

Incentivizing Quantity and Quality of Output: An Experimental Investigation of the Quantity-Quality Trade-off

Jared Rubin ^a
Anya Samek ^b
Roman M. Sheremeta ^{c,*}

^a Argyros School of Business and Economics, Chapman University
One University Drive, Orange, CA 92866, U.S.A.

^b Center for Economic and Social Research, University of Southern California
635 Downey Way, Los Angeles, CA 90098, U.S.A.

^c Weatherhead School of Management, Case Western Reserve University
11119 Bellflower Road, Cleveland, OH 44106, U.S.A.

January 27, 2016

Abstract

Firms face an optimization problem that requires a maximal quantity output given a quality constraint. How firms should incentivize quantity and quality to meet these dual goals remains an open question, potentially due to limitations of field data. We provide a theoretical model and conduct an experiment in which participants are paid for both quantity and quality of a real effort task. Consistent with the theoretical predictions, higher quality incentives encourage participants to shift their attention from quantity to quality, and higher quality incentives reduce inefficient decision-making. We also observe behavioral components in responsiveness to the quality incentive.

JEL Classifications: D24, J24, J31, J41

Keywords: quantity, quality, experiment, incentives, real effort, loss aversion

* Corresponding author: Roman Sheremeta, rshereme@gmail.com

We want to thank Kevin Guo, Christa Gibbs, Kathryn Carroll and students at the Behavioral and Experimental Economics Research Group for excellent research assistance. Any remaining errors are ours.

1. Introduction

Firms face a quantity-quality output tradeoff. For instance, an orchard owner wants to incentivize her workers to pick as many apples as possible, but if workers are paid only based on the number of apples they pick, they may be careless and bruise much of the crop. Yet if the owner rewards workers solely based on the number of non-bruised apples picked, much of the orchard may remain unpicked. Understanding how workers respond to such quantity-quality tradeoffs is essential for understanding the conditions under which different wage schemes are efficient.

How to incentivize workers is a question fundamental to economics, and an active literature exists on the effect of different incentive compensation schemes on worker effort.¹ Worker productivity and quantity of output has been a focus of theoretical and empirical economic research for decades (Laffont and Martimort, 2009; Syverson, 2011). Some important works also consider the quality side of the tradeoff. Holmstrom and Milgrom (1991) and Baker (1992) lay out seminal principal-agent models that incorporate the multi-dimensional aspects of work incentives, and explain why incentivizing quantity may cause agents to ignore the quality of their output. However, to the best of our knowledge, an empirical investigation of how workers respond to different quantity-quality incentives is missing from the literature.

Recent work has begun to investigate the optimal incentive contracts for workers in situations when the firm cares about multiple dimensions of worker output. A series of papers have used existing data or field experiments to investigate the relative merits of flat rate versus piece rate incentive schemes in the workplace (Lazear, 2000; Paarsch and Shearer, 2000;

¹ For instance, economists have used behavioral economics theories of gift exchange and framing to induce greater productivity of workers in a field setting (see Gneezy and List, 2006 for gift exchange, and Hossain and List, 2012 on framing). Other notable papers include the merits of competitive or piece rate incentive schemes, including the gender gap in competitiveness (Gneezy et al., 2003), and various profit-sharing compensation schemes (Nalbantian and Schotter, 1997). While many of these papers have incorporated quality considerations into their work, none of them have evaluated quality of output directly.

Shearer, 2004; Copeland and Monnet, 2009; Ederer and Manso, 2013; Al-Ubaydli et al., 2015).²

The above papers find a positive impact of piece rates on quantity of output, but the evidence is mixed for its impact on quality. For instance, Al-Ubaydli et al. (2015) find increases in quality, while Johnson et al. (2015) and Ederer and Manso (2013) find decreases in quality from pay-for performance compensation. Confounds in the field may explain these disparate results – for example, Al-Ubaydli et al. (2015) note that workers may have uncertainty about the observability of the quality of output.

We overcome the confounds inherent in field data by designing a laboratory experiment aimed at investigating the trade-offs of incentivizing quantity and quality in worker effort. The laboratory environment gives us control over problems that exist in the field, such as difficulty in monitoring quality of output and the effects of selection from workers moving in and out of the firm in response to changes in their compensation package. In addition, while the existing literature examines how different contracts (e.g., pay-for-performance or fixed wage) affect quantity or quality, none of these address the quantity-quality tradeoffs directly.

Importantly, our results broaden our understanding of incentivizing the multiple dimensions of productivity. Consider the following example. The firm hires a worker to produce a certain product (as in our introductory example). If the worker's payment depends only on *quantity* of units produced, she may choose to put all of her effort in maximizing quantity, while ignoring the quality of the production, as predicted by the theoretical models of Holmstrom and Milgrom (1991) and Baker (1992). On the other hand, if the worker's payment depends mainly on *quality* of units produced, she may choose to put too much effort in maintaining high quality,

² Additional related work includes Eriksson et al. (2009) who use a real-effort experiment to examine how feedback about performance of others impacts quantity and quality under pay-for-performance and tournament payment schemes, Charness and Grieco (2014) who investigate the impact of incentives on creativity, and Bracha and Fershtman (2013) who study how competitive incentive schemes affect the combination of cognitive and labor efforts provided by workers.

while producing low quantity (an example of such a setup would be workers paid mostly on commission from successful endeavors). The optimal compensation scheme should involve a balance between rewarding quantity and quality, although the optimal tradeoff depends on how important quality considerations are in the market for the final product.

In this study, we examine how individuals respond to different incentives in the quantity-quality context. How do quality incentives impact productivity? Does incentivizing quality increase the quality of output? Does the quantity-quality tradeoff differ by the characteristics of the agent (e.g., ability, risk aversion, loss aversion, and overconfidence)? What type of labor contracts should employers offer workers given the conditions faced by the employers? The theoretical model we outline provides insights into the answers to these questions, while the experiment we conduct provides empirical evidence. In our experiment, individuals solve math problems and their output quantity (number of problems attempted) and quality (number of problems answered correctly) is measured when (i) only quantity is incentivized, (ii) some quality is incentivized, and (iii) the bulk of the incentives are on quality.

We find that, consistent with the theoretical predictions, higher quality incentives encourage participants to shift their attention from quantity to quality by increasing quality of output and decreasing the number of problems answered incorrectly at the expense of lowering quantity of output. We also find that higher quality incentives encourage participants to make better quantity-quality trade-offs, reducing inefficient decision-making. Moreover, we find that, consistent with the theoretical predictions, higher ability participants choose to focus more on quality and have lower error rates (i.e., number incorrect/number attempted). Based on these findings, we discuss what the potential implications are for the optimal type of labor contracts that employers should offer workers.

We also observe a behavioral component in responsiveness to the quality incentive. There is heterogeneity in the impact of treatment, with more loss-averse participants displaying greater changes to their output from a change in quality incentives. Overall, we find that loss aversion leads participants to focus more on quantity and less on quality, while increasing the number of problems answered incorrectly. In addition, we characterize participants by whether they focus on pursuing quality or quantity during the experiment, and find that higher quality incentives increase the number of subjects whose primary focus is quality.

In what follows, Section 2 describes the theoretical model and predictions. Section 3 outlines the experimental design. Section 4 summarizes the results, and Section 5 provides a discussion and conclusion.

2. Theoretical Model and Predictions

2.1. Theoretical Model

In this model, we provide insight into how economic agents exert effort under different rewards for the quantity and quality of their output. Consider an agent who exerts two-dimensional effort $e = (e_1, e_2)$, where $e_1 \geq 0$ is effort used to produce quantity and $e_2 \geq 0$ is effort used to produce quality. The agent has one unit of effort to provide, so $e_1 + e_2 = 1$. The agent has ability $a > 0$, and agents with higher ability produce high quality output at lower cost (for a given level of effort).

The quantity of high-quality output produced, $q^H = e_1 p(e_2)$, depends on effort e_1 used to produce quantity and effort e_2 used to increase the probability of successful production $p(e_2)$, where $p' > 0$, $p'' < 0$, $p(0) = 0$, and $p(1) = 1$. The low-quality output is produced with a remaining probability, i.e., $q^L = e_1(1 - p(e_2))$. The cost of exerting effort to produce quality is

$c(e_2, a)$, where $c_1 > 0$, $c_2 < 0$, $c_{11} > 0$, $c_{12} < 0$, and $c(0, a) = 0$. We use a simplifying assumption that the cost to produce quantity is not a function of ability and it is normalized to zero.³ The agent receives wage $w_1 \geq 0$ for each output (payment for quantity) and wage $w_2 \geq 0$ for each high-quality output (payment for quality). We assume that quality is perfectly verifiable.

The payoff of the agent is:

$$U = w_1(q^H + q^L) + w_2q^H - c(e_2, a) = w_1e_1 + w_2e_1p(e_2) - c(e_2, a) \quad (1)$$

Since $e_1 = 1 - e_2$ in equilibrium, the agent's first order condition $\frac{\partial U}{\partial e_2} = 0$ is:

$$-w_1 - w_2p(e_2) + w_2(1 - e_2)p'(e_2) - c_1(e_2, a) = 0 \quad (2)$$

2.2. Predictions

From the first order condition in (2), we can derive comparative statics predictions related to how optimal effort levels respond to changes in relative wages. Consider first how effort changes as the relative return from producing quality increases (i.e., w_2 increases relative to w_1). From (2), there are increasing differences in $\{e_2, w_2\}$ if and only if $e_2 \leq 1 - \frac{p(e_2)}{p'(e_2)}$. Note that this also means that there are increasing differences in $\{e_2, w_2\}$ if and only if $\frac{\partial q^H}{\partial e_2} \geq 0$, since $\frac{\partial q^H}{\partial e_2} = -p(e_2) + (1 - e_2)p'(e_2)$ and thus $\frac{\partial q^H}{\partial e_2} \geq 0$ implies $e_2 \leq 1 - \frac{p(e_2)}{p'(e_2)}$. Intuitively, it must be true that $\frac{\partial q^H}{\partial e_2} \geq 0$ at any level of e_2 chosen by the agent: otherwise, increasing e_2 would decrease both high-quality output q^H and low-quality output q^L .⁴ Hence, there are increasing differences in $\{e_2, w_2\}$, and e_2 is increasing in w_2 . Finally, there is some e^* , which solves $e^* = 1 - \frac{p(e^*)}{p'(e^*)}$,

³ Including the cost of e_1 in the agent's utility function would change none of the comparative static results.

⁴ To see this, note that $q^L = (1 - e_2)(1 - p(e_2))$, which is clearly decreasing in e_2 .

which the optimal value of e_2 never exceeds. This intuition is summarized in the following proposition and represented graphically in Figure 1.

Proposition 1: As w_2 increases, the optimal value e_1^* weakly decreases and e_2^* weakly increases.

It follows directly from Proposition 1 that $\frac{\partial(q^H + q^L)}{\partial w_2} \leq 0$ since $q^H + q^L = e_1$ and $\frac{\partial e_1}{\partial w_2} \leq 0$,

implying that higher quality incentives decrease the total output (the sum of high-quality and low-quality output). We state this as Prediction 1:

Prediction 1: The total quantity of output $q^H + q^L$ is weakly decreasing in w_2 .

It also follows from Proposition 1 that $\frac{\partial q^H}{\partial w_2} \geq 0$. Recall that $q^H = e_1 p(e_2) = (1 - e_2)p(e_2)$. Therefore, $\frac{\partial q^H}{\partial w_2} = \frac{\partial e_2}{\partial w_2}[-p(e_2) + (1 - e_2)p'(e_2)] \geq 0$ since $\frac{\partial e_2}{\partial w_2} \geq 0$ and the term in brackets is always non-negative in equilibrium. The intuition behind this result is that an increase in w_2 encourages the agent to spend more effort in a manner where more high-quality units are produced. Sometimes this means reducing the effort e_1 spent on producing quantity, as noted above. This brings us to the next prediction:

Prediction 2: The level of high-quality output q^H is weakly increasing in w_2 .

