

## Chapman University Chapman University Digital Commons

ESI Working Papers

Economic Science Institute

8-23-2019

# Lying and Shirking Under Oath

Nicolas Jacquemet  
*Université de Lorraine*


Alexander James  
*University of Alaska, Anchorage*

Stéphane Luchini  
*Aix-Marseille University*

James J. Murphy  
*Chapman University, [murphy@uaa.alaska.edu](mailto:murphy@uaa.alaska.edu)*

Jason F. Shogren  
*University of Wyoming*

Follow this and additional works at: [https://digitalcommons.chapman.edu/esi\\_working\\_papers](https://digitalcommons.chapman.edu/esi_working_papers)

 Part of the [Econometrics Commons](#), [Economic Theory Commons](#), and the [Other Economics Commons](#)

### Recommended Citation

Jacquemet, N., James, A., Luchini, S., Murphy, J., & Shogren, J. F. (2019). Lying and shirking under oath. ESI Working Paper 19-19. Retrieved from [https://digitalcommons.chapman.edu/esi\\_working\\_papers/278/](https://digitalcommons.chapman.edu/esi_working_papers/278/)

This Article is brought to you for free and open access by the Economic Science Institute at Chapman University Digital Commons. It has been accepted for inclusion in ESI Working Papers by an authorized administrator of Chapman University Digital Commons. For more information, please contact [laughtin@chapman.edu](mailto:laughtin@chapman.edu).

---

# Lying and Shirking Under Oath

## **Comments**

ESI Working Paper 19-19

# Lying and Shirking Under Oath\*

Nicolas Jacquemet<sup>†</sup>    Alexander James<sup>‡</sup>    Stéphane Luchini<sup>§</sup>  
James Murphy<sup>¶</sup>    Jason F. Shogren<sup>||</sup>

August 23, 2019

## Abstract

This study explores whether an oath to honesty can reduce both shirking and lying among crowd-sourced internet workers. Using a classic coin-flip experiment, we first show that a substantial majority of Mechanical Turk workers both shirk and lie when reporting the number of heads flipped. We then demonstrate lying can be reduced by first asking each worker to swear voluntarily on his or her honor to tell the truth in subsequent economic decisions. The oath, however, did not reduce shirking as measured by time-at-coin-flip-task, although it did increase the time they spent answering a demographic survey. Conditional on response, MTurk shirkers and liars were less likely to agree to an ex post honesty oath. Our results suggest oaths may help elicit more truthful behavior in on-line crowd-sourced environments.

*Keywords:* Experimental Economics; Honesty; Intrinsic Costs; Field Experiment; Solemn Oath; Mechanical Turk; MTurk; Lying; Shirking; Labor economics

*JEL Classification:* D91; C81; C90; C93; D01; D82; J20; J30; J40

---

\*Thanks to Marie Claire Villeval for helpful comments and suggestions at a preliminary stage. Thanks to Kyle Borash for technical support. All errors are our own.

<sup>†</sup>Paris School of Economics, Université de Lorraine (BETA), 3 Place Carnot, 54035 Nancy, France

<sup>‡</sup>Department of Economics, University of Alaska Anchorage, Anchorage, AK 99508

<sup>§</sup>Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS, Centre de la Vieille Charité, 13236 Marseille Cedex 02, France

<sup>¶</sup>Department of Economics, University of Alaska Anchorage, Anchorage, AK. 99508 and The Economic Science Institute, Chapman University, Orange, CA., 92866.

<sup>||</sup>Department of Economics, University of Wyoming, Laramie, WY. 82071

# 1 Introduction

Some workers shirk and they often misrepresent the value of their contribution when monitoring is weak (see Akerlof, 1970; Maskin, 1999). When they can blur the link between payoffs and effort with strategically insincere statements, the shirkers gain at society's expense. On net, society is earning fewer benefits at greater costs. The basic idea of mechanism design is to address this inefficiency. In theory, a well-designed demand-revealing mechanism attempts to address this challenge. A typical mechanism uses financial incentives to overcome this strategic behavior and reveal preferences/skill by either providing a menu that links rewards and effort (see e.g., Laffont and Tirole, 1988), or by separating what people say from what they earn in pay (e.g., Vickrey, 1961).

