after the high piece rate period, so it could be the case that subjects are producing more output because they have more experience rather than because dismissal threats are highly motivating. (Alternatively, exhaustion could reduce output over time, but this is not a plausible potential explanation for the results shown above because output and the results were qualitatively similar (the distribution of predicted output is always to the left of the distribution of observed output, and no hypothesis tests change significance levels).

\footnote{Figure 5A includes subjects who are risk averse and increased output in response to the piece rate increase. Figure 5B includes subjects who are risk averse, increased output in response to the piece rate increase, and were retained after dismissal period 1.}
Figure 5A

CDF of Output, 1st Dismissal Period
Predicted (Risk Averse) vs. Observed

Figure 5B

CDF of Output, 2nd Dismissal Period
Predicted (Risk Averse) vs. Observed
is higher in later periods as dismissal threats are introduced.) Comparing worker output between discrete periods with different levels and types of incentives will not shed light on how much learning occurred because incentives are changing along with worker experience.

The best way to analyze worker learning during the real effort task is to compare output over time within the same work period. If workers are truly improving their ability during the work task, a 10-minute period would likely be enough time to pick up a difference in productivity. The experiment software stored the exact time at which each work task page was submitted, allowing analysis of how many tasks were submitted in each minute of every work period. Figures 6A and 6B plot mean worker output over time for both 5-minute piece rate periods and for the first two 10-minute dismissal work periods. (The third dismissal period is less informative about learning because a large proportion of retained subjects shirk.)

The figures show that apart from the first minute, output is relatively constant over time within each work period. However, there is a strong mechanical reason to expect output to increase from the first to the second minute. In the first minute, workers must start a fresh task; in all subsequent minutes, there is typically a task in progress that carries over from the previous minute. Another trend observed in the data explains why output slightly increases in the final minute: many subjects try to submit a final answer just before the timer ticks down to zero, even if that answer is incorrect.

Keeping these mechanical trends in mind relating to the first and last minutes of the period, hypothesis tests were run on the null hypothesis that output in the first half of the work period (omitting the first minute) is equal to output in the second half of the work period (omitting the final minute). Both rank-sum tests and t-tests were run on all four work periods, and every test failed to reject the null at a 95% confidence level. The conclusion from the figures and hypothesis tests is that neither learning to be more efficient at the task nor exhaustion played important roles in determining worker output.

**Result 1.3: Learning.** Subjects do not increase productivity over time, suggesting that learning does not explain the high productivity under dismissal threats.

Another alternative hypothesis is that subjects worked very hard under dismissal threats either due to intrinsic motivation or experimenter demand effects, but this explanation does not hold up to scrutiny after examining worker output during the final dismissal period with full job security. Because all fixed payments are up-front and there is no fourth dismissal period, there is no monetary incentive to work at all during the third period. Thus, the possible motivations to produce output during this period were only intrinsic motivation.
Figure 6A

Figure 6B
(i.e., enjoyment of the work task), experimenter demand effects, or a desire to avoid a public announcement of one’s dismissal at the end of the period.

Most workers who made it to the final tenured work period produced lower output than in previous periods. Of the 39 workers (out of 101 total) who made it to the final period, 71.8% of them produced fewer units of output than in period 2. Mean output in period 2 among workers who would become tenured was 34.5 units, and mean output in period 3 dropped to 20.6 units. This decline in output is highly statistically significant (signed-rank test, p = 0.0002). Some of the tenured workers displayed severe levels of shirking: 25.6% of tenured workers produced 0 units of output, while 43.6% of them produced 5 units or fewer. Keeping in mind that these are systematically the most productive subjects in the experiment, these levels of output are quite low.

The final period turned off most but not all of the behavioral motivations involved in firing threats. There was no reason to work to avoid further social exclusion, jealousy of coworkers who would continue to earn money, or any potential self-image impacts of dismissal (the signal of being fired is uninformative if the worker was not exerting effort in the first place). Because the public announcement of firing remained, if anything the productivity drop-off is an overestimate of any remaining demand effects or intrinsic motivation that might have been driving the high productivity in periods 1 and 2. The reasonable conclusion is that the high productivity when all behavioral and monetary mechanisms were active was due to the combined strength of these motivators, and not merely due to intrinsic motivation or experimenter demand effects.

**Result 1.4: Intrinsic Motivation.** Subjects are significantly less productive in the final dismissal period, suggesting that intrinsic motivation does not explain the high productivity under dismissal threats.

The above analysis suggests that dismissal threats produce a large extensive margin effect on productivity which cannot be rationalized by subject responses to piece rate changes or risk preferences. Although the vast majority of subjects produce more output under dismissal threats than the parametric model with risk aversion predicts, the magnitude of this difference varies considerably by subject. An interesting follow-up question is which if any of the measured behavioral traits of subjects predicts the size of this residual. In other words, do the behavioral traits of workers predict how strongly they will react to the existence of dismissal threats?

Table 2 addresses this question by regressing the output residual on worker behavioral traits. The output residual for each worker is defined as observed output in the first dismissal period minus output predicted by the parametric model with risk aversion. Given
the parametric model’s restrictions, only subjects who increase output between the low piece rate and high piece rate work periods are included. Additionally, only risk averse subjects are included. The variables for worker behavioral traits include incentivized measures of loss aversion, patience, overconfidence, altruism, trust, and reciprocity.

**Table 2:** Linear Regression of Output Residual on Worker Traits

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output Residual</td>
<td></td>
</tr>
<tr>
<td>Loss Choice</td>
<td>0.725</td>
</tr>
<tr>
<td></td>
<td>(0.612)</td>
</tr>
<tr>
<td>Patience Choice</td>
<td>2.620*</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
</tr>
<tr>
<td>Confidence: Actual - Reported Rank</td>
<td>0.339</td>
</tr>
<tr>
<td></td>
<td>(0.331)</td>
</tr>
<tr>
<td>Social Prefs: Altruism</td>
<td>-0.734</td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
</tr>
<tr>
<td>Social Prefs: Trust</td>
<td>0.639</td>
</tr>
<tr>
<td></td>
<td>(0.540)</td>
</tr>
<tr>
<td>Social Prefs: Reciprocity</td>
<td>0.362</td>
</tr>
<tr>
<td></td>
<td>(0.800)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.491</td>
</tr>
<tr>
<td></td>
<td>(0.826)</td>
</tr>
<tr>
<td>Observations</td>
<td>52</td>
</tr>
</tbody>
</table>

*Notes: The table shows the results of an OLS regression. The dependent variable is observed output in the first dismissal threat work period minus output predicted by the parametric model including risk aversion. The subsample only includes subjects who increase output under a higher piece rate and only subjects who display risk aversion in decisions over gambles. Standard errors are clustered at the session level; p-values are shown in parentheses.