Next, we define the error rate as the fraction of low-quality output relative to total output, or $\frac{q^L}{q^H+q^L}$. From Proposition 1, it follows that $\frac{\partial(q^L/(q^H+q^L))}{\partial w_2} \leq 0$. To show this, note that $\frac{q^L}{q^H+q^L} = \frac{e_1(1-p(e_2))}{e_1} = 1 - p(e_2)$. Therefore, $\frac{\partial(q^L/(q^H+q^L))}{\partial w_2} = \frac{\partial(1-p(e_2))}{\partial w_2} = -\frac{\partial e_2}{\partial w_2} p'(e_2) \leq 0$ since $\frac{\partial e_2}{\partial w_2} \geq 0$ and $p'_1 > 0$. This brings us to the next prediction:

Prediction 3: The error rate $\frac{q^L}{q^H+q^L}$ is weakly decreasing in w_2 .

Finally, consider how the agent's ability affects her decision to focus on quality effort e_2 . Since ability only enters into the cost function, it follows directly from (2) that there are increasing differences in $\{e_2, a\}$, and hence e_2 is increasing in a . The intuition underlying this result is straight-forward: higher ability agents face a lower marginal cost from exerting effort used to produce quality, so they choose greater e_2 . This intuition is summarized in the following proposition:

Proposition 2: As a increases, the optimal value e_1^* weakly decreases and e_2^* weakly increases.

Predictions 4, 5, and 6 follow the same mathematical logic as Predictions 1, 2, and 3, respectively. For the sake of brevity we do not repeat the mathematics nor the intuition, but simply note that the sign of $\frac{\partial e_1^*}{\partial w_2}$ equals the sign of $\frac{\partial e_1^*}{\partial a}$ and the sign of $\frac{\partial e_2^*}{\partial w_2}$ equals the sign of $\frac{\partial e_2^*}{\partial a}$, entailing that the above comparative statics are the same with respect to a as they are with respect to w_2 .

Prediction 4: The total quantity of output $q^H + q^L$ is weakly decreasing in α .

Prediction 5: The level of high-quality output q^H is weakly increasing in α .

Prediction 6: The error rate $\frac{q^L}{q^H+q^L}$ is weakly decreasing in α .

3. Experimental Design and Procedures

The experiment used participants drawn from the population of undergraduate students at the University of Wisconsin-Madison. Computerized experimental sessions were run using z-Tree (Fischbacher, 2007) at the Behavioral Research Insights Through Experiments (BRITE) Laboratory. A total of 197 participants participated in 14 experimental sessions. Upon arriving at the laboratory, participants were randomly assigned to a computer station. The experiment proceeded in seven parts. All participants were given written instructions, available in Appendix A, at the beginning of each part, and an experimenter also read the instructions aloud.

In the first part, participants performed a real effort task: adding up sets of five randomly generated 2-digit numbers by hand, as quickly as possible, with no assistance other than a pen and paper (no calculators), for 5 minutes. The 2-digit numbers task is commonly used in the experimental literature because it is easy to explain, does not require previous experience and performance is not associated with a particular gender, socioeconomic background, or physical conditioning (Niederle and Vesterlund, 2007; Cason et al., 2010). In each treatment, participants were provided with up to 60 problems (one at a time) they could attempt to solve during 5 minutes. Participants could see only one problem at a time and they could not skip any problems. Each time a participant arrived at a new problem, she had 5 seconds to review it before the

submit button appeared. After spending at least 5 seconds, the computer allowed participants to enter their answers. The 5 second delay can be considered an opportunity cost of skipping a problem by submitting any random answer.

In all treatments, as shown in Table 1, participants received $w_1 = \$0.10$ for each attempted problem (i.e., for quantity). Depending on the treatment, participants also received an additional bonus for each attempted problem answered correctly (i.e., for quality), varying from $w_2 = \$0.00$ in the T-0.00 treatment to $w_2 = \$3.00$ in the T-3.00 treatment for each attempted problem answered correctly.

In the second part, we elicited beliefs about output quality by asking participants to provide a guess about how many of the attempted problems in the first part they solved correctly. Participants received an additional \$3 if their guess was equal to the number of correct answers they provided in the first part. Participants were not aware of the second part until after they finished the first part of the experiment. The main purpose of eliciting participants' beliefs about their performance was to test whether the measured quantity and quality of output from part 1 matched the participants' own beliefs about how much quality they attempted.

In order to learn whether behavioral motivations play a role in responsiveness to quality incentives, in parts 3-5, we elicited participants' preferences toward ambiguity, risk and loss. While our theoretical model did not make a clear prediction about the role of behavioral motivations, such motivations have been shown to be important in principal-agent relationships in the field (Haigh and List, 2005; Hossain and List, 2012).

In part 3, we elicited participants' preferences toward ambiguity from a set of 20 lotteries (see Table B1 in Appendix B). In each lottery, participants were asked to state whether they prefer an ambiguous option A (\$0.00 or \$10.00 with unknown chance each) or a safe option B

(increasing monotonically from \$0.50 to \$10.00).⁵ Parameters were set in such a way that more ambiguity-averse participants would choose safer options (and switch earlier to a safe option) than less ambiguity-averse participants. Again, participants were not aware of this part until after they finished the preceding part.

In part 4, we elicited participants' preferences toward risk from a set of 20 lotteries (see Table B2 in Appendix B). In each lottery, participants were asked to state whether they prefer a risky option A (\$0.00 or \$10.00 with 50% chance each) or a safe option B (increasing monotonically from \$0.50 to \$10.00). As in previous parts, participants were not aware of this part until after they finished the preceding part.

In part 5, we elicited participants' preferences toward losses from a set of 20 lotteries (see Table B3 in Appendix B). In each lottery, participants were asked to state whether they prefer a risky option A (50% chance of losing a certain amount between -\$0.50 to -\$10.00) or a safe option B of \$0. As in previous parts, participants were not aware of this part until after they finished the preceding part.

Part 6 was used to obtain a measure of a participants' ability on the math task, independent of incentive concerns. In this part, participants again performed a real effort task (as in the first part of the experiment): adding up sets of five randomly generated 2-digit numbers by hand, as quickly as possible. This time, participants had only 2.5 minutes to complete the task. The computer provided participants with up to 30 math problems (one at a time) that they could attempt to solve during the allotted time. As before, participants could see only one problem at a time and they could not skip any problems. Each time a participant arrived at a new problem, she had 5 seconds to review it before the submit button appeared. Participants received \$0.50 for

⁵ Our elicitation procedure is similar to Shupp et al. (2013).

each problem answered correctly, regardless of the treatment. Contrary to the first part, subjects made no earnings from attempted problems that were incorrect.

Finally, in part 7, participants were asked to provide a guess about how many of the attempted problems in the sixth part they solved correctly. Participants received an additional \$3 if their guess was equal to the number of correct answers they provided in the sixth part. Participants were not aware of this task until after they finished the preceding part of the experiment. The main purpose of eliciting participants' beliefs about their performance in the sixth part was to obtain a measure of confidence, which may be linked to participants' decision to put more effort into quality or quantity. This measure is comparable across treatments, since it is not affected by the quantity-quality incentives that differ across treatments (unlike the guess in part 2, which may be a function of the different quantity-quality tradeoffs faced in part 1).

At the end of the experiment, each participant received earnings from parts 1, 2, 6 and 7. For parts 3-5, in order to avoid portfolio effects, only one part and line was paid out at random. Each session lasted approximately 90 minutes. Participants' earnings ranged from \$10.50 to \$119.70, with a median of \$25.60. In addition to their earnings in the experiment, participants also received a \$7.00 show-up fee.

4. Results

4.1. How Incentives Impact Quantity and Quality

The summary statistics of our experiment are reported in Table 2 and represented graphically in Figures 2-4. First, we examine how higher quality incentives (i.e., higher reward for solving problems correctly) impact quantity (i.e., the number of problems attempted). Prediction 1 states

that the level of total output (i.e., $q^H + q^L$) should decrease with higher quality incentives (i.e., w_2).

We begin by noting that there is no significant difference in the number of problems attempted between treatments T-0.00 and T-0.05 where quality incentives are low (25.63 versus 27.58; Wilcoxon rank-sum test, p-value = 0.43). Similarly, there are no differences between treatments T-0.25 and T-0.50 where quality incentives are medium (18.15 versus 18.51; Wilcoxon rank-sum test, p-value = 0.96) and treatments T-1.00 and T-3.00 where quality incentives are high (13.85 versus 13.20; Wilcoxon rank-sum test, p-value = 0.35). In the analysis that follows, we report pooled data from the low quality incentive treatments T-0.00 and T-0.05, the medium quality incentive treatments T-0.25 and T-0.50, and the high quality incentive treatments T-1.00 and T-3.00.

Figure 2 suggests that there are clear differences in the number of problems attempted between treatments with low quality incentives (i.e., T-0.00 and T-0.05), medium quality incentives (i.e., T-0.25 and T-0.50) and high quality incentives (i.e., T-1.00 and T-3.00). Pairwise comparisons show that the differences in distributions are statistically significant (Wilcoxon rank-sum test, six p-values < 0.01 and two p-values < 0.08). We also find significant differences when comparing all treatments jointly (Kruskal-Wallis test, p-value < 0.01). Therefore, consistent with Prediction 1, we find that higher incentives for quality decrease quantity of output.

Result 1: Higher quality incentives decrease quantity of output.

Second, we examine how higher quality incentives impact quality (i.e., the number of problems solved correctly). Recall that Prediction 2 states that the level of high-quality output (i.e., q^H) should increase with higher quality incentives (i.e., w_2).

We find no significant difference in quality between T-0.00 and T-0.05 (Wilcoxon rank-sum test, p-value = 0.65). Also, there are no differences in quality between T-0.25, T-0.50, T-1.00, and T-3.00 (Wilcoxon rank-sum test, four p-values > 0.35). However, we see significant difference in quality when comparing treatments with low quality incentives to treatments with medium and high quality incentives (Wilcoxon rank-sum test, eight p-values < 0.04). The differences are also significant when comparing all treatments jointly (Kruskal-Wallis test, p-value < 0.01). Therefore, consistent with Prediction 2, we find that higher incentives for quality increase quality of output.

Result 2: Higher quality incentives increase quality of output.

Third, we examine how higher quality incentives impact the error rate. To calculate the error rate, we use the ratio of the number of problems solved incorrectly to the number of problems attempted. Recall that Prediction 3 states that the error rate (i.e., $\frac{q^L}{q^H+q^L}$) should decrease with higher quality incentives (i.e., w_2).

There are no significant differences in the error rate between T-0.00 and T-0.05 (Wilcoxon rank-sum test, p-value = 0.48), between T-0.25 and T-0.50 (Wilcoxon rank-sum test, p-value = 0.49), and between T-1.00 and T-3.00 (Wilcoxon rank-sum test, p-value = 0.35). However, we find that the error rate in treatments with high quality incentives is significantly lower than in treatments with medium quality incentives (Wilcoxon rank-sum test, three p-values

< 0.01 and one p-value < 0.07). Similarly, the error rate in treatments with medium quality incentives is significantly lower than in treatments with low quality incentives (Wilcoxon rank-sum test, three p-values < 0.01 and one p-value < 0.03). The differences are also significant when comparing all treatments jointly (Kruskal-Wallis test, p-value < 0.01), suggesting that, consistent with Prediction 3, the error rate decreases with higher quality incentives.

Result 3: Higher quality incentives decrease the error rate.

Together, Results 1, 2, and 3 provide strong support for the theoretical predictions of our model: higher quality incentives encourage participants to shift their attention from quantity to quality by increasing quality of output and decreasing the error rate at the expense of lowering quantity of output.

4.2. Individual Characteristics

Next, we explore whether individual characteristics impact the choice of quality versus quantity. To answer this question, we elicited different individual characteristics summarized in Table 3. Specifically, in part 6 of the experiment, we elicited an independent measure of participants' ability by having participants perform a real effort task during 150 seconds. In part 7, we elicited participants' beliefs about their performance in part 6 (see Appendix A for details). Using these beliefs, we compute an individual measure of overconfidence, defined as the predicted number of problems solved correctly in part 6 minus the number of problems actually solved correctly. From Table 3, we see that the median participant is overconfident, overestimating his

performance by 1 correct problem (the mean participant overestimates performance by 0.84 correct problems).