In practice, however, not all demand revealing mechanisms need to be financial. The solemn oath to honesty is an ancient—and time tested—mechanism designed to eliminate strategic misbehavior by asking a person to commit to the truth (see Tyler, 1835; Kiesler and Sakumura, 1966; Joule and Beauvois, 1998; Joule, Fabien, and Bernard, 2007). Evidence suggests that in economic contexts people under oath lie less, coordinate better, trust each other, and cooperate more (Jacquemet et al. 2013, 2017, 2018a, 2018b; Stevens, Tabatabaei, and Lass, 2013; 2018; Hergueux et al., 2016; Beck *et al.*, 2016).<sup>1</sup>

The open question we address herein is whether a non-financial honesty oath works to reduce shirking and lying about payoffs in an anonymous financial task. Using a unique field experiment carried out on Amazon Mechanical Turk (MTurk), we explore whether the

---

<sup>1</sup>Ellingsen and Johannesson (2004) similarly find that making a promise enhances cooperation in a hold-up game.

honesty oath works for reducing both shirking and lying. MTurk workers “MTurkers” are independent piece-rate contractors popular nowadays in many economics experiments (see for example Horton, Rand, and Zeckhauser, 2011; Garbarino, Slonim and Villeval, 2019). But there remains significant concerns about the quality of data that is collected in online environments where monitoring is difficult (Chandler, Mueller, and Paolacci, 2014; Fleisher, Mead, and Huange, 2015; Smith *et al.* 2016). According to Ford (2017), “The potential for encountering speeders and cheaters appear to be significant for research using MTurk”.<sup>2</sup>

Using a variant of the classic coin flipping-for-payment experiment (see e.g., Bucciol and Piovesan, 2011; Abeler *et al.*, 2014)<sup>3</sup>, we build upon earlier work and utilize an experimental design that allows for greater privacy and allows us to more precisely estimate intrinsic lying costs. We recruited MTurkers and asked them to perform a simple task: flip a coin ten times. They were paid ten cents for each head they reported flipped. Our design offers three advantages over other coin-flip/die role experiments that took place in a lab or over the phone (e.g., Abeler *et al.*, 2014; Beck *et al.*, 2016).<sup>4</sup> First, we had complete anonymity and privacy; it was clearly impossible for us to observe whether an MTurker performed the task or truthfully reported the outcome. Moreover, MTurkers are only identified by their user name which we cannot possibly link to their actual identity. In contrast, although participants were anonymous in Abeler *et al.*, the authors acknowledge that the participants’ might not

---

<sup>2</sup>It should be pointed out that some existing research shows that subjects on MTurk behave similarly to those in more traditional laboratory settings and are at least as attentive to instructions as other subjects (Paolacci, Chandler, and Ipeirotis, 2010; Horton, Rand, and Zeckhauser, 2011; Buhrmester, Kwang, and Gosling, 2011; Goodman, Cryder, and Cheema, 2013; Hauser and Schwarz, 2016).

<sup>3</sup>Fischbacher and Föllmi-Heusi (2013) similarly examine lying behavior by asking subjects to role dice rather than flip a coin.

<sup>4</sup>In the Beck *et al.* study subjects were required to sign an honesty oath. This is potentially problematic as oaths are considered to be most effective when freely signed (Jacquemet, Joule, Luchini, and Shogren, 2013).

perceive this since the researcher already had their land line phone number (see footnote 9 of their paper). By administering the experiment in this environment, we relieve subjects of additional social pressures that can bias estimated intrinsic lying costs. Second, we asked subjects to carry out a time-consuming task (flipping a coin ten times). By using a hidden timer, we can identify subjects that surely did not actually carry out the requested task, i.e., we can identify people that clearly shirked. This also allows us to evaluate the behavior of workers that are known to be acting dishonestly, rather than focusing purely on group-level effects. Third, MTurk workers also tend to be more representative of the U.S. population than many other in-person samples (Berinsky, Huber and Lenz, 2012) reducing questions of external validity.