* p < 0.05, ** p < 0.01, *** p < 0.001

The results of this regression show that patience is a significant predictor of the extensive margin effect of dismissal threats. The patience choice variable represents how much money a subject would be willing to accept in one week instead of receiving $3 today, so subjects with a higher value are more impatient. The coefficient implies that a subject who requires $1 additional dollar received in one week will produce 2.6 additional units of output under dismissal threats than the parametric model predicts.
Result 1.5: Behavioral Predictors of Extensive Margin. Subjects who are more impatient display larger extensive margin productivity gains under dismissal threats.

This subsection has established that the mixture of monetary and behavioral motivators invoked with dismissal threats in this experiment were highly motivating. Output under dismissal threats was significantly higher than would be predicted under a variety of assumptions and methods, all tilted to give the alternative hypothesis the best chance whenever possible. The results cannot be explained by the level of risk aversion observed in decisions over gambles in the experiment. They also cannot be explained by learning because subjects do not improve output even over the course of long 10-minute work periods while incentives remain constant. Instead, it appears that dismissal threats are a much stronger motivating incentive device than piece rates.

4.2 Intensive Margin: The Incentive and Selection Effects of Stronger Dismissal Threats

This subsection decomposes the productivity gains from stronger dismissal threats into the pure incentive effect from a higher marginal benefit of output and the selection effect from attracting more productive workers. (Note that the following analysis will only involve output from the first work period with dismissal threats unless stated otherwise.) Recall that the firing chance reduction per unit of output is twice as high in the Risky Job: a 2% decline in the chance of being fired per unit of output, as opposed to only a 1% decline in the Safe Job. Additionally, subjects in the Risky Job have more to lose by being fired because the wage per period is twice as high. These steeper incentives suggest that output should be higher in the Risky Job.

Result 2.1: Intensive Margin. Workers are more productive under stronger dismissal threats with higher pay than under weaker dismissal threats with lower pay.

Indeed, mean output of workers who preferred the Risky Job and were assigned the Risky Job was 30.9 units, while mean output of workers who preferred the Safe Job and were assigned the Safe Job was only 25.8 units. A Wilcoxon rank-sum test rejects the null hypothesis that these samples of output levels come from the same distribution ($p = 0.028$). However, comparing output between these two groups only reveals the combined effect of both the incentive effect and the selection effect for the subset of subjects who were assigned their preferred job.

To decompose the selection and incentive effects, each of the four job preference-assignment combinations will be exploited in data analysis. In terms of job preference, there was a fairly
even split: 60.4% of subjects preferred the Risky Job and 39.6% of subjects preferred the
Safe Job. Seventeen subjects preferred the Safe Job and were assigned the Safe Job; 23
subjects preferred the Safe Job and were assigned the Risky Job; 25 subjects preferred the
Risky Job and were assigned the Safe Job; and 36 subjects preferred the Risky Job and were
assigned the Risky Job.

Result 2.2: Incentive and Selection Effects. The aggregate productivity gain from stronger
dismissal threats is caused largely by the selection effect, not the incentive effect.

The story is clear whichever method of analysis is used: the aggregate incentive effect
of stronger dismissal threats is small and statistically insignificant, but the aggregate se-
lection effect on output is large and significant. Mean output is 25.7 among workers who
are assigned the Safe Job, while it is 26.6 among workers who are assigned the Risky Job
(statistically insignificant with a t-test). Thus there appears to be little difference in mean
output indicating a pure incentive effect from stronger dismissal threats. Mean output is
22.3 among workers who prefer the Safe Job and 28.8 among workers who prefer the Risky
Job (with a t-test p-value of 0.003). Therefore, the selection effect on mean output is both
statistically significant and large in magnitude (roughly a 30% increase in output between
job preferences).

Figures 7A and 7B visually display the output differences caused by the incentive and
selection effects by showing the CDFs of output by either job preference or job assignment.
Figure 7A shows roughly overlapping distributions of output between job assignments, indi-
cating little or no incentive effect. A Wilcoxon rank-sum test does not indicate a statistically
significant difference between these two distributions of output. On the other hand, Figure
7B shows a clear separation of the distribution of output between job preferences: the dis-
tribution of output among those who prefer the Risky Job is clearly to the right of the
distribution for those who prefer the Safe Job. This difference is highly statistically signifi-
cant (rank-sum test: p = 0.001).

These aggregate results hide substantial heterogeneity in subject responses to stronger
dismissal threats. It seems that subjects who prefer the Risky Job do produce more output
when randomly assigned the steeper incentives, although this result is only marginally sig-
nificant. On the other hand, subjects who prefer the Safe Job but are thrust into the steep
incentives of the Risky Job actually perform worse than their counterparts who prefer the
Safe Job and are assigned the Safe Job.

Result 2.3: Heterogeneity in Incentive Effects. Subjects who prefer the Safe Job are
less productive when randomly assigned to the Risky Job, suggesting that imposing strong
dismissal threats on workers who do not prefer them can backfire.
Figures 8A and 8B illustrate this point by plotting the distribution of output between job assignment while only including a subset of subjects with the same job preference. Figure 8A only includes subjects who prefer the Safe Job and plots the distribution of output separately by job assignment. Notice that the expected placement of the distributions is reversed: the distribution of output under the steeper incentives in the Risky Job is actually further to the left among this subset of subjects. This difference in output is statistically significant using both a t-test (p = 0.02) and a rank-sum test (p = 0.025).

This negative incentive effect may seem surprising at first, but many of these workers were placed into a job in which they were not expected to perform well. The median output of workers who preferred the Safe Job during the high piece rate period was only 8 units in 5 minutes. A worker who maintained this same level of output per minute during an entire work period in the Risky Job would end up with a dismissal probability of 68%. In fact, fully 85% of workers who preferred the Safe Job would end up with a dismissal probability of greater than 50% even when maintaining their output-per-minute from the high piece rate period. A logical conclusion is that placing workers under dismissal threats that are too strong compared to their ability level can actually be demotivating.

Figure 8B shows suggestive evidence that the incentive effect of stronger dismissal threats was positive among the subset of subjects who did prefer the Risky Job. The distribution of output under the Risky Job incentives is clearly to the right of the distribution of output under the Safe Job incentives, but this difference between distributions is only marginally significant (t-test: 0.078; rank-sum test: 0.079).