Knowing that preferences towards uncertainty play an important role in decision making, we also elicited preferences regarding ambiguity, risk, and losses using multiple lottery choice mechanisms (see Table B1, Table B2, and Table B3 in Appendix B). Parameters of the elicitation procedure were set in such a way that the more ambiguity-, risk- and loss-averse participants would choose ‘safer’ options relative to ‘uncertain’ options (and switch earlier from an ‘uncertain’ option to a ‘safe’ option) than the less ambiguity-, risk- and loss-averse participant. Although the three elicitation tasks are not directly comparable, in all three tasks, a higher number (of ‘safe’ options) implies a higher level of aversion toward ambiguity, risk and losses.⁶

Next, we examine whether the elicited characteristics of participants are predictive of the number of problems attempted, the number of problems solved, and the error rate. First, we examine what factors influence the number of problems attempted. Table 4 shows the estimation results of different OLS regressions in which the dependent variable is quantity of output (the number of problems attempted), and the independent variables are the dummy for treatments with medium quality incentives (i.e., T-0.25 and T-0.50), the dummy for treatments with high quality incentives (i.e., T-1.00 and T-3.00), and ability. Specification (2) supports the non-parametric results by showing that medium and high quality incentives decrease quantity of output relative to low quality incentives, and that high quality incentives decrease quantity of output relative to medium quality incentives (as indicated by the p-value at the bottom of Table 4). Specifications (1) and (2) indicate that there is no statistically significant relationship between

⁶ A simple correlation analysis shown in Table B4 in Appendix B indicates that there is a strong correlation between ambiguity-aversion and risk-aversion ($\rho = 0.67$), and somewhat weaker correlation between loss-aversion and ambiguity-aversion ($\rho = 0.30$) and loss-aversion and risk-aversion ($\rho = 0.35$).

the participant's ability and quantity of output. This is contrary to Prediction 4, which suggested that total quantity should be decreasing in ability. A potential explanation is that the relationship between ability and quantity is non-linear: subjects of sufficiently high ability may have been able to exert little effort on quality while still answering problems correctly, thus achieving higher quality *and* quantity. This insight does not follow directly from the model,⁷ but it does make intuitive sense. To address this possibility, we include an ability-squared term. As expected, the coefficient on this term is positive and statistically significant, indicating that at sufficiently high ability levels subjects can focus on quantity (without, presumably, losing much quality; see Table 6). The coefficient on ability is negative, though imprecisely estimated.

We next examine the impact of elicited individual characteristics on quantity of output, reported in Table 5. We find that overconfidence, risk, and ambiguity are not predictive of quantity (specifications 1, 2, and 3). However, we find that loss aversion is a significant predictor of quantity (specification 4), with participants who are more loss-averse choosing to focus on quantity by attempting more problems. This finding is intuitive. By focusing on quantity, participants can always guarantee a certain amount of payment for their performance, i.e., $w_1(q^H + q^L) = w_1e_1$, while focusing on quality involves the possibility of not solving the problem correctly, i.e., $w_2q^H = w_2e_1p(e_2, a)$. Therefore, a loss-averse participant may choose to focus mainly on quantity in order to minimize potential losses. To check the robustness of our findings, we include additional interaction terms (specification 5).⁸ Besides confirming our previous findings, we also find that higher quality incentives affect loss-averse participants less. The most straight-forward interpretation of this result is that loss aversion is less salient when

⁷ Although a simple addition to the model, where the probability of producing high-quality output is in part a function of ability, could yield this result.

⁸ Including interaction terms with overconfidence, risk, and ambiguity does not yield any statistically significant results, and we therefore do not report these results for the sake of brevity. These results are available upon request.

incentives for quality are high, because the loss incurred from spending more time on a problem is smaller relative to the potential gains of getting the problem correct. We report additional robustness checks with individual treatment dummies in Table B5 in Appendix B.

Result 4: Participants focus more on quantity (attempt more problems) if they are more loss-averse.

Next, we examine what factors influence the choice of quality (the number of problems solved). The estimation results reported in Tables 6 and 7 provide support for the non-parametric results that medium and high quality incentives increase quality of output. Also, we find that consistent with Prediction 5, ability is positively and significantly correlated with quality, suggesting that participants of higher ability are more likely to focus on quality of output. This is not simply a matter of participants who are better at math in part 6 being better at math in part 1: whether participants focus on quantity is a *decision* in part 1. Furthermore, we find that loss aversion is a significant predictor of quality (specifications 4 and 5 in Table 7), with participants who are less loss-averse choosing to focus on quality by solving more problems. Again, this finding is intuitive, since focusing on quality entails losing out on the sure wage (i.e., $w_1 = \$0.10$) associated with focusing on quantity. Hence, more loss averse participants are less willing to take such a loss. As was the case in the quantity regressions, loss aversion only shows up as salient in the low quality incentive treatments (see specification 5 in Table 7). Intuitively, in the low quality incentive treatments, the benefit of focusing on quality is low relative to the loss of the sure wage associated with focusing on quantity. As the quality incentive increase, the

latter loss becomes relatively less salient. We report additional robustness checks with individual treatment dummies in Table B6 in Appendix B.

Result 5: Participants focus more on quality (solve more problems correctly) if they have higher ability or if they are less loss-averse.

Finally, we examine what factors influence the error rate. The estimation results reported in Tables 8 and 9 provide support for the non-parametric results that medium and high quality incentives decrease the error rate. Moreover, consistent with Prediction 6, in all specifications we find that ability is negatively and significantly correlated with the error rate, suggesting that participants of higher ability have lower error rates. There is a positive and significant relationship between loss aversion and the error rate in specifications (4) and (5) of Table 9, confirming our previous findings relating loss aversion to quantity and quality. We report additional robustness checks with individual treatment dummies in Table B7 in Appendix B.

Result 6: Participants have higher error rates if they have lower ability or if they are more loss-averse.

To summarize, Results 4, 5, and 6 indicate that there are important individual characteristics impacting the quantity-quality tradeoff. First, we find that there is heterogeneity in the impact of treatment, with more loss-averse individuals displaying greater changes to their output from a change in quality incentives. Loss aversion leads participants to focus more on quantity and less on quality, while increasing the error rates. Also, we find that, consistent with

the theoretical predictions, higher ability participants choose to focus more on quality and have lower error rates.

4.3. Classification of Participants

Next, we characterize participants by response time to identify how treatment differences affected the incentives of participants to focus primarily on quantity or quality. We begin by examining how much time participants spend on average on a given problem, which we consider an indicator of how much effort participants exert on quality. We assume that participants who spent more time on a problem than the average were more likely to be focusing on quality. As shown in Table 10, there are no significant differences in the average time spent on a problem between the low quality incentive treatments T-0.00 and T-0.05 (15.83 versus 13.73; Wilcoxon rank-sum test, p-value = 0.57). Similarly, there are no differences between medium quality incentives treatments T-0.25 and T-0.50 (18.19 versus 18.37; Wilcoxon rank-sum test, p-value = 0.98) and high quality incentives treatments T-1.00 and T-3.00 (22.83 versus 23.28; Wilcoxon rank-sum test, p-value = 0.36). However, there are clear differences in the average time spent per problem when comparing low quality incentive treatments with medium and high quality incentive treatments, and there is likewise a difference comparing medium incentive treatments with high incentive treatments. Pairwise comparisons show that the differences in distributions are statistically significant (Wilcoxon rank-sum test, ten p-values < 0.01 and two p-values < 0.07). The difference are also significant when comparing all treatment jointly (Kruskal-Wallis test, p-value < 0.01).

Table 10 also reports the fraction of problems answered ‘quickly’ (signifying that a participant is focusing on quantity) by treatment. Recall that each participant had to spend a

minimum of 5 seconds on each problem since the ‘submit’ button did not appear on the screen until 5 seconds had passed. We therefore look at different cut-off points – 6, 7, and 10 seconds – to see whether participants answer more quickly when quality is not incentivized. We find that 25% of problems were answered within 6 seconds when the reward for solving problems is not incentivized (i.e., T-0.00), while only 2% answer within 6 seconds when the reward is highly incentivized (i.e., T-1.00 and T-3.00) (Kruskal-Wallis test across all six treatments, p-value < 0.01). A similar pattern is observed for participants answering within 7 seconds (Kruskal-Wallis test, p-value < 0.01) and within 10 seconds (Kruskal-Wallis test, p-value < 0.01).

Finally, the last column in Table 10 shows the fraction of participants choosing to focus only on quality. We define a participant as focusing on quality on a specific question if they either answered the question correctly or they spent at least as much time answering the problem as the minimum time it took them to answer a question in part 6 (where quantity was not incentivized and payouts were the same across treatments).⁹ As expected, we find that higher quality incentives increase the number of quality types (Kruskal-Wallis test, p-value < 0.01).

Result 7: Higher quality incentives increase the number of participants focusing on quality and decrease the number of participants focusing on quantity.

⁹ We do not count decisions made in the last 30 seconds or decisions made in the participant’s last answer because the decision-making calculus at the end of the five minute period may be different than in the first four minutes. For instance, one who can correctly answer a problem in 10 seconds (meaning that she should focus on quality in most of the treatments) has incentive to input a quick answer if there are only 6 seconds remaining. Moreover, we derived numerous definitions for quality and all give similar qualitative results. The statistics associated with other metrics are available upon request.

4.4. Optimal Choice of Quantity and Quality

A participant making a decision of whether to focus on quantity or quality should take into account her ability to perform the task. As we have already shown, such ability is indeed important in making this decision. However, another important factor is the price the participant is rewarded for quality. For example, when the reward is $w_2 = \$0.25$, the participant should expect to earn $\$0.35$ ($w_1 = \$0.10$ for quantity and $w_2 = \$0.25$ for quality) for successfully completing a task, which comes at the cost of spending time on that task (say x seconds depending on the ability). However, the participant also has an option to focus solely on quantity, which would result in a reward of $w_1 = \$0.10$ at the cost of a minimum 5 seconds spent on the task. Therefore, each participant should make a choice of whether to focus on quantity or quality depending on their relative ability to complete the task in x seconds and prices w_1 and w_2 . Specifically, if $(w_1 + w_2)/x > w_1/5$ then a participant should focus on quality, and otherwise they should focus on quantity. One immediate implication is that higher w_2 should lead participants to pay more attention to quality. See a more formal discussion of this argument in Section 2, Proposition 1.

We can therefore calculate how many participants *should* have chosen to focus on quality given their ability. As a proxy for ability, we use the average time a participant needs to solve one problem correctly in part 6. Table 11 summarizes the average ability of participants across treatments: the first column reports the average number of seconds subjects spent on each problem in part 6 in each treatment. Not surprisingly, since participants were randomly assigned to each treatment, there is no difference in ability between treatments (Kruskal-Wallis test, p-value = 0.49). However, since the reward for quality is different across treatments, the expected earnings are different. For example, when the reward is \$0.25 per correct answer, a participant

who spends 30 seconds to solve one problem correctly should expect to earn \$3.50 if she chooses to focus on quality, i.e., $(\$0.25 + \$0.10) \times 300/30 = \$3.50$. However, if instead, such a participant chooses to focus solely on quantity, she can earn \$6 since the opportunity cost is 5 seconds of moving to the next problem, i.e., $\$0.10 \times 300/5 = \6.00 . Therefore, a rational decision maker who can solve only one problem during 30 seconds should choose to focus on quantity when the reward for quality is \$0.25. Similar computations can be performed for all participants in each treatment. Table 11 reports the fraction of participants who should choose quality over quantity based on their ability and quality incentives.