We are not the first to measure lying behavior online using the common coin-flipping technique. Similar to our analysis, Garbarino, Slonim, and Villeval (2019) use a mind coin-tossing game in which subjects on MTurk are paid according to their prediction of what their toss will be and document clear lying behavior. Suri and Mason (2011) similarly ask MTurkers to roll a die and offer to pay them  $\$0.25 \times$  the outcome of the roll. They find that too many fives and sixes—and too few ones and twos—are rolled in this environment. Our most obvious and significant contribution to this literature is that we evaluate the effectiveness of both ex-ante and ex-post honesty oaths to induce truthful responses and reveal both dishonest behavior (shirking) and dishonest answers (lying) in an online labor market.

Four key results emerge from our experimental design regarding shirking and lying. First, do MTurkers shirk? Yes, we find that about 58 percent of MTurkers completed the coin flipping task in under 30 seconds, which we estimate is the minimum required time to flip a coin ten

times. Second, do MTurkers lie? Yes, consistent with other studies, our results show that the distribution of flips is skewed significantly towards more “heads” than one would expect from random draws of a binomial distribution. Third, does the oath reduce lying and shirking? Yes and no. Yes, the oath reduced lying, though it did not eliminate it. No, our results show that the oath did not reduce shirking (a similar percentage of MTurkers completed the task in under 30 seconds). Fourth, we find that young and male MTurkers were more likely to shirk and lie without the oath, and cynical MTurkers (those who said they do not trust others, or believe other people are liars and cheaters) were more likely to lie themselves.

## 2 Experimental Design

The experiment was administered on Amazon’s Mechanical Turk, which is an online platform that connects employers (or “requesters”) with potential workers. The tasks (or “human intelligence tasks” (HITs)) that MTurkers complete are typically simple and straightforward (e.g., answering a questionnaire) and can be completed privately and anonymously at any location.

MTurkers were randomly assigned to one of two treatments (Oath and No Oath). To avoid possible temporal bias, we released a batch of 200 HITs every two hours until the budget was exhausted. MTurkers could only participate once. The only restriction on participation was that the MTurkers had to be at least 18 years of age.

The experiment was advertised on Mechanical Turk with the title, “Answer a survey consisting of approximately 25 questions” and with the description, “Carry out a simple coin flipping exercise, then answer a set of survey questions”. This page also read, “We are conducting

an academic survey. We need your input to help us understand general human behaviors. Select the link below to complete the survey. At the end of the survey, you will receive a code to paste into the box below to receive credit for taking the survey”.

After agreeing to the task, MTurkers were redirected to a Qualtrics survey. The first page thanked them for participating, and elicited informed consent to participate in an academic study. This page also informed them that, “You will be paid \$1.00 for completing the survey and will have an opportunity to earn additional money during the survey. At the end of the survey you will be given a code to enter into Mechanical Turk to receive this payment. This survey should take no more than 3 minutes to complete”. Additional information was provided, such as a requirement for MTurkers to be at least eighteen years old and that all of the information obtained would be kept anonymous and confidential.