Although these graphs and hypothesis tests are visually useful and do tell most of the story, one must keep in mind that they do not fully isolate the selection and incentive effects because the random assignment to incentives did not occur with 50% probability. For example, in the plots of output distributions illustrating the incentive effect, roughly 60% of the subjects working in the Risky Job will also have preferred the Risky Job due to the random job assignment design mechanism. As a result, regressions are run to fully isolate the impact of job preference versus job assignment.

Table 3 displays the results of regressions which isolate the incentive and selection effects on output. These specifications are simple OLS regressions with output during the first 10-minute dismissal work period as the dependent variable. Standard errors are clustered at the session level throughout. Column 1 includes all subjects and includes two indicator variables as covariates: whether the subject was assigned the Risky Job, and whether the subject preferred the Risky Job. The coefficient on being assigned the risky job is small and statistically insignificant, indicating that there were not strong incentive effects despite facing a much larger marginal benefit of output in the Risky Job compared to the Safe
Figure 8A

CDF of Output, by Job Assignment, Only Prefer Safe

Figure 8B

CDF of Output, by Job Assignment, Only Prefer Risky
Job. In contrast, the selection effect on output is large and statistically significant. After controlling for job assignment, indicating a preference for the Risky Job is associated with a 6.4-unit increase in output ($p < 0.01$). This magnitude is large enough to be economically impactful, given that mean output in the first dismissal period was 26.2 units.

**Table 3:** Linear Regression of Output under Dismissal Threats on Job Preference and Assignment

<table>
<thead>
<tr>
<th></th>
<th>(1) All Subjects</th>
<th>(2) Only Prefer Safe</th>
<th>(3) Only Prefer Risky</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assigned Risky Job=1</td>
<td>0.725</td>
<td>-6.084*</td>
<td>5.237</td>
</tr>
<tr>
<td></td>
<td>(0.729)</td>
<td>(0.018)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>Prefer Risky Job=1</td>
<td>6.434**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>21.91***</td>
<td>25.82***</td>
<td>25.68***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>101</td>
<td>40</td>
<td>61</td>
</tr>
</tbody>
</table>

*Notes:* The table shows the results of OLS regressions. The dependent variable is output during the first 10-minute work period with dismissal threats. Column 1 includes all subjects. Columns 2 and 3 include only subjects who preferred the Safe Job or the Risky Job, respectively. Standard errors are clustered at the session level; p-values are shown in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regressions were also run to understand the heterogeneity in incentive effects between subjects with different job preferences. Columns 2 and 3 of Table 3 show the results of regressing output on being assigned the risky job among only subjects who preferred the Safe Job and only subjects who preferred the Risky Job, respectively. These columns confirm the results of the previous CDF plots: among subjects who preferred the Safe Job, being assigned the Risky Job is associated with about a 6-unit reduction in output ($p = 0.018$). There is again suggestive evidence that the incentive effect of stronger dismissal threats is positive among subjects who preferred the Risky Job, although the coefficient is not statistically significant at the 5% level ($p = 0.103$).

The analysis so far has focused on productivity during the first period under dismissal threats, but retention over time also plays an important role in the selection effect in the real labor market. More productive worker are systematically retained in jobs with strong dismissal threats, and this of course occurs in this experiment as well. The mean output in period 1 was 26.2 units, while mean output among retained workers in period 2 was 31.1 units. Additionally, more workers are fired from the Risky Job due to the higher chance of
being fired at all levels of output: 58.4% of workers start out in the Risky Job, but by period 2 only 49.2% of retained workers are in the Risky Job. This confirms that the selection out of lower-productivity workers is an important element of the long-term impact of strong dismissal threats on productivity.

The lack of aggregate pure incentive effects from being randomly assigned stronger dismissal threats is an interesting behavioral puzzle. It appears that once the mixture of monetary and behavioral motivators of dismissal threats is activated, subjects generally work their hardest regardless of adjustments on the margin to the dismissal probability function. A notable exception to this is that when the dismissal threats are too strong for the ability level of a worker, they can become demotivating.

Analyzing the differences in output between periods 1 and 2 also speak to this lack of responsiveness to changes in dismissal threats on the margin. Any model of rational worker behavior predicts that workers should reduce output from period 1 to period 2. This is because in the first period, there are two remaining payments to earn, while in the second period, only one potential payment remains. However, this does not occur. Restricting to the sample of workers who were be retained after period 1, their mean output was 28.7 in period 1 and 31.1 in period 2. It appears that at the parameters explored in this experimental design, workers were far more sensitive to the mere presence of credible dismissal threats than they were to the marginal strength of dismissal threats.

Additional analysis was conducted to determine whether selection into dismissal threats occurred on worker traits other than productivity. Although some univariate differences were detected between workers who preferred the Safe Job and the Risky Job, these traits were all correlated with productivity. For example, workers who preferred the Risky Job tended to be more confident but also more accurate in those beliefs. However, both confidence and accuracy of beliefs were highly correlated with productivity in the high piece rate work period. A probit regression of preference for the Risky Job on all incentivized worker traits results in productivity being the only trait which is significantly predictive of job preference. The full explanation of this analysis including tables and regression specifications may be found in the appendix.

5 Discussion

This paper makes two significant contributions to the literature on employee motivation and selection. First, the experiment design imports the threat of being fired to the lab in a clearly defined probabilistic fashion, linking worker output to the chance of being fired. This brings tractability to an incentive that is ubiquitous in the modern labor market, but
rarely explicitly quantified in labor contracts and thus difficult to study in the field. This tractability allows for clean comparisons between dismissal threats and other monetary incentives such as piece rates, along with the ability to measure worker responses to exogenous, quantifiable changes in the strength of dismissal threats. Further, the design purposefully activates many of the behavioral mechanisms concomitant with the threat of being fired: the loss framing, the fear of exclusion from a social group, the embarrassment and self-image implications of public announcements of firings, and the jealousy of retained coworkers who will continue to work and earn money. These design elements produce a rich implementation of dismissal threats in the lab, enabling detailed analysis of this unique and understudied type of incentive.