Obviously, when the reward for quality is \$0.00, nobody should choose to focus on quality. The same is true when the reward is only \$0.05 for all but the most mathematically gifted (none of whom took part in this treatment). When the reward is \$0.25, 12% of participants should choose to focus on quality. When the reward is \$0.50 this number increases to 79%, and further to 97% when the reward is \$1.00. Finally, when the reward is \$3.00, all participants should focus on quality. Using this information, we can calculate the portion of participants in each treatment that chose to correctly focus on quantity or quality. We first calculate their average earnings from focusing on quality, as measured by the average time they spent deriving a correct answer in part 6 (see Table 11). Using this measure, we calculate their expected earnings from focusing on quality, which equals $(300 / \text{average seconds per correct answer}) \times (\$0.10 + w_2)$, where w_2 differs by treatment. Any participant whose expected earnings from focusing on quality exceed \$6 (the amount one could earn from solely focusing on quantity) should focus on quality; otherwise they should focus on quantity. We consider it a mistake for a participant to focus on quality (even once) when she should focus on quantity or for a participant

to focus on quantity (even once) when she should focus on quality.¹⁰ Table 11 shows that 85% of participants make mistakes when quality is not at all incentivized (i.e., T-0.00), while only 7% make mistakes when the reward is most highly incentivized (i.e., T-3.00) (Kruskal-Wallis test, p-value < 0.01). The same conclusion stands when we drop the first 34 seconds of experiment, which is one standard deviation above the mean time taken to answer a question in part 6 (Kruskal-Wallis test, p-value < 0.01).

Result 8: Higher quality incentives encourage participants to make better trade-offs between quantity and quality, reducing inefficient decision making.

4.5. Optimal Wage Schemes for Employers

Finally, our experiment has implications for the optimal type of labor contracts that employers offer workers. Of course, the optimal contract will depend on the conditions faced by the employer. Employers whose product receives a high price for quality will want to incentivize quality more than employers whose product's price is less dependent on quality.

Although we do not wish to push the specific results of our experiment too far, these results have implications for the type of contracts that employers offer their workers. Specifically, assume that employers face the following profit equation:

$$\pi = p_L q_L + p_H q_H - w_1 q_L - w_2 q_H = (p_L - w_1)q_L + (p_H - w_2)q_H, \quad (4)$$

where $q_H \geq 0$ and $q_L \geq 0$ are the quantities of high quality and low quality output, $p_H \geq 0$ is the price received for high quality output, and $p_L \in \mathbb{R}$ the price received for low quality output. Note that p_L may be negative (if a company has to refund low quality or is sued) or positive. $w_1 \geq 0$ is

¹⁰ Similar to before, we do not count decisions made in the last 30 seconds or decisions made in the participant's last answer in this calculation.

the wage paid for low quality output and $w_2 \geq 0$ is the wage paid for high quality output. In our experiment, $w_1 = \$0.10$ and w_2 varies by treatment.

In this profit equation, p_L and p_H are exogenous parameters (assuming the market is competitive), and w_1 and w_2 are choice variables. The question that our experiment answers – for a *specific* set of parameters – is which wage scheme among six possible ones is optimal for a given set $\{p_L, p_H\}$. Figure 7 reports the results. Employing $w_1 = \$0.10$ and the summary statistics of each treatment to determine q_L and q_H , these figures indicate that providing low quantity incentives (i.e., $w_2 = \$0.00$ and $\$0.05$) is an optimal strategy when the price for quantity is positive and the price for quality is not sufficiently large. When there is a penalty for low quality output, it is optimal to incentivize for quality. As long as the penalty is not too large, it is optimal to provide a medium quality incentive (i.e., $w_2 = \$0.25$) – participants produce enough high-quality output that it is not worthwhile to pay them more to receive marginally more high-quality output. Also note that $w_2 = \$0.50$ is not optimal in the price range shown in Figure 7, indicating diminishing returns to w_2 once quality is sufficiently incentivized.¹¹ When the costs associated with low-quality output are high (i.e., p_L is sufficiently negative), it is optimal for employers to impose greater quality incentives (i.e., $w_2 = \$1.00$). But the marginal benefits attained by much higher wages (i.e., $w_2 = \$3.00$) are only worth it if p_L is highly negative or p_H is very large (indeed, this figure would need a much larger range to show where $w_2 = \$3.00$ is optimal). In other words, although greater quality incentives are optimal to impose when the return on quality is large, the return on higher wages is rapidly diminishing past a certain point. What this point depends on the characteristics of the workers (e.g., their ability and loss aversion) and is likely only determinable through employer trial and error.

¹¹ It is true, however, that if the range of p_L and p_H is sufficiently large, $w_2 = \$0.50$ and $w_2 = \$3.00$ are optimal wages in part of the parameter space.

The specifics of this exercise should obviously not be taken too far; the actual numbers apply only to the present experiment. But one upshot of the analysis is that when firms are punished for low quality – via lawsuit, return policies, or warehouse fees for unsold goods – even modest quality incentives are enough to encourage high quality effort. When punishment is minimal in our experiment, “medium” quality incentives are optimal, with the less expensive incentive, $w_2 = \$0.25$, dominating the more expensive incentive, $w_2 = \$0.50$. Likewise, when punishment for low quality is severe, high quality incentives are optimal, again with the less expensive incentive, $w_2 = \$1.00$, dominating the more expensive incentive, $w_2 = \$3.00$.

5. Discussion and Conclusion

Firms face an optimization problem that requires a maximal quantity output given a quality constraint. It is not trivial to incentivize economic agents to care not only about quantity but also about quality of output. A large literature suggests that incentives designed to encourage certain behaviors may backfire (Bowles, 2009; Gneezy et al., 2011; Bowles and Polania-Reyes, 2012). For example, incentives that are ‘too small’ may crowd out intrinsic motivation to put forth effort (Gneezy and Rustichini, 2000).¹² The problem becomes even more complicated when quality is incorporated into consideration. On the one hand, incentivizing quality may increase individual performance. On the other hand, incentivizing quality may decrease quantity of output. Moreover, incentivizing quality may be too costly.

We provide a theoretical model and conduct an experiment to examine how incentivizing quality impacts individual decisions to focus on quality versus quantity. Consistent with

¹² Along these lines, Mellstrom and Johannesson (2008) provide evidence that monetary incentives may decrease (instead of increasing) blood donations. Rietz et al. (2013) show that imposing restrictive rules may have a detrimental impact on a gift-exchange relationship.

theoretical predictions, we find that higher quality incentives encourage participants to shift their attention from quantity to quality by increasing quality of output and decreasing the error rate – at the expense of lowering quantity of output. We also find that higher quality incentives encourage participants to make better quantity-quality trade-offs, reducing inefficient decision making.

Our findings have direct practical relevance for managers and employers. First, we show both theoretically and experimentally that there are important quantity-quality trade-offs that should be taken into account when designing contracts. For example, a manager who is highly concerned with the quality of output may choose to incentivize high-quality output. This should lead to higher quality of output and a lower error rate. However, it will most likely decrease quantity of output. Moreover, the results of our experiment show that although greater quality incentives are optimal to impose when the return on quality is large, the return on higher wages diminishes rapidly past a certain point. Therefore, the optimal compensation scheme should involve a balance between rewarding quantity and quality. For instance, Mauboussin (2012) provides an example of the Wallace Company, a pipe and valve distributor which won the prestigious Malcolm Baldrige National Quality Award in 1990 but had to file for bankruptcy two years later. Mauboussin concludes that “both too little and too much quality can be bad for a company’s financial performance.” Our study provides both theoretical and empirical evidence for this insight.

Second, our findings contribute to the literature examining how behavioral components can be used to improve work outcomes (Haigh and List, 2005; Hossain and List, 2012). Hossain and List (2012), for example, show that a manager can significantly improve the performance of workers by framing contracts in terms of “losses”. We also find in our experiment that more

loss-averse participants display greater changes to their output from a change in quality incentives. Participants who are more loss-averse choose to focus more on quantity, increasing the error rates. Therefore, a manager who is highly concerned with the quality of output may choose to frame contracts in terms of losses to reduce the tendency of loss averse workers to focus on quantity rather than quality.

Third, one practical application of our findings relates to an ongoing discussion in health economics on how to reward physicians in order to improve medical practice and increase social welfare. One part of the debate is whether to reward physicians solely for the volume of services they order (quantity) or to incorporate certain quality measures (quality).¹³ Our findings suggest that rewarding quality is indeed effective in increasing quality of output and decreasing the error rate. However, in our experiment, it is easy to define and measure quality, which is not always the case in the medical field where quality is ill-defined and may be difficult to measure (Hennig-Schmidt et al., 2011; Godager et al., 2016).¹⁴

¹³ See the following article in the New York Times: http://www.nytimes.com/2013/01/12/nyregion/new-york-city-hospitals-to-tie-doctors-performance-pay-to-quality-measures.html?pagewanted=2&_r=1&hp

¹⁴ Incorporating quality incentives into a payment scheme can be difficult because measures of quality are not well defined and it can be difficult to monitor the quality. Other concerns with using pay-for-performance payment schemes are that some doctors will choose to pass on hard patients, while others will choose to perform too many treatments to assure good quality.

References

- Al-Ubaydli, O., Andersen, S., Gneezy, U., & List, J.A. (2015). Carrots that look like sticks: Toward an understanding of multitasking incentive schemes. *Southern Economic Journal*, 81, 538-561.
- Baker, G.P. (1992). Incentive Contracts and Performance Measurement. *Journal of Political Economy*, 100, 598-614.
- Bowles, S. (2009). When economic incentives backfire. *Harvard Business Review*, 87, 22-23.
- Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, 50, 368-425.
- Bracha, A., & Fershtman, C. (2013). Competitive incentives: Working harder or working smarter? *Management Science*, 59, 771-781.
- Cason, T.N., Masters, W.A., & Sheremeta, R.M. (2010). Entry into winner-take-all and proportional-prize contests: An experimental study. *Journal of Public Economics*, 94, 604-611.
- Charness, G., & Grieco, D. (2014). Creativity and financial incentives. Working Paper.
- Copeland, A., & Monnet, C. (2009). The welfare effects of incentive schemes. *Review of Economic Studies*, 76, 93-113.
- Dickinson, D., & Villeval, M.C. (2008). Does monitoring decrease work effort? The complementarity between agency and crowding-out theories. *Games and Economic Behavior*, 63, 56-76.
- Ederer, F.P., & Manso, G. (2013). Is pay-for-performance detrimental to innovation? *Management Science*, 59, 1496-1513.
- Eriksson, T., Poulsen, A., & Villeval, M.C. (2009). Feedback and incentives: Experimental evidence. *Labour Economics*, 16, 679-688.
- Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96, 1611-1630.
- Fehr, E. and Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, 422, 137-140.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10, 171-178.
- Frey, B.S. (1993). Does monitoring increase work effort? The rivalry with trust and loyalty. *Economic Inquiry*, 31, 663-670.
- Gneezy, U., & List, J.A. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74, 1365-1384.
- Gneezy, U., & Rustichini, A. (2000a). A fine is a price. *The Journal of Legal Studies*, 29, 1-17.
- Gneezy, U., & Rustichini, A. (2000b). Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115, 791-810.
- Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, 25, 191-209.
- Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*, 118, 1049-1074.
- Godager, G., Henning-Schmidt, H., & Iversen, T. (2016). Does performance disclosure influence physicians' medical decisions? An experimental analysis. *Journal of Economic Behavior and Organization*, forthcoming.

- Griffith, T.L. (1993). Monitoring and performance: A comparison of computer and supervisor monitoring. *Journal of Applied Social Psychology*, 23, 549-572.
- Haigh, M.S., & List, J.A. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *Journal of Finance*, 60, 523-534.
- Hennig-Schmidt, H., Selten, R., & Wiesen, D. (2011). How payment systems affect physicians' provision behavior – An experimental investigation. *Journal of Health Economics*, 30, 637-646.
- Holmstrom, B., & Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization*, 7, 24-52.
- Hossain, T., & List, J.A. (2012). The behavioralist visits the factory: Increasing productivity using simple framing manipulations. *Management Science*, 58, 2151-2167.
- Johnson, R.M., Reiley, D.H., & Muñoz, J.C. (2015). "The war for the fare": How driver compensation affects bus system performance. *Economic Inquiry*, 53, 1401-1419.
- Laffont, J.J., & Martimort, D. (2009). *The theory of incentives: The principal-agent model*. Princeton University Press.
- Lazear, E.P. (2000). Performance pay and productivity. *American Economic Review*, 90, 1346-1361.
- Mellstrom, C., & Johannesson, M. (2008). Crowding out in blood donation: Was Titmuss right? *Journal of the European Economic Association*, 6, 845–863.
- Mauboussin, M.J. (2012). *The success equation: Untangling skill and luck in business, sports, and investing*. Harvard Business Press.
- Nalbantian, H.R., & Schotter, A. (1997). Productivity under group incentives: An experimental study. *American Economic Review*, 87, 314-341.
- Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122, 1067-1101.
- Paarsch, H.J., & Shearer, B. (2000). Piece rates, fixed wages, and incentive effects: Statistical evidence from payroll records. *International Economic Review*, 41, 59-92.
- Rietz, T., Schniter, E., Sheremeta, R.M., & Shields, T.W. (2013). Trust, reciprocity and rules. Working Paper.
- Shearer, B. (2004). Piece rates, fixed wages and incentives: Evidence from a field experiment. *Review of Economic Studies*, 71, 513-534.
- Shupp, R., Sheremeta, R.M., Schmidt, D., & Walker, J. (2013). Resource allocation contests: Experimental evidence. *Journal of Economic Psychology*, 39, 257-267.
- Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 49, 326-365.
- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 106, 1039-1061.

Figure 1: Optimal values of e_1^* and e_2^* over different values of w_2

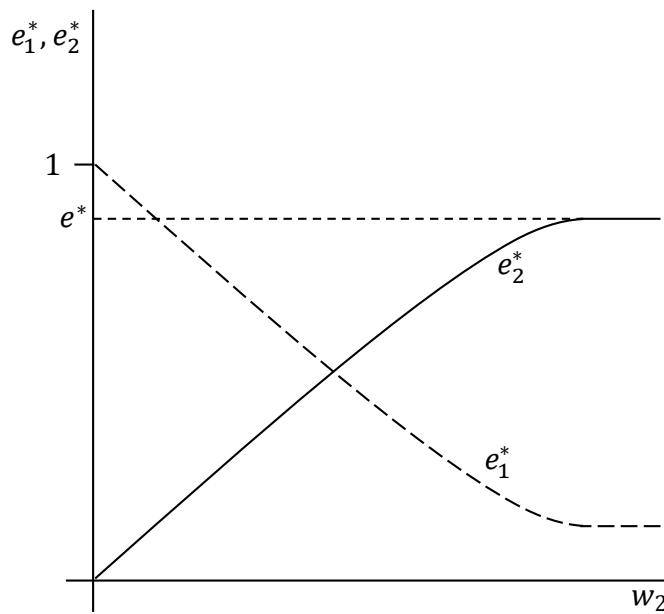


Figure 2: Measure of quantity (average problems attempted) by treatment

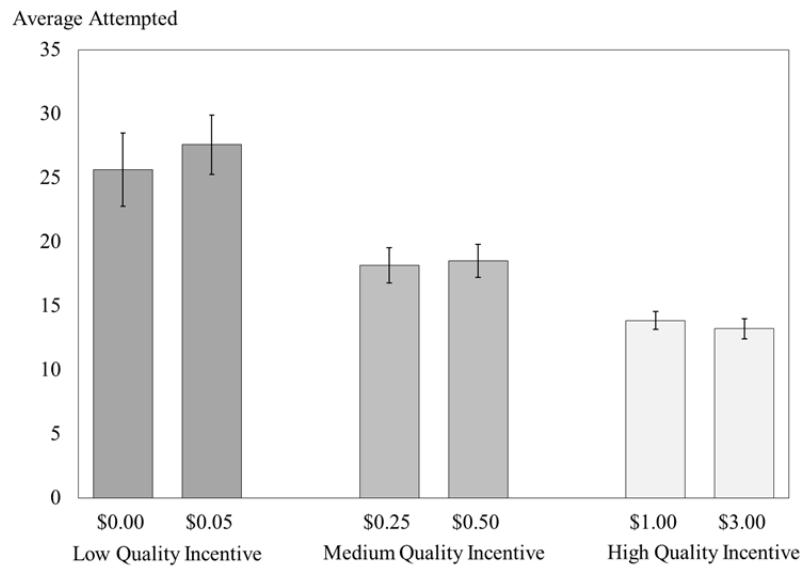


Figure 3: Measure of quality (average problems correct) by treatment

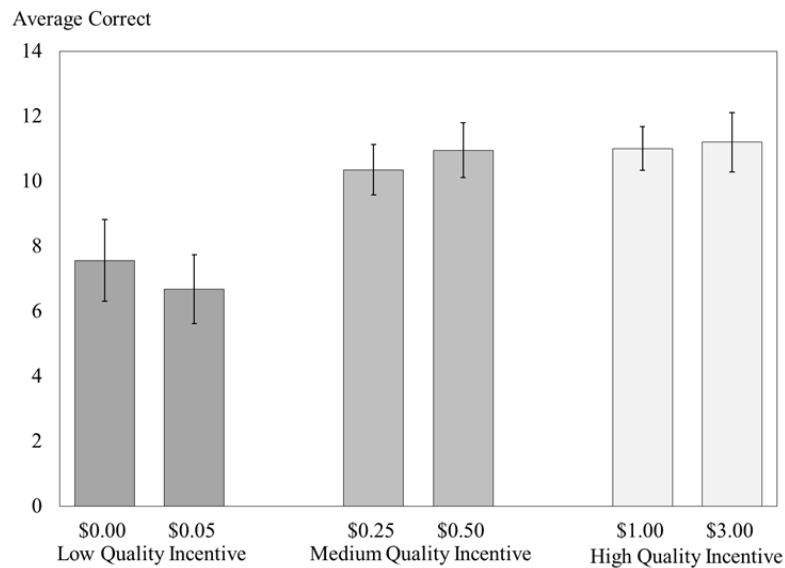


Figure 4: Error rate by treatment

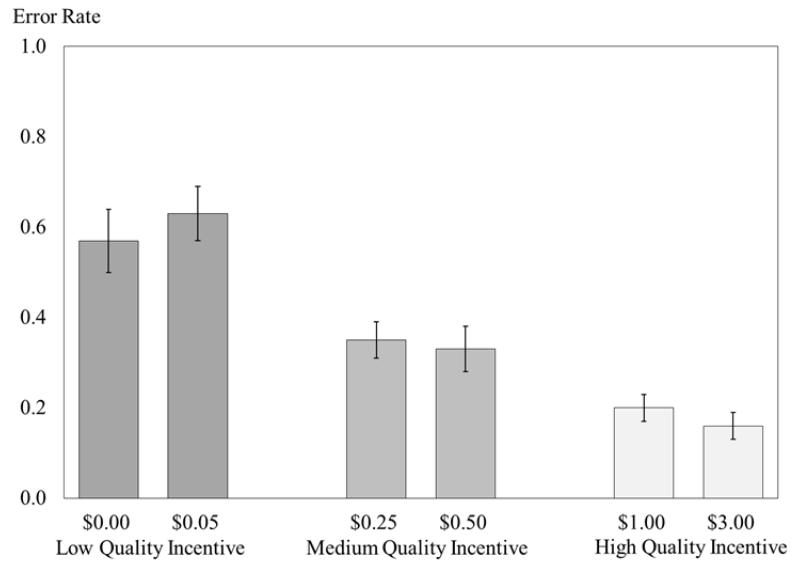


Figure 5: Fraction of problems answered in less than 10 seconds

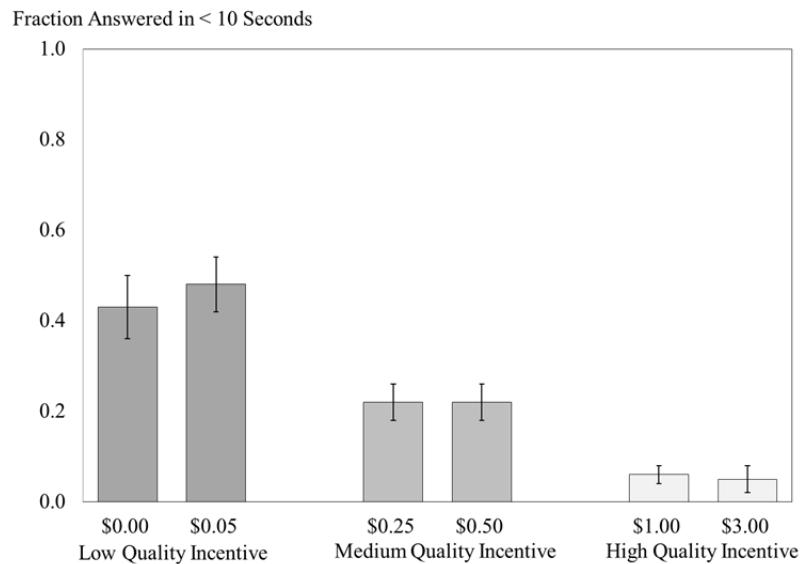


Figure 6: Fraction of subjects focusing only on quality

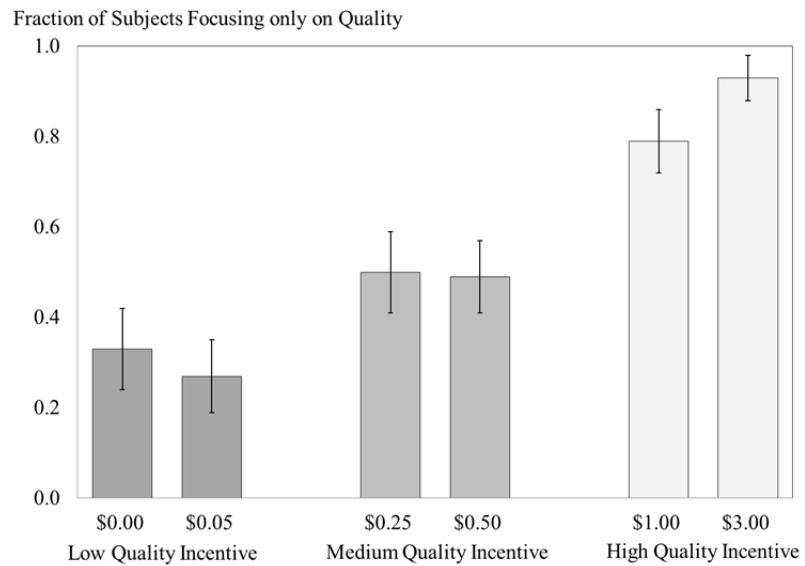


Figure 7: Optimal wage scheme for $p_H \leq 5, p_L \in [-5, 5]$

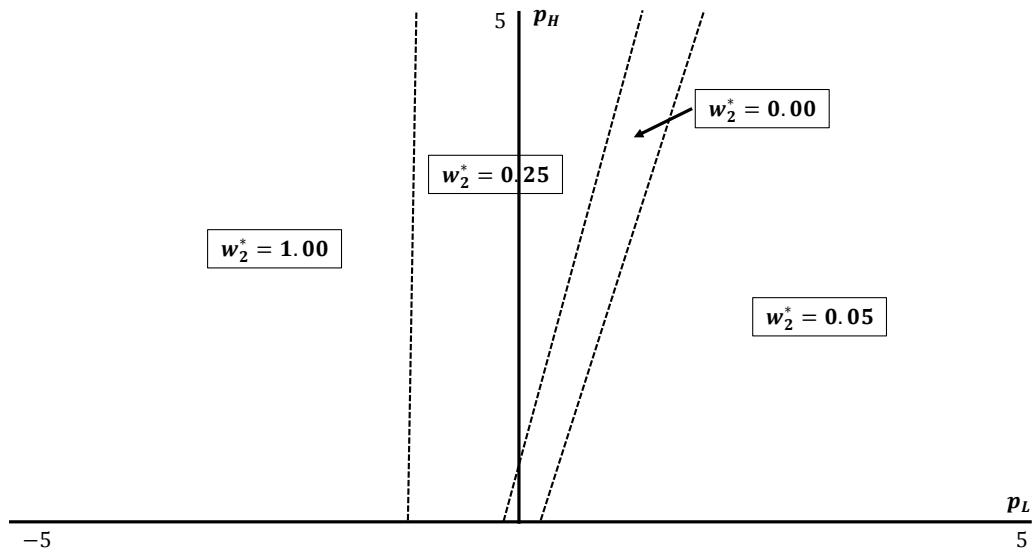


Table 1: Summary of treatments

Treatment	Payment for each problem		
	Attempted	Correct	N
T-0.00	\$0.10	\$0.00	27
T-0.05	\$0.10	\$0.05	33
T-0.25	\$0.10	\$0.25	34
T-0.50	\$0.10	\$0.50	39
T-1.00	\$0.10	\$1.00	34
T-3.00	\$0.10	\$3.00	30

Table 2: Summary statistics

Reward	Average attempted	Average correct	Average incorrect	Error rate = incorrect/attempted	N
\$0.00	25.63 (2.86)	7.56 (1.26)	18.07 (3.60)	0.57 (0.07)	27
\$0.05	27.58 (2.32)	6.67 (1.06)	20.91 (3.14)	0.63 (0.06)	33
\$0.25	18.15 (1.37)	10.35 (0.78)	7.79 (1.76)	0.35 (0.04)	34
\$0.50	18.51 (1.30)	10.95 (0.84)	7.56 (1.63)	0.33 (0.05)	39
\$1.00	13.85 (0.70)	11.00 (0.67)	2.85 (0.66)	0.20 (0.03)	34
\$3.00	13.20 (0.78)	11.20 (0.91)	2.00 (0.43)	0.16 (0.03)	30

Standard errors are in parentheses.