Second, the design includes a mechanism which allows the clean decomposition of productivity gains into the pure incentive effect versus the selection effect. This mechanism is unique in the literature on worker selection, but quite simple to understand for both researchers and subjects: workers select an incentive scheme to improve the odds that they are randomly assigned to that scheme. This retains incentive compatibility because self-interested subjects are motivated to honestly reveal their preferences over incentive schemes. Crucially, it also allows researchers to observe counterfactual work arrangements that are difficult to observe in the field: subjects working under the incentive scheme they did not prefer. The previous literature has almost entirely opted for either guaranteed assignment to one’s preferred incentives, or fully random assignment, which each have drawbacks compared to this method. By eliciting worker preferences over incentives and randomly assigning them to incentives within the same round, this mechanism saves time and avoids order contamination effects of previous approaches in the literature. This mechanism is ready to import into other experiments on selection into a wide variety of incentive schemes such as competitions, teams, goal-setting, commitment contracts, and more.

The results paint a clear picture: on the extensive margin, dismissal threats are highly motivating as an incentive device. Introducing credible dismissal threats results in large productivity gains over the most basic monetary incentive of piece rates. The incentive schemes in the experiment were designed such that the monetary benefit of output under dismissal threats was always lower than the high piece rate. However, the vast majority of subjects increase productivity under dismissal threats compared to their own productivity under the high piece rate. The mean increase in productivity is 23.6% in the first dismissal work period, an economically important magnitude. Additionally, parametric analysis shows that this result is unlikely to be caused purely by risk aversion. Although incorporating subject risk preferences does increase predicted output under dismissal threats compared to a risk neutral benchmark, risk aversion alone does not explain the large increase in productivity.
induced by dismissal threats.

Turning to the intensive margin effects of steeper incentives, stronger dismissal threats result in higher output than weaker ones, as predicted. The interesting results appear in the decomposition of these productivity gains into the pure incentive effect versus the self-selection effect. Analysis of the data clearly shows that the aggregate productivity gains from stronger dismissal threats come largely from selection effects. Workers who prefer higher pay with a steeper firing probability function are significantly more productive than other workers. However, being randomly assigned to these steeper incentives appears to have little effect on productivity on average.

A likely explanation for this result is that the mixture of financial and behavioral motivators invoked under firing threats was so strongly motivating that effort was not very sensitive to marginal changes in the pure financial incentive. In other words, as long at least a moderate threat of being fired is present, along with the social, self-image, and loss-framing motivators, subjects will put forth high levels of effort. If this is the case, introducing credible firing threats is more powerful than making modifications to existing firing threats on the margin. Further evidence for this interpretation is the fact that retained workers did not show declines in productivity in the second dismissal period, despite the decline in the marginal benefit of output due to fewer payments remaining to be collected.

However, the lack of aggregate incentive effects of steeper dismissal threats masks important heterogeneity. Subjects who preferred higher pay along with stronger dismissal threats did produce higher output on average when randomly assigned to this incentive scheme (however, this difference was not statistically significant). But subjects who were willing to accept lower pay to avoid strong dismissal threats are significantly less productive when assigned to strong dismissal threats. Clearly, these workers were trying to self-select into the incentive scheme under which they expected to perform better. In a sense, the negative and positive incentive effects among these two subsets of subjects cancel out in the aggregate results.

These results contain important implications for firms and policymakers. Firms that rely largely on monetary incentives other than dismissal threats may experience large productivity gains by shifting to moderate-strength dismissal threats. Firms with existing credible dismissal threats may consider strengthening them to improve productivity, but they should realize that these effects are likely to occur over time as the composition of workers changes through hiring and retention decisions. It is important to avoid imposing dismissal threats.

\(^{15}\)This does not contradict the results on the extensive margin. Even the weaker dismissal threats implemented in this experiment resulted in significantly higher productivity than the high piece rate. This was true among both the workers who preferred weaker dismissal threats and the ones who preferred stronger dismissal threats.
threats that are too strong for the ability level of a worker given the evidence that this can backfire and cause negative incentive effects. This is an important consideration when imposing stronger uniform dismissal threats on an existing workforce of mixed ability levels. Policymakers considering implementing widespread employment protection may expect to experience muted productivity losses given that the aggregate effect operates largely through selection, but this hides differential productivity impacts on workers with different preferences over dismissal threats.

There are a number of potentially fruitful avenues for future research. In this experiment design, only two levels of pay and dismissal threats were offered. This was necessary to maintain incentive compatibility of the job preference elicitation and allow a simple decomposition of the selection effect from the incentive effect. However, it seems quite possible that a more complex design that offered many different levels of pay, baseline chance of firing, and relationship between output and firing chance could uncover much richer evidence on worker preferences over dismissal threats. Other modifications to the framing of firing threats might be considered, such as displaying the chance of retaining one’s job rather than the chance of losing it.

This paper has intentionally activated many of the behavioral motivators of firing threats in the lab, yet still more mechanisms could be explored in future research. The present design included a strong loss framing, public notifications of firing, the fear of exclusion from a group of coworkers, and jealousy of retained workers who continue to earn money. Firing threats in the real labor market may be even more motivating when the rejection comes from a human boss with discretion compared to an impersonal mass layoff. Social motivators may be stronger when worker identities are revealed and social connections are developed, rather than the anonymous setting implemented in this design. It also seems possible that a sense of ambiguity plays an important role in the field as well – workers never quite know precisely what it takes to avoid being fired. This ambiguity might produce even stronger motivation for workers than a known, quantified link between productivity and the probability of being fired.

The results of this experiment suggest that the threat of being fired is a highly motivating incentive device, and this may partially explain why this form of incentive is so ubiquitous in the modern labor market compared to simpler incentives such as piece rates. The behavioral motivators activated in this lab experiment did not operate at the full strength at which they exist in the real labor market. This suggests that the productivity gains from dismissal threats estimated in this paper are a lower bound on the motivating effect they

\[\text{Without a clever design solution, using a human subject as a “boss” involves losing control over the ability to quantify and manipulate the strength of dismissal threats.}\]
may have in the field. A fruitful avenue for future research would be to explore which of the various behavioral components of firing threats are the most powerful and why. This sort of decomposition of the motivating mechanisms of firing threats would be fascinating for labor and personnel economists and informative for management teams trying to motivate their employees.
References


A  Real Effort Task

In the real effort task used in this experiment, subjects are shown a randomly generated image with 24 regular polygons (3 rows of 8 shapes each). Each shape may have 7, 8, 9, or 10 sides. In each of the 24 spots for shapes, each of these four polygons is equally likely to appear. Additionally, every shape is rotated a random amount ranging from 0 to 360 degrees. The real effort task for the subjects was to count the number of 7-sided shapes in each image. Although the answer could theoretically range from 0 to 24, in practice most answers were between 4 and 8.