Table 3: Elicited characteristics

Percentile	Ability (correct in part 6)	Overconfidence (guess – correct)	Ambiguity	Risk	Loss
Min	0	-3.00	0	1	4
25%	5	0.00	10	10	14
50%	6	1.00	11	11	15
75%	7	2.00	13	12	17
Max	17	4.00	20	20	20

Table 4: OLS regressions of quantity

Specification	(1)	(2)	(3)
Dependent variable = Quantity (number attempted)			
Ability	0.26	0.41	-0.95
(correct in part 6)	(0.30)	(0.26)	(0.84)
Ability ²		0.09*	(0.05)
Medium quality incentives (T-0.25 and T-0.50)		-8.76*** (1.65)	-8.56*** (1.65)
High quality incentives (T-1.00 and T-3.00)		-13.28*** (1.69)	-13.16*** (1.68)
Constant	17.74*** (1.96)	24.38*** (1.91)	28.58*** (3.11)
Observations	197	197	197
p-value, Medium = High	--	0.01	0.00
R-squared	0.00	0.25	0.26

Table 5: OLS regressions of quantity, with behavioral characteristics

Specification	(1)	(2)	(3)	(4)	(5)
Dependent variable = Quantity (number attempted)					
Ability	-1.02	-1.02	-1.01	-1.11	-0.70
(correct in part 6)	(0.87)	(0.84)	(0.84)	(0.83)	(0.81)
Ability ²	0.09* (0.05)	0.09* (0.05)	0.09* (0.05)	0.10* (0.05)	0.08 (0.05)
Medium quality incentives (T-0.25 and T-0.50)	-8.57*** (1.65)	-8.45*** (1.65)	-8.59*** (1.65)	-9.29*** (1.65)	17.42* (8.88)
High quality incentives (T-1.00 and T-3.00)	-13.19*** (1.69)	-12.92*** (1.69)	-13.13*** (1.68)	-13.33*** (1.66)	13.62* (7.65)
Overconfidence	-0.19 (0.58)				
Risk aversion		0.29 (0.24)			
Ambiguity aversion			0.26 (0.22)		
Loss aversion				0.56** (0.23)	1.77*** (0.38)
Loss aversion × Medium quality incentives					-1.81*** (0.58)
Loss aversion × High quality incentives					-1.87*** (0.52)
Constant	29.07*** (3.46)	25.60*** (3.99)	25.78*** (3.90)	21.11*** (4.32)	2.30 (6.35)
Observations	197	197	197	197	197
p-value, Medium = High	0.00	0.01	0.01	0.01	--
R-squared	0.26	0.27	0.27	0.28	0.34

Table 6: OLS regressions of quality

Specification	(1)	(2)	(3)
Dependent variable = Quality (number correct)			
Ability	1.51*** (0.11)	1.48*** (0.10)	0.81** (0.32)
(correct in part 6)			0.04** (0.02)
Ability ²			
Medium quality incentives (T-0.25 and T-0.50)		2.17*** (0.64)	2.26*** (0.63)
High quality incentives (T-1.00 and T-3.00)		3.58*** (0.65)	3.63*** (0.64)
Constant	0.51 (0.71)	-1.24* (0.73)	0.82 (1.19)
Observations	197	197	197
p-value, Medium = High	--	0.02	0.03
R-squared	0.51	0.58	0.59

Table 7: OLS regressions of quality, with behavioral characteristics

Specification	(1)	(2)	(3)	(4)	(5)
Dependent variable = Quality (number correct)					
Ability	0.96*** (0.33)	0.82** (0.32)	0.82** (0.32)	0.86*** (0.32)	0.78** (0.32)
(correct in part 6)					
Ability ²	0.04* (0.02)	0.04** (0.02)	0.04** (0.02)	0.04** (0.02)	0.05** (0.02)
Medium quality incentives (T-0.25 and T-0.50)	2.27*** (0.63)	2.25*** (0.63)	2.27*** (0.63)	2.48*** (0.64)	-4.65 (3.51)
High quality incentives (T-1.00 and T-3.00)	3.69*** (0.64)	3.60*** (0.65)	3.63*** (0.64)	3.68*** (0.64)	-1.38 (3.02)
Overconfidence	0.37* (0.22)				
Risk aversion		-0.04 (0.09)			
Ambiguity aversion			-0.03 (0.08)		
Loss aversion				-0.17* (0.09)	-0.43*** (0.15)
Loss aversion × Medium quality incentives					0.48** (0.23)
Loss aversion × High quality incentives					0.35* (0.20)
Constant	-0.14 (1.31)	1.22 (1.53)	1.14 (1.50)	3.08* (1.66)	7.17*** (2.51)
Observations	197	197	197	197	197
p-value, Medium = High	0.02	0.03	0.03	0.05	--
R-squared	0.59	0.59	0.59	0.59	0.60

Table 8: OLS regressions of the error rate

Specification	(1)	(2)	(3)
Dependent variable = Error rate (incorrect/attempted)			
Ability	-0.05*** (0.01)	-0.04*** (0.01)	-0.08*** (0.02)
Ability ²		0.00* (0.00)	
Medium quality incentives (T-0.25 and T-0.50)		-0.22*** (0.05)	-0.21*** (0.05)
High quality incentives (T-1.00 and T-3.00)		-0.41*** (0.05)	-0.41*** (0.05)
Constant	0.65*** (0.06)	0.84*** (0.05)	0.96*** (0.09)
Observations	197	197	197
p-value, Medium = High	--	0.00	0.00
R-squared	0.13	0.38	0.39

Table 9: OLS regressions of the error rate, with behavioral characteristics

Specification	(1)	(2)	(3)	(4)	(5)
Dependent variable = Error rate (incorrect/attempted)					
Ability	-0.08*** (0.02)	-0.08*** (0.02)	-0.08*** (0.02)	-0.08*** (0.02)	-0.08*** (0.02)
Ability ²	0.00* (0.00)	0.00* (0.00)	0.00* (0.00)	0.00* (0.00)	0.00* (0.00)
Medium quality incentives (T-0.25 and T-0.50)	-0.21*** (0.05)	-0.21*** (0.05)	-0.22*** (0.05)	-0.23*** (0.05)	0.26 (0.25)
High quality incentives (T-1.00 and T-3.00)	-0.40*** (0.05)	-0.40*** (0.05)	-0.40*** (0.05)	-0.41*** (0.05)	-0.06 (0.22)
Overconfidence	0.00 (0.02)				
Risk aversion		0.01 (0.01)			
Ambiguity aversion			0.01 (0.01)		
Loss aversion				0.01* (0.01)	0.03*** (0.01)
Loss aversion × Medium quality incentives					-0.03** (0.02)
Loss aversion × High quality incentives					-0.02 (0.01)
Constant	0.95*** (0.09)	0.87*** (0.11)	0.90*** (0.11)	0.81*** (0.12)	0.53*** (0.18)
Observations	197	197	197	197	197
p-value, Medium = High	0.00	0.00	0.00	0.00	--
R-squared	0.39	0.40	0.40	0.40	0.42

Table 10: Classification of participants by response time

Reward	Average time per problem	Fraction guessed < 6 seconds	Fraction guessed < 7 seconds	Fraction guessed < 10 seconds	Fraction only choosing quality
\$0.00	15.83 (1.85)	0.25 (0.07)	0.33 (0.07)	0.42 (0.07)	0.33 (0.09)
\$0.05	13.73 (1.11)	0.27 (0.05)	0.38 (0.06)	0.48 (0.06)	0.27 (0.08)
\$0.25	18.19 (1.01)	0.13 (0.04)	0.17 (0.04)	0.21 (0.04)	0.50 (0.09)
\$0.50	18.37 (1.06)	0.09 (0.03)	0.16 (0.04)	0.21 (0.04)	0.49 (0.08)
\$1.00	22.83 (1.11)	0.02 (0.01)	0.03 (0.02)	0.06 (0.02)	0.79 (0.07)
\$3.00	23.28 (1.28)	0.02 (0.02)	0.02 (0.02)	0.05 (0.03)	0.93 (0.05)

Standard errors are in parentheses.

Table 11: Ability and expected earnings from quality

Reward	Avg. Seconds per Correct Answer (Part 6)	Expected earnings from quality	Fraction should focus on quality	Mistake	Mistake (dropping first 34 seconds)
\$0.00	26.03 (1.80)	1.31 (0.10)	0.00 (0.00)	0.85 (0.07)	0.81 (0.08)
\$0.05	33.88 (4.35)	1.73 (0.13)	0.00 (0.00)	0.82 (0.07)	0.82 (0.07)
\$0.25	23.82 (1.28)	4.85 (0.27)	0.12 (0.06)	0.88 (0.06)	0.88 (0.06)
\$0.50	25.64 (2.05)	8.43 (0.57)	0.79 (0.07)	0.49 (0.08)	0.49 (0.08)
\$1.00	28.93 (3.18)	13.69 (0.83)	0.97 (0.03)	0.21 (0.07)	0.21 (0.07)
\$3.00	28.48 (3.49)	39.94 (3.17)	1.00 (0.00)	0.07 (0.05)	0.07 (0.05)

Standard errors are in parentheses.

Appendix A: Instructions

GENERAL INSTRUCTIONS

This is an experiment in the economics of decision-making. Various research agencies have provided funds for this research. The instructions are simple.

The experiment will proceed in 7 parts. Each part contains decision problems that require you to make a series of choices that determine your total earnings. The currency used in all parts of the experiment is U.S. Dollars. You have already received a \$7.00 participation fee. Your earnings from 7 parts of the experiment will be added to your participation fee. At the end of today's experiment, you will be paid in private and in cash.

It is very important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your cooperation.

At this time we proceed to Part 1 of the experiment.

PART 1

In this part of the experiment, you will work on your own and have the chance to earn money by solving 2-digit math problems. At the end of the whole experiment, your entire earnings will be paid out to you immediately and in cash.

You will have 5 minutes (300 seconds) for this part. The computer will provide you with up to 60 math problems (one at a time) that you can attempt to solve during this 5 minutes. Each problem will consist of adding 5, two-digit numbers. All of the problems are about the same level of difficulty. You will see the problems one at a time and you will not be able to skip any problems. You will not be able to go back to any problems.

Your earnings for each problem depend on your responses in the following way:

- For each problem you **attempt**, you will receive **\$0.10**.
- For each attempted problem you answer **correctly**, you will receive a bonus of **\$0.50**.
- For each attempted problem you answer **incorrectly**, you will receive a penalty of **-\$0.00**.

Answering a problem correctly means that you have provided the correct answer, for example, $2+2=4$ is correct while $2+2=3$ is incorrect.

The time remaining will be displayed on the overhead. When 5 minutes are up, time will be called. You will not be able to respond to any more problems after time is up because your computer will be on pause. After time is called, you will need to enter "0" to move on to the outcome screen, and the last problem you answer will not count as an attempt. An example of a problem screen is shown below.

Problem 1

Add up the numbers:				
28	39	48	23	25
Your answer:				
<input type="text"/>				
Submit this sum?				
<input type="button" value="Submit"/>				

Note that you will know which problem you are on. The 5 numbers that you should add are listed in the middle of the screen. In this example, you should be adding $28+39+48+23+25$.

Each time you arrive at a new problem, you will have 5 seconds to review it before the submit button appears. After spending at least 5 seconds, the computer will allow you enter your answer. Although you will be required to spend *at least* 5 seconds on each problem, you can also spend *more than* 5 seconds.

Press "Submit" when you are ready to go on to the next problem.

You will not know if you answered any one problem correctly or incorrectly until the end of the experiment, when you will learn your total number of correct and incorrect responses.

The actual earnings for this part of the experiment will be determined at the end of the experiment, and will be independent of other parts of the experiment.

PART 2

In this part of the experiment, you will be asked to provide a guess about how many of the attempted problems in Part 1 you solved **correctly**. You will receive an additional \$3 if your guess is equal to the number of correct answers that you provided us in Part 1.

Please enter your guess on your screen. Record your answer (and outcome) below. The actual earnings for this part of the experiment will be determined at the end of the experiment, and will be independent of other parts of the experiment.

Use the following table for records:

	Record your Results Here
Number of Problems Attempted	
Guess About the Number of Problems Correct	

PARTS 3-5

In PARTS 3-5 of the experiment, you will be asked to make a series of choices in decision problems. How much you receive will depend partly on **chance** and partly on the **choices** you make.

In each PART, you will see a table with 20 lines. You will state whether you prefer Option A or Option B in each line. You should think of each line as a separate decision you need to make. However, only **one line** in PARTS 4-6 will be the 'line that counts' and will be paid out.

- At the end of the experiment, we will draw a card from a deck of cards numbered 3, 4, 5. Depending on which card is chosen, either PART 3, PART 4, or PART 5 will "count"
- Then, we will draw a card from a deck of cards numbered 1, 2, ..., 20. The number on the card chosen indicates which **line** in that part will be paid out

Because each line is equally likely to be selected, and because you do not know which line will be selected when you make your choices, you should pay close attention to the choices you make in each line. **In some lines, depending on the decisions you make, you may earn up to \$10.**

PART 3

For each line in the table, please state whether you prefer option A or option B. Notice that there are a total of **20 lines** in the table – you should think of each line as a separate decision you need to make.

Your earnings for the selected line depend on which option you chose: If you chose option B in that line, you will receive an amount of money specified by option B – between **\$0.50** and **\$10**, depending on the line. If you chose option A in that line, you will receive either **\$10** or **\$0**. To determine your earnings in the case you chose option A we will randomly draw a ball from a bag containing twenty balls. The balls are either **white** or **orange**, but you do not know the exact number of white and orange balls before you make your decision. Before you draw the ball you choose a color. For example, suppose that you choose white. If the drawn ball is really white, you will receive \$10. If the drawn ball is orange, you will receive \$0.

While you have all the information in the table, you should input all your 20 decisions into the computer. The actual drawing of the ball for this part of the experiment will be determined at the end of the experiment.

Use the following tables for records:

	Record Your Response Here	
CHOOSE YOUR COLOR:	<input type="checkbox"/> WHITE	<input type="checkbox"/> ORANGE

Decision Number	Option A		Option B	Choose A or B
1	\$10.00 with unknown chance	\$0.00 with unknown chance	\$0.50 for sure	
2	\$10.00 with unknown chance	\$0.00 with unknown chance	\$1.00 for sure	
3	\$10.00 with unknown chance	\$0.00 with unknown chance	\$1.50 for sure	
4	\$10.00 with unknown chance	\$0.00 with unknown chance	\$2.00 for sure	
5	\$10.00 with unknown chance	\$0.00 with unknown chance	\$2.50 for sure	
6	\$10.00 with unknown chance	\$0.00 with unknown chance	\$3.00 for sure	
7	\$10.00 with unknown chance	\$0.00 with unknown chance	\$3.50 for sure	
8	\$10.00 with unknown chance	\$0.00 with unknown chance	\$4.00 for sure	
9	\$10.00 with unknown chance	\$0.00 with unknown chance	\$4.50 for sure	
10	\$10.00 with unknown chance	\$0.00 with unknown chance	\$5.00 for sure	
11	\$10.00 with unknown chance	\$0.00 with unknown chance	\$5.50 for sure	
12	\$10.00 with unknown chance	\$0.00 with unknown chance	\$6.00 for sure	
13	\$10.00 with unknown chance	\$0.00 with unknown chance	\$6.50 for sure	
14	\$10.00 with unknown chance	\$0.00 with unknown chance	\$7.00 for sure	
15	\$10.00 with unknown chance	\$0.00 with unknown chance	\$7.50 for sure	
16	\$10.00 with unknown chance	\$0.00 with unknown chance	\$8.00 for sure	
17	\$10.00 with unknown chance	\$0.00 with unknown chance	\$8.50 for sure	
18	\$10.00 with unknown chance	\$0.00 with unknown chance	\$9.00 for sure	
19	\$10.00 with unknown chance	\$0.00 with unknown chance	\$9.50 for sure	
20	\$10.00 with unknown chance	\$0.00 with unknown chance	\$10.00 for sure	

PART 4

For each line in the table, please state whether you prefer option A or option B. Notice that there are a total of **20 lines** in the table – you should think of each line as a separate decision you need to make.

Your earnings for the selected line depend on which option you chose: If you chose option B in that line, you will receive an amount of money specified by option B – between **\$0.50** and **\$10**, depending on the line. If you chose option A in that line, you will receive either **\$10** or **\$0**. To determine your earnings in the case you chose option A we will randomly draw a ball from a bag containing twenty balls. There are **ten orange** and **ten white** balls in the bag. That means that when we draw a ball, there is a 50% chance that it is white and a 50% chance that it is orange. Before you draw the ball you choose a color. For example, suppose that you choose white. If the drawn ball is really white, you will receive \$10. If the drawn ball is orange, you will receive \$0.

While you have all the information in the table, you should input all your 20 decisions into the computer. The actual drawing of the ball for this part of the experiment will be determined at the end of the experiment.

Use the following tables for records:

Record Your Response Here	
CHOOSE YOUR COLOR:	<input type="checkbox"/> WHITE <input type="checkbox"/> ORANGE

Decision Number	Option A		Option B	Choose A or B
1	\$10.00 with 50% chance	\$0.00 with 50% chance	\$0.50 for sure	
2	\$10.00 with 50% chance	\$0.00 with 50% chance	\$1.00 for sure	
3	\$10.00 with 50% chance	\$0.00 with 50% chance	\$1.50 for sure	
4	\$10.00 with 50% chance	\$0.00 with 50% chance	\$2.00 for sure	
5	\$10.00 with 50% chance	\$0.00 with 50% chance	\$2.50 for sure	
6	\$10.00 with 50% chance	\$0.00 with 50% chance	\$3.00 for sure	
7	\$10.00 with 50% chance	\$0.00 with 50% chance	\$3.50 for sure	
8	\$10.00 with 50% chance	\$0.00 with 50% chance	\$4.00 for sure	
9	\$10.00 with 50% chance	\$0.00 with 50% chance	\$4.50 for sure	
10	\$10.00 with 50% chance	\$0.00 with 50% chance	\$5.00 for sure	
11	\$10.00 with 50% chance	\$0.00 with 50% chance	\$5.50 for sure	
12	\$10.00 with 50% chance	\$0.00 with 50% chance	\$6.00 for sure	
13	\$10.00 with 50% chance	\$0.00 with 50% chance	\$6.50 for sure	
14	\$10.00 with 50% chance	\$0.00 with 50% chance	\$7.00 for sure	
15	\$10.00 with 50% chance	\$0.00 with 50% chance	\$7.50 for sure	
16	\$10.00 with 50% chance	\$0.00 with 50% chance	\$8.00 for sure	

17	\$10.00 with 50% chance	\$0.00 with 50% chance	\$8.50 for sure	
18	\$10.00 with 50% chance	\$0.00 with 50% chance	\$9.00 for sure	
19	\$10.00 with 50% chance	\$0.00 with 50% chance	\$9.50 for sure	
20	\$10.00 with 50% chance	\$0.00 with 50% chance	\$10.00 for sure	

PART 5

For each line in the table, please state whether you prefer option A or option B. Notice that there are a total of 20 lines in the table – you should think of each line as a separate decision you need to make.

Your earnings for the selected line depend on which option you chose: If you chose option B in that line, you will receive **\$0**. If you chose option A in that line, you can receive either a loss between -\$1 and -\$20, depending on the line, or a gain of **\$10**. To determine your earnings in the case you chose option A we will randomly draw a ball from a bag containing twenty balls. There are **ten orange** and **ten white** balls in the bag. Before you draw the ball you choose a color. For example, suppose that you choose white. If the drawn ball is really white, you will receive -\$x (the exact amount depends on the line chosen). If the drawn ball is orange, you will receive \$10.

While you have all the information in the table, you should input all your 20 decisions into the computer. The actual drawing of the ball for this part of the experiment will be determined at the end of the experiment.

Use the following tables for records:

Record Your Response Here	
CHOOSE YOUR COLOR:	<input type="checkbox"/> WHITE <input type="checkbox"/> ORANGE

Decision Number	Option A		Option B	Choose A or B
1	-\$1.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
2	-\$2.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
3	-\$3.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
4	-\$4.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
5	-\$5.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
6	-\$6.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
7	-\$7.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
8	-\$8.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
9	-\$9.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
10	-\$10.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
11	-\$11.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
12	-\$12.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
13	-\$13.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
14	-\$14.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
15	-\$15.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
16	-\$16.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
17	-\$17.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
18	-\$18.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
19	-\$19.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	
20	-\$20.00 with 50% chance	\$10.00 with 50% chance	\$0.00 for sure	

PART 6

In this part of the experiment, you will work on your own and have the chance to earn money by solving 2-digit math problems.

You will have 2 and a half minutes (150 seconds) for this part. The computer will provide you with up to 30 math problems (one at a time) that you can attempt to solve during this 2 and a half minutes. Each problem will consist of adding 5, two-digit numbers. All of the problems are about the same level of difficulty. You will see the problems one at a time and you will not be able to skip any problems. You will not be able to go back to any problems.

Your earnings for each problem depend on your responses in the following way:

- For each problem you answer **correctly**, you will receive **\$0.50**.
- There is no penalty for incorrect problems, and no earnings from attempted problems that are not correct

Answering a problem correctly means that you have provided the correct answer, for example, $2+2=4$ is correct while $2+2=3$ is incorrect.

The time remaining will be displayed on the overhead. When 2 and a half minutes are up, time will be called. You will not be able to respond to any more problems after time is up because your computer will be on pause. After time is called, you will need to enter “0” to move on to the outcome screen, and the last problem you answer will not count as an attempt.

The actual earnings for this part of the experiment will be determined at the end of the experiment, and will be independent of other parts of the experiment.

PART 7

In this part of the experiment, you will be asked to provide a guess about how many of the attempted problems in Part 6 you solved **correctly**. You will receive an additional \$3 if your guess is equal to the number of correct answers that you provided us in Part 6.

Please enter your guess on your screen. Record your answer (and outcome) below. The actual earnings for this part of the experiment will be determined at the end of the experiment, and will be independent of other parts of the experiment.