All work periods were timed. To penalize guessing, when a subject submitted an incorrect answer, it did not count as a completed task and a 10-second delay occurred before the next image appeared.

Importantly, each subject was shown the same randomly generated images in the same order as every other subject within the same work period. That means there cannot be subjects who face particularly difficult or easy images compared to other subjects in the same work period.

A screenshot of the work interface during the low piece rate period is shown below.
B Worker Traits

B.1 Measurement Procedures

Part 2 of the experiment involved a series of incentivized decisions to measure worker traits along with a set of unincentivized survey questions. Only one incentivized task or decision was randomly selected for payment from Part 2, and this was explained to subjects once they got to this part of the experiment. Some tasks were single-player decisions, while others were two-player games. Subjects were told that only one task would be randomly selected for payment, but that if a two-player game was selected, that game would also be selected for a counterpart subject in the experiment and the payoffs would be implemented for both players. An algorithm was designed to make it as close to equally likely as possible for each task to be selected for payment, under the constraint that all two-player decisions selected required a counterpart assigned to the same game.

Standard methods were used to elicit the following worker traits in an incentive-compatible way: confidence, performance under high stakes, risk aversion, loss aversion, patience, altruism, trust, and reciprocity. Confidence was measured by asking subjects to predict the rank they earned in a group of 10 randomly selected subjects; if the rank they reported was within 1 rank of the truth, they would be paid $5. Subjects faced a high-stakes work period in which they completed the same shape-based work task from earlier in the experiment. They would earn $5 if they completed 4 tasks within 45 seconds. To increase feelings of pressure and stress, the work period timer was significantly enlarged and presented with a bright red background at the top of the screen.

Risk aversion, loss aversion, and patience were all elicited using a variant of the Multiple Price List method called the staircase procedure (see Cornsweet, 1962, for an early definition). In this method, the subject is presented with a series of binary decisions that adjust the options presented in subsequent screens based on past decisions. For example, the initial choice in the risk aversion elicitation is between a 50/50 gamble for $0 or $6, or taking a certain amount of $2.90. If the subject selects the gamble, the certain amount increases for the next offer. If the subject selects the certain amount, the certain amount decreases for the next offer. This staircase method allows precise measurements of subject preferences while requiring much fewer decisions than a standard Multiple Price List would require. For each of these three trait measurements, subjects made 5 decisions but could end up with any of $2^5 = 32$ potential values. The parameters of the staircase choice sets were constructed similarly to those in Falk et al.’s (2018) global survey of economic preferences.

Risk aversion was measured with a series of decisions between a 50/50 gamble for winning $6 or winning $0, or taking a variable certain amount. The final value of this certainty
equivalent could range from $0.10 to $6.30. The loss aversion decision was between a 50/50 gamble for winning $5 or losing a variable amount of money, or staying at $0 for this task. The final value of the potential loss could range from $0.10 to $6.30. The patience decision was between being paid $3 today or a variable amount in one week. The final value of the variable amount could range between $2.90 and $6.00. Any future dollar amounts selected for payment from this task were actually paid out on Venmo exactly one week after the session ended.

Subjects played two different two-player games, and the strategy method was used to observe the decisions of each subject in both games and both roles (sender and receiver). First, subjects played a Dictator Game in which one player received $8.00 and could split any amount of it with another player who received $0 (Forsythe et al., 1994). Second, subjects played a Trust Game in which each player started with $3 (Berg, Dickaut, and McCabe, 1995). The sender could send a restricted choice set of $0, $1, $2, or $3 to the receiver, and any amount sent would be tripled. The receiver then input any amount (including decimals) from their new balance to send back to the sender for each potential amount initially sent. The strategy method for the receiver role involved all subjects inputting a dollar amount for each of the four potential amounts initially sent by the sender, allowing analysis of the relationship between the amount sent and the amount returned.

A series of unincentivized survey questions were also included in Part 2 to measure other worker traits. Subjects were asked to report the level of stress they felt during the high-stakes work period, along with whether they work well under pressure. They were also asked two questions about preferences for competition: how much they agree with the statement that “Competition brings out the best in me,” and whether they agree more with the statement that “Competition is harmful. It brings out the worst in people,” or “Competition is good. It stimulates people to work hard and develop new ideas.” The phrasing of these questions comes from Fallucchi, Nosenzo, and Reuben (2020) who find that these two survey questions are the most predictive of incentivized measurements of preferences to compete. All of these questions were asked on a scale of one to seven. Finally, a short version of the Big 5 personality trait test was administered (Rammstedt and John, 2007). Subjects answered two questions corresponding to each trait which were combined to create a score from one

---

17 To prevent the possibility of subjects earning below the minimum payment of $5, subjects were told that $3 would be added to their earnings only if this task was the one selected for payment. Although this could interfere with the loss framing, it was presented in a less salient manner on a separate screen; the actual gains and losses displayed on the decision page were shown in bold green or red text and did not take into account this extra payment. Based on subject decisions, this appeared to have been successful in invoking the loss framing even when the $3 payment would have brought them out of the loss domain in this task.

18 In a survey of experimental methodology, Brandts and Charness (2011) do not find strong evidence of the strategy method changing subject behavior.
to five. Finally, subjects were asked to report a few basic demographic variables including age, gender, race, and ethnicity.

At the end of the experiment, total earnings were displayed to subjects including a breakdown of how much they earned from each part of the experiment, and which decision was selected for payment in Part 2. Nothing was explicitly reported to subjects about cumulative earnings up until this point, when all subjects had completed all decisions in the experiment.