Use the following table for records:

Record your Results Here	
Number of Problems Attempted	
Guess About the Number of Problems Correct	

Earnings Sheet

		Result	Your Earnings
PART 1 – Adding Numbers	Number of Problems Attempted		
	Number of Problems Correct		
	Number of Problems Incorrect		
PART 2 – Guessing Game	Guess About the Number of Problems Correct		
	Actual Number of Problems Correct		
PARTS 3 - 5 Line Games	Which Part is Chosen? <input type="checkbox"/> PART 3 <input type="checkbox"/> PART 4 <input type="checkbox"/> PART 5		
	Line that Counts (1-20)		
	Color Chosen (White or Orange)		
PART 6 – Adding Numbers	Number of Problems Attempted		
	Number of Problems Correct		
	Number of Problems Incorrect		
PART 7 - Guessing Game	Guess About the Number of Problems Correct		
	Actual Number of Problems Correct		
TOTAL:			\$

Appendix B: Additional Information

Table B1: Elicitation of ambiguity aversion preferences

Choice	Option A ambiguous option	Option B safe option
# 1	\$0.00 or \$10.00 with unknown chance	\$0.50 for sure
# 2	\$0.00 or \$10.00 with unknown chance	\$1.00 for sure
# 3	\$0.00 or \$10.00 with unknown chance	\$1.50 for sure
# 4	\$0.00 or \$10.00 with unknown chance	\$2.00 for sure
# 5	\$0.00 or \$10.00 with unknown chance	\$2.50 for sure
# 6	\$0.00 or \$10.00 with unknown chance	\$3.00 for sure
# 7	\$0.00 or \$10.00 with unknown chance	\$3.50 for sure
# 8	\$0.00 or \$10.00 with unknown chance	\$4.00 for sure
# 9	\$0.00 or \$10.00 with unknown chance	\$4.50 for sure
# 10	\$0.00 or \$10.00 with unknown chance	\$5.00 for sure
# 11	\$0.00 or \$10.00 with unknown chance	\$5.50 for sure
# 12	\$0.00 or \$10.00 with unknown chance	\$6.00 for sure
# 13	\$0.00 or \$10.00 with unknown chance	\$6.50 for sure
# 14	\$0.00 or \$10.00 with unknown chance	\$7.00 for sure
# 15	\$0.00 or \$10.00 with unknown chance	\$7.50 for sure
# 16	\$0.00 or \$10.00 with unknown chance	\$8.00 for sure
# 17	\$0.00 or \$10.00 with unknown chance	\$8.50 for sure
# 18	\$0.00 or \$10.00 with unknown chance	\$9.00 for sure
# 19	\$0.00 or \$10.00 with unknown chance	\$9.50 for sure
# 20	\$0.00 or \$10.00 with unknown chance	\$10.00 for sure

Participants choose between an ambiguous option A (\$0.00 or \$10.00 with unknown chance) or a safe option B (a certain amount for sure).

Table B2: Elicitation of risk preferences

Choice	Option A ambiguous option	Option B safe option
# 1	-\$1.00 or \$10.00 with 50% chance	\$0.00 for sure
# 2	-\$2.00 or \$10.00 with 50% chance	\$0.00 for sure
# 3	-\$3.00 or \$10.00 with 50% chance	\$0.00 for sure
# 4	-\$4.00 or \$10.00 with 50% chance	\$0.00 for sure
# 5	-\$5.00 or \$10.00 with 50% chance	\$0.00 for sure
# 6	-\$6.00 or \$10.00 with 50% chance	\$0.00 for sure
# 7	-\$7.00 or \$10.00 with 50% chance	\$0.00 for sure
# 8	-\$8.00 or \$10.00 with 50% chance	\$0.00 for sure
# 9	-\$9.00 or \$10.00 with 50% chance	\$0.00 for sure
# 10	-\$10.00 or \$10.00 with 50% chance	\$0.00 for sure
# 11	-\$11.00 or \$10.00 with 50% chance	\$0.00 for sure
# 12	-\$12.00 or \$10.00 with 50% chance	\$0.00 for sure
# 13	-\$13.00 or \$10.00 with 50% chance	\$0.00 for sure
# 14	-\$14.00 or \$10.00 with 50% chance	\$0.00 for sure
# 15	-\$15.00 or \$10.00 with 50% chance	\$0.00 for sure
# 16	-\$16.00 or \$10.00 with 50% chance	\$0.00 for sure
# 17	-\$17.00 or \$10.00 with 50% chance	\$0.00 for sure
# 18	-\$18.00 or \$10.00 with 50% chance	\$0.00 for sure
# 19	-\$19.00 or \$10.00 with 50% chance	\$0.00 for sure
# 20	-\$20.00 or \$10.00 with 50% chance	\$0.00 for sure

Participants choose between a risky option A (\$0.00 or \$10.00 with 50% chance) or a safe option B (a certain amount for sure).

Table B3: Elicitation of loss aversion preferences

Choice	Option A risky option	Option B safe option
# 1	-\$0.50 or \$5.00 with 50% chance	\$0.00 for sure
# 2	-\$1.00 or \$5.00 with 50% chance	\$0.00 for sure
# 3	-\$1.50 or \$5.00 with 50% chance	\$0.00 for sure
# 4	-\$2.00 or \$5.00 with 50% chance	\$0.00 for sure
# 5	-\$2.50 or \$5.00 with 50% chance	\$0.00 for sure
# 6	-\$3.00 or \$5.00 with 50% chance	\$0.00 for sure
# 7	-\$3.50 or \$5.00 with 50% chance	\$0.00 for sure
# 8	-\$4.00 or \$5.00 with 50% chance	\$0.00 for sure
# 9	-\$4.50 or \$5.00 with 50% chance	\$0.00 for sure
# 10	-\$5.00 or \$5.00 with 50% chance	\$0.00 for sure
# 11	-\$5.50 or \$5.00 with 50% chance	\$0.00 for sure
# 12	-\$6.00 or \$5.00 with 50% chance	\$0.00 for sure
# 13	-\$6.50 or \$5.00 with 50% chance	\$0.00 for sure
# 14	-\$7.00 or \$5.00 with 50% chance	\$0.00 for sure
# 15	-\$7.50 or \$5.00 with 50% chance	\$0.00 for sure
# 16	-\$8.00 or \$5.00 with 50% chance	\$0.00 for sure
# 17	-\$8.50 or \$5.00 with 50% chance	\$0.00 for sure
# 18	-\$9.00 or \$5.00 with 50% chance	\$0.00 for sure
# 19	-\$9.50 or \$5.00 with 50% chance	\$0.00 for sure
# 20	-\$10.00 or \$5.00 with 50% chance	\$0.00 for sure

Participants choose between a risky option A (which has 50% chance of losing certain amount) or a safe option B (\$0.00 for sure).

Table B4: Correlation between ambiguity, risk and loss-aversion

Variable	Observations	Average	Correlations		
			Ambiguity aversion	Risk aversion	Loss aversion
Ambiguity aversion	122	11.69	1		
[# safe choices]		(3.08)			
Risk aversion	122	10.89	0.67***	1	
[# safe choices]		(2.79)			
Loss aversion	122	14.97	0.30***	0.35***	1
[# safe choices]		(2.95)			

*** significant at 1%; standard errors are in parentheses.

Table B5: OLS regressions of quantity, with individual treatment dummies

Specification	(1)	(2)	(3)	(4)
Dependent variable = Quantity (number attempted)				
Ability	0.26 (0.30)	0.42 (0.26)	0.39 (0.26)	0.28 (0.26)
Treatment T-\$0.05		2.16 (2.45)	1.54 (2.44)	29.09** (11.60)
Treatment T-\$0.25		-7.74*** (2.44)	-8.88*** (2.46)	43.31*** (13.09)
Treatment T-\$0.50		-7.45*** (2.37)	-8.42*** (2.38)	25.72* (13.19)
Treatment T-\$1.00		-11.79*** (2.43)	-11.88*** (2.40)	31.30*** (10.89)
Treatment T-\$3.00		-12.45*** (2.50)	-13.44*** (2.51)	33.60*** (12.88)
Loss aversion			0.54** (0.24)	3.06*** (0.63)
Loss aversion x Treatment T-\$0.05				-2.05** (0.80)
Loss aversion x Treatment T-\$0.25				-3.62*** (0.87)
Loss aversion x Treatment T-\$0.50				-2.48*** (0.88)
Loss aversion x Treatment T-\$1.00				-3.13*** (0.77)
Loss aversion x Treatment T-\$3.00				-3.32*** (0.87)
Constant	17.74*** (1.96)	23.13*** (2.39)	15.87*** (3.94)	-18.04** (8.74)
Observations	197	197	197	197
R-squared	0.13	0.39	0.39	0.43

** significant at 5%, *** significant at 1%; standard errors are in parentheses.

Table B6: OLS regressions of quality, with individual treatment dummies

Specification	(1)	(2)	(3)	(4)
Dependent variable = Quality (number correct)				
Ability	1.51*** (0.11)	1.48*** (0.10)	1.49*** (0.10)	1.54*** (0.11)
Treatment T-\$0.05		-0.16 (0.94)	0.06 (0.94)	-9.86** (4.63)
Treatment T-\$0.25		1.89** (0.94)	2.29** (0.95)	-13.96*** (5.23)
Treatment T-\$0.50		2.24** (0.91)	2.58*** (0.92)	-6.43 (5.26)
Treatment T-\$1.00		3.41*** (0.94)	3.44*** (0.93)	-6.12 (4.35)
Treatment T-\$3.00		3.58*** (0.96)	3.92*** (0.97)	-7.19 (5.14)
Loss aversion			-0.19** (0.09)	-0.87*** (0.25)
Loss aversion x Treatment T-\$0.05				0.72** (0.32)
Loss aversion x Treatment T-\$0.25				1.11*** (0.35)
Loss aversion x Treatment T-\$0.50				0.66* (0.35)
Loss aversion x Treatment T-\$1.00				0.70** (0.31)
Loss aversion x Treatment T-\$3.00				0.79** (0.35)
Constant	0.51 (0.71)	-1.14 (0.92)	1.39 (1.52)	10.46*** (3.49)
Observations	197	197	197	197
R-squared	0.51	0.58	0.59	0.61

** significant at 5%, *** significant at 1%; standard errors are in parentheses.

Table B7: OLS regressions of the error rate, with individual treatment dummies

Specification	(1)	(2)	(3)	(4)
Dependent variable = Error rate (incorrect/attempted)				
Ability	-0.05*** (0.01)	-0.04*** (0.01)	-0.04*** (0.01)	-0.05*** (0.01)
(correct in part 6)				
Treatment T-\$0.05	0.04 (0.07)	0.03 (0.07)	0.58* (0.33)	
Treatment T-\$0.25	-0.19*** (0.07)	-0.21*** (0.07)	0.94** (0.37)	
Treatment T-\$0.50	-0.20*** (0.06)	-0.22*** (0.07)	0.27 (0.38)	
Treatment T-\$1.00	-0.37*** (0.07)	-0.37*** (0.07)	0.30 (0.31)	
Treatment T-\$3.00	-0.40*** (0.07)	-0.42*** (0.07)	0.32 (0.37)	
Loss aversion		0.01 (0.01)	0.05*** (0.02)	
Loss aversion x			-0.04* (0.02)	
Treatment T-\$0.05			-0.08*** (0.02)	
Loss aversion x			-0.04 (0.03)	
Treatment T-\$0.25			-0.05** (0.02)	
Loss aversion x			-0.04 (0.02)	
Treatment T-\$0.50			-0.05** (0.02)	
Loss aversion x			-0.05** (0.02)	
Treatment T-\$1.00			-0.05** (0.02)	
Loss aversion x			-0.05** (0.02)	
Treatment T-\$3.00			-0.05** (0.02)	
Constant	0.65*** (0.06)	0.82*** (0.07)	0.67*** (0.11)	0.09 (0.25)
Observations	197	197	197	197
R-squared	0.13	0.39	0.39	0.43

** significant at 5%, *** significant at 1%; standard errors are in parentheses.