B.2 Descriptive Statistics

A short description of worker traits and demographics will be provided in this subsection. The mean and standard deviation of each incentivized trait and survey measure are displayed in Appendix Tables 1A and 1B, respectively. Subjects reported a mean rank in the confidence measurement of 4.93, and they were on average slightly overconfident (overestimating their placement by 0.83 ranks). Note that the confidence measurement is a rank out of 10, with rank 1 being the most productive, so a lower number corresponds to higher confidence. The mean certainty equivalent for a 50/50 gamble between $0 and $6 was $3.04, but this mean is increased by a small number of outliers who appeared highly risk-loving. The median certainty equivalent was $2.90, and 61.4% of subjects display risk aversion over these small dollar amounts. Subjects show loss aversion on average because the mean acceptable 50/50 gamble involving the loss domain was between winning $5 or losing $2.54. In fact, 98% of subjects rejected gambles involving losses with expected values greater than $0. With regards to patience, many subjects appeared to want to maximize the dollar amount received without regard to timing. Nearly half of subjects (47.5%) were willing to forgo $3 today to receive any amount greater than or equal to $3.05 in 1 week. However, no subject accepted $2.95 in 1 week over $3 today, suggesting that subjects were attentive and not simply clicking one option repeatedly without thought. Because of this high level of patience among subjects, there was a strong mode in the patience variable and a small standard deviation. However, among the 52.5% of subjects who were not simply selecting the higher dollar amount regardless of timing, there was much more dispersion in patience, ranging from accepted future values of $3.10 to the maximum of $6.
**Appendix Table 1A: Descriptive Statistics (Incentivized Traits)**

<table>
<thead>
<tr>
<th>Worker Traits</th>
<th>Mean/SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Piece Rate Tasks</td>
<td>10.61</td>
</tr>
<tr>
<td>Responsiveness to Piece Rate</td>
<td>0.82</td>
</tr>
<tr>
<td>High Stakes Tasks</td>
<td>2.44</td>
</tr>
<tr>
<td>Confidence: Rank Choice</td>
<td>4.93</td>
</tr>
<tr>
<td>Confidence: Actual - Reported Rank</td>
<td>0.83</td>
</tr>
<tr>
<td>Risk Choice</td>
<td>3.04</td>
</tr>
<tr>
<td>Loss Choice</td>
<td>2.54</td>
</tr>
<tr>
<td>Patience Choice</td>
<td>3.40</td>
</tr>
<tr>
<td>Social Prefs: Altruism</td>
<td>2.40</td>
</tr>
<tr>
<td>Social Prefs: Trust</td>
<td>1.99</td>
</tr>
<tr>
<td>Social Prefs: Reciprocity</td>
<td>1.46</td>
</tr>
<tr>
<td>Observations</td>
<td>101</td>
</tr>
</tbody>
</table>

Notes: The table shows the mean value of each trait for subjects in the experiment. Standard deviations are shown in parentheses.

Measures of altruism, trust, and reciprocity were calculated from the two-player games. The mean contribution in the Dictator Game was $2.40, and 24.8% of dictators sent nothing. High levels of trust were observed by senders in the Trust Game: the mean amount sent
was $1.99, and 42.6% of subjects sent the maximum amount of $3. Reciprocity is not as straightforward to summarize given that it is a function of four different choices by the receiver – one for each potential dollar amount initially sent. To measure reciprocity, a linear relationship was calculated between the amount sent and the amount returned. This was accomplished with an OLS regression run separately for each subject with the amount returned as the outcome variable and the amount initially sent as the explanatory variable (with the intercept forced to zero). The results of this data exercise indicate that for each dollar sent by the Sender (multiplied into $3 for the Receiver), the Receiver sends back $1.46 on average (i.e., the Receivers send roughly half of the multiplied money back).
### Appendix Table 1B: Descriptive Statistics (Survey Traits)

<table>
<thead>
<tr>
<th>Worker Traits</th>
<th>mean/sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personality: Extraversion</td>
<td>2.89 (0.86)</td>
</tr>
<tr>
<td>Personality: Agreeableness</td>
<td>3.29 (0.85)</td>
</tr>
<tr>
<td>Personality: Conscientiousness</td>
<td>3.57 (0.88)</td>
</tr>
<tr>
<td>Personality: Neuroticism</td>
<td>3.16 (0.96)</td>
</tr>
<tr>
<td>Personality: Openness</td>
<td>3.58 (0.91)</td>
</tr>
<tr>
<td>Survey: High Stakes Stress</td>
<td>5.55 (1.28)</td>
</tr>
<tr>
<td>Survey: Well Under Pressure</td>
<td>4.26 (1.39)</td>
</tr>
<tr>
<td>Survey: Competition Best In Me</td>
<td>4.64 (1.51)</td>
</tr>
<tr>
<td>Survey: Competition Is Good</td>
<td>4.74 (1.53)</td>
</tr>
<tr>
<td>Demographic: Female</td>
<td>0.59 (0.49)</td>
</tr>
<tr>
<td>Demographic: White</td>
<td>0.29 (0.45)</td>
</tr>
<tr>
<td>Demographic: Asian</td>
<td>0.50 (0.50)</td>
</tr>
<tr>
<td>Demographic: Hispanic</td>
<td>0.23 (0.42)</td>
</tr>
</tbody>
</table>

**Observations**: 101

**Notes**: The table shows the mean value of each trait for subjects in the experiment. Standard deviations are shown in parentheses. Survey questions are either on a scale from 1 to 5 or from 1 to 7, as indicated below each trait.
Subject demographics were as follows. The pool of participants consisted of 37.6% male, 59.4% female, and 3.0% other. The race and ethnicity of the participants was 28.7% White, 50.5% Asian, 1% Black, and 19.8% other. The demographic variable on Hispanic identification was collected separately (in line with Census guidelines); 22.8% of subjects identified as Hispanic. The mean age of the subject pool was 20.7 years old, with a minimum age of 18 and a maximum age of 29.

B.3 Selection into Dismissal Threats on Worker Traits

This subsection will analyze whether selection into high-paying, strong-dismissal-threat jobs occurs based on worker traits other than productivity. To help answer this question, Appendix Tables 2A and 2B present the mean of each worker trait separately for subjects who prefer the Safe Job and subjects who prefer the Risky Job. Standard deviations are displayed below each mean in parentheses. Additionally, the third column presents the p-value resulting from a rank-sum test on the null hypothesis that the two samples of worker traits came from the same distribution. Appendix Table 2A shows only incentivized measurements of worker traits, while Appendix Table 2B shows survey questions and demographics.
Appendix Table 2A: Job Preference and Worker Traits (Incentivized)

<table>
<thead>
<tr>
<th></th>
<th>(1) Prefer Safe Job</th>
<th>(2) Prefer Risky Job</th>
<th>(3) Ranksum Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean/sd</td>
<td>mean/sd</td>
<td>p-value</td>
</tr>
<tr>
<td>High Piece Rate Tasks</td>
<td>8.38</td>
<td>12.08</td>
<td>0.0002***</td>
</tr>
<tr>
<td></td>
<td>(3.87)</td>
<td>(4.99)</td>
<td></td>
</tr>
<tr>
<td>Responsiveness to Piece Rate</td>
<td>1.06</td>
<td>0.67</td>
<td>0.8041</td>
</tr>
<tr>
<td></td>
<td>(1.95)</td>
<td>(1.01)</td>
<td></td>
</tr>
<tr>
<td>High Stakes Tasks</td>
<td>2.13</td>
<td>2.64</td>
<td>0.0597+</td>
</tr>
<tr>
<td></td>
<td>(1.34)</td>
<td>(1.23)</td>
<td></td>
</tr>
<tr>
<td>Confidence: Rank Choice</td>
<td>5.67</td>
<td>4.44</td>
<td>0.0028**</td>
</tr>
<tr>
<td></td>
<td>(2.12)</td>
<td>(2.01)</td>
<td></td>
</tr>
<tr>
<td>Confidence: Actual - Reported Rank</td>
<td>1.45</td>
<td>0.43</td>
<td>0.0431*</td>
</tr>
<tr>
<td></td>
<td>(2.07)</td>
<td>(2.69)</td>
<td></td>
</tr>
<tr>
<td>Risk Choice</td>
<td>2.94</td>
<td>3.10</td>
<td>0.5256</td>
</tr>
<tr>
<td></td>
<td>(1.18)</td>
<td>(1.40)</td>
<td></td>
</tr>
<tr>
<td>Loss Choice</td>
<td>2.42</td>
<td>2.61</td>
<td>0.6887</td>
</tr>
<tr>
<td></td>
<td>(1.15)</td>
<td>(1.41)</td>
<td></td>
</tr>
<tr>
<td>Patience Choice</td>
<td>3.42</td>
<td>3.38</td>
<td>0.9912</td>
</tr>
<tr>
<td></td>
<td>(0.71)</td>
<td>(0.63)</td>
<td></td>
</tr>
<tr>
<td>SocialPrefs: Altruism</td>
<td>2.40</td>
<td>2.40</td>
<td>0.8606</td>
</tr>
<tr>
<td></td>
<td>(1.87)</td>
<td>(1.70)</td>
<td></td>
</tr>
<tr>
<td>SocialPrefs: Trust</td>
<td>1.85</td>
<td>2.08</td>
<td>0.1973</td>
</tr>
<tr>
<td></td>
<td>(0.98)</td>
<td>(1.05)</td>
<td></td>
</tr>
<tr>
<td>SocialPrefs: Reciprocity</td>
<td>1.53</td>
<td>1.42</td>
<td>0.5417</td>
</tr>
<tr>
<td></td>
<td>(0.59)</td>
<td>(0.79)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>40</td>
<td>61</td>
<td>101</td>
</tr>
</tbody>
</table>

Notes: The table shows the mean value of each trait for the subset of workers who preferred the Safe Job in Column 1 and the Risky Job in Column 2. Standard deviations are shown in parentheses. Column 3 shows the p-value from a Wilcoxon rank-sum test comparing the distribution between prefer Safe and prefer Risky workers of the trait in that row.

\[ + p < 0.1, \; * p < 0.05, \; ** p < 0.01, \; *** p < 0.001 \]
In Appendix Table 2A, there is strong evidence of sorting on productivity into different jobs, but little evidence of sorting on any other incentivized worker traits. There is a large difference in productivity during the high piece rate period between subjects who preferred the Safe Job versus those who preferred the Risky Job, but this merely echoes the evidence already presented above of selection on worker productivity. Similarly, subjects who prefer the Risky Job tended to do slightly better on during the short high-stakes work period, but this difference was only marginally significant (p = 0.0597). Although subjects who prefer the Risky Job were significantly more likely to report a lower rank in the confidence elicitation (i.e., on average they believed they were more productive, since rank 1 is the highest productivity rank), this is likely due to the reported rank’s strong correlation with overall productivity in the work task. Turning to overconfidence and underconfidence, although both groups of subjects are overconfident on average, the subjects who prefer the Safe Job were more overconfident than subjects who preferred the Risky Job (p = 0.0431). In other words, subjects who prefer the high-paying, strong-dismissal-threat job had more accurate relative evaluations of their productivity. No other incentivized worker traits appear to be significantly different by job preference.

Appendix Table 2B presents evidence of sorting on the various survey measures or demographics. Only one personality trait displays a statistically significant difference between subjects who preferred the Safe Job and those who preferred the Risky Job: conscientiousness. This difference implies that subjects who prefer the Risky Job tend to be more organized, careful, and diligent. There is also suggestive evidence of higher levels of extraversion (sociability and talkativeness), although this is only marginally significant (p < 0.1).

It is important to consider that many of these worker traits may be correlated with productivity. This is not to say that the univariate comparisons do not matter; firms offering higher pay along with stronger dismissal threats are still likely to observe the differences in traits uncovered in the univariate analysis above. However, it’s useful to determine whether workers are attracted to stronger dismissal threats independently based on these traits, or whether it’s simply the case that more productive workers happen to also have these traits. An analysis of correlation matrices between these traits (not shown) confirms that most of the traits that show univariate differences by worker preference are indeed correlated with output under the high piece rate.

Probit regressions are run to further establish whether selection on worker traits occurs independent of productivity. Appendix Table 3 presents the results. The latent dependent variable is the propensity to select the Risky Job. Appendix Table 3 only includes incentivized worker trait measurements as independent variables. Standard errors are clustered by session. Rather than coefficients, marginal effects evaluated at the mean of each trait...
**Appendix Table 2B: Job Preference and Worker Traits (Survey)**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Prefer Safe Job</th>
<th>Prefer Risky Job</th>
<th>Ranksum Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean/sd</td>
<td>mean/sd</td>
<td>p-value</td>
</tr>
<tr>
<td>Personality: Extraversion</td>
<td>2.70 / 0.80</td>
<td>3.01 / 0.88</td>
<td>0.0965*</td>
</tr>
<tr>
<td>Personality: Agreeableness</td>
<td>3.33 / 0.80</td>
<td>3.27 / 0.88</td>
<td>0.8241</td>
</tr>
<tr>
<td>Personality: Conscientiousness</td>
<td>3.34 / 0.78</td>
<td>3.72 / 0.92</td>
<td>0.0214*</td>
</tr>
<tr>
<td>Personality: Neuroticism</td>
<td>3.35 / 0.83</td>
<td>3.03 / 1.02</td>
<td>0.1387</td>
</tr>
<tr>
<td>Personality: Openness</td>
<td>3.45 / 0.83</td>
<td>3.66 / 0.96</td>
<td>0.1698</td>
</tr>
<tr>
<td>Survey: High Stakes Stress</td>
<td>5.72 / 1.20</td>
<td>5.44 / 1.32</td>
<td>0.3057</td>
</tr>
<tr>
<td>Survey: Well Under Pressure</td>
<td>4.17 / 1.43</td>
<td>4.31 / 1.37</td>
<td>0.6205</td>
</tr>
<tr>
<td>Survey: Competition Best In Me</td>
<td>4.60 / 1.48</td>
<td>4.67 / 1.54</td>
<td>0.8191</td>
</tr>
<tr>
<td>Survey: Competition Is Good</td>
<td>4.67 / 1.23</td>
<td>4.79 / 1.71</td>
<td>0.4310</td>
</tr>
<tr>
<td>Demographic: Female</td>
<td>0.65 / 0.48</td>
<td>0.56 / 0.50</td>
<td>0.3563</td>
</tr>
<tr>
<td>Demographic: White</td>
<td>0.23 / 0.42</td>
<td>0.33 / 0.47</td>
<td>0.2661</td>
</tr>
<tr>
<td>Demographic: Asian</td>
<td>0.55 / 0.50</td>
<td>0.48 / 0.50</td>
<td>0.4656</td>
</tr>
<tr>
<td>Demographic: Hispanic</td>
<td>0.28 / 0.45</td>
<td>0.20 / 0.40</td>
<td>0.3613</td>
</tr>
</tbody>
</table>

**Notes:** The table shows the mean value of each trait for the subset of workers who preferred the Safe Job in Column 1 and the Risky Job in Column 2. Standard deviations are shown in parentheses. Column 3 shows the p-value from a Wilcoxon rank-sum test comparing the distribution between prefer Safe and prefer Risky workers of the trait in that row.

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
are displayed. (Another specification was run which additionally included control variables for all survey-based trait measurements, but none of these are significant and their inclusion does not impact the results in any meaningful way.)

**Appendix Table 3: Probit Regression of Job Preference on Worker Traits**

<table>
<thead>
<tr>
<th></th>
<th>(1) Prefer Risky Job</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Piece Rate Tasks</td>
<td>0.0444*** (0.001)</td>
</tr>
<tr>
<td>Responsiveness to Piece Rate</td>
<td>-0.0143 (0.425)</td>
</tr>
<tr>
<td>High Stakes Tasks</td>
<td>-0.0387 (0.448)</td>
</tr>
<tr>
<td>Confidence: Rank Choice</td>
<td>-0.0256 (0.280)</td>
</tr>
<tr>
<td>Risk Choice</td>
<td>0.0355 (0.448)</td>
</tr>
<tr>
<td>Loss Choice</td>
<td>-0.00993 (0.787)</td>
</tr>
<tr>
<td>Patience Choice</td>
<td>0.0510 (0.543)</td>
</tr>
<tr>
<td>Social Prefs: Altruism</td>
<td>0.00730 (0.840)</td>
</tr>
<tr>
<td>Social Prefs: Trust</td>
<td>0.0792 (0.219)</td>
</tr>
<tr>
<td>Social Prefs: Reciprocity</td>
<td>-0.0979 (0.192)</td>
</tr>
<tr>
<td>Observations</td>
<td>98</td>
</tr>
</tbody>
</table>

*Notes: The table shows the results of a probit regression. The latent dependent variable is the subjects’ propensity to prefer the Risky Job. Each row displays the marginal effect of the variable evaluated at the mean of that variable. Standard errors are clustered at the session level; p-values are shown in parentheses.  
* p < 0.05, ** p < 0.01, *** p < 0.001*  

It turns out that controlling for worker productivity eliminates any association between
job preference and other worker traits. Worker productivity during the piece rate period is highly predictive of preferences over pay and dismissal threats, but none of the other incentivized trait measurements are significant once also controlling for productivity.

Again, this does not mean multidimensional sorting is not occurring – it simply means that the observed sorting on traits is due to their correlation with productivity and not an independent relationship between the trait and job preference. Firms offering steeper dismissal incentives will still observe differences on average in the worker traits mentioned above. This is even more likely to be true when workers have less ability to predict their productivity in a given job, thus muting selection on productivity and empowering selection on other worker traits and preferences.

The experiment design placed all worker trait measurements after the work periods, and therefore a natural question to ask is whether the history observed by subjects may impact these trait measurements and thus change the conclusions on multidimensional sorting into stronger dismissal threats. Two examples of this stand out. First, whether a subject was fired along with how many other subjects were announced to have been fired are likely to impact stated confidence. Second, the amount of money earned so far may impact risk and loss aversion. The following analysis will reexamine multidimensional sorting after adjusting for session-specific history. It turns out that even after these adjustments, no qualitative differences are found in sorting patterns from previous analysis.

To determine whether session history impacts statements of confidence, a specification search was conducted to isolate the impact of exogenous events. OLS regressions were run with elicited confidence as the outcome variable and a variety of explanatory variables (standard errors always clustered at the session level). The results consistently indicate that whether a subject was dismissed has an important impact on the subject’s statement of confidence, even after controlling for the probability of dismissal. In other words, being exogenously dismissed hurts a subject’s confidence regardless of whether the dismissal was warranted. On the other hand, neither the number nor percentage of other subjects who were dismissed in a session seemed to have any impact on statements of confidence in any reasonable specifications.

The final preferred specification with strong explanatory power was used to create an “adjusted confidence” measurement for each subject. The specification controlled for output during the high piece rate period, job preference, job assignment, the probability of being dismissed in each period, and an indicator for whether the subject was actually dismissed in that period. The coefficients on being dismissed in the first and second periods were statistically significant and had values of 1.31 and 0.86, respectively. The interpretation is that controlling for the chance of being dismissed, actually being dismissed results in subjects
reporting a worse rank by 1.31 units if dismissed in the first period, and a worse rank by 0.86 units if dismissed in the second period. The adjusted confidence measure was created by simply subtracting these respective values from the reported ranks of subjects who were fired in periods 1 or 2 (in a sense, adjusting them to be more confident if they were randomly fired).

The results of the analysis of multidimensional sorting does not change in any meaningful way when using this new adjusted confidence measure. There is still a highly statistically significant difference in adjusted confidence between subjects who prefer the Risky Job versus the Safe Job, and this difference is not significant in a probit regression that controls for productivity along with all other incentivized worker traits.

Similar methodology was used to determine whether the amount of money earned during the work periods impacted subject risk, loss, and patience preferences. For a series of OLS regressions with each of these three preference as an outcome, the cumulative amount of money earned had no impact on these preferences at the 5% confidence level for any reasonable specification. Therefore, no further adjustments or analysis was done using these traits.