Upon agreeing to the task, MTurkers in the Oath treatment were shown a screen that read, “Before we begin, do you swear upon your honor to answer the following questions truthfully? (You will be allowed to continue with this survey regardless of your answer to this question)”. MTurkers did not sign the oath, they simply clicked “yes” or “no”. While agreeing to the oath was optional, only three MTurkers did not agree to this oath. The oath was administered prior to MTurkers knowing anything about the task at hand, or the financial penalty associated with being honest. Regardless of how MTurkers answered the oath question, the next screen read “Flip a coin 10 times. You will be paid a bonus of 10 cents per head that is flipped. How many heads did you flip?”.<sup>5</sup> MTurkers were then provided a drop-down menu where they

---

<sup>5</sup>While the monetary reward for lying is small, note that \$1.00 amounts to roughly thirty minutes of work for a typical MTurker (the median wage on Mechanical Turk is just \$2.00 per hour according to Hara *et al.* (2018)).



selected their answer. The No Oath MTurkers were not presented with the oath screen, and instead went directly to the coin-flipping task. After the coin-flipping task, in both treatments, MTurkers then completed a survey which contained demographic questions and a subset of questions from the World Values Survey.<sup>6</sup> The survey was identical across the two treatments except that No Oath MTurkers were given an “ex-post” oath immediately following the coin-flipping stage that read, “Do you swear upon your honor that the number of heads you reported flipping is truthful? (You will be paid according to the number of heads you reported flipping regardless of your answer to this question).” MTurkers in the Oath treatment were instead asked “Did swearing upon your honor to tell the truth affect the number of heads you reported flipping?”. MTurkers were required to answer each question to proceed to the next question, and were not allowed to go back and change previous answers. A hidden timer recorded how long it took to complete the coin-flipping task and the survey; MTurkers did not know they were being timed. This feature allows us to identify those MTurkers who could not possibly have carried out the coin flipping task because it was completed too quickly. Hereafter we refer to these MTurkers as “shirkers” as they clearly did not complete the task as requested. Conversely, for MTurkers that used a sufficient amount of time to have possibly completed the task are referred to as “workers”.<sup>7</sup>

A common concern with Mechanical Turk is that workers are not thinking carefully about the questions being asked, or that automated programs (“bots”) that are designed to mimic human behavior contaminate results. We address these concerns in two ways. First, about

---

<sup>6</sup>See online appendix.

<sup>7</sup>Note that we can only identify those MTurkers that certainly did not complete the task, and cannot identify those that certainly did complete it.

halfway through the survey, we asked MTurkers, “We want to make sure you are not a robot. What is the number two plus the number three equal to?”. Those MTurkers who gave an incorrect answer to this question were dropped from the analysis.<sup>8</sup> Second, since assignment to treatment was random, inattentive MTurkers (or even non-human MTurkers) were equally likely to be assigned to a particular treatment, and therefore should not bias the estimated treatment effect.

### 3 Results

We collected data from 1,410 MTurkers. Of these, we dropped the 43 (3%) MTurkers who failed to correctly answer the question about what the sum of 2+3 equals. In addition, one MTurker who spent 1,700 seconds answering the coin flipping question was dropped to minimize outlier bias when we examine flipping times. This leaves 1,366 observations (681 in the Oath treatment and 685 in the No Oath treatment).

Table 1 gives variable names and definitions. Table 2 gives the summary statistics. Across the Oath and No Oath treatments, MTurkers were predominantly male ( $\approx 60\%$ ), white ( $\approx 63\%$ ) and physically located in the USA (82%). The average age was 35 ( $\sigma=10.7$ ). None of the MTurkers in the No Oath treatment were Hispanic, whereas 4.7% were Hispanic in the Oath treatment. Across all other characteristics, MTurkers in the Oath and No Oath treatments were similar, suggesting that the random assignment of MTurkers to treatment was successful.

---

<sup>8</sup>This method of bot detection is commonly used on Mechanical Turk (Crump, McDonnell, and Gureckis, 2013).

### 3.1 Do MTurkers Shirk and Lie?

MTurkers are rarely asked to take an oath before completing a task, we therefore begin by using the No Oath treatment as a baseline to address this question of whether MTurkers shirk and/or lie. Shirking is the failure to perform the agreed upon task, i.e. not flipping the coin ten times as instructed. Some, but not all, shirking can be detected at the individual level based on the amount of time an individual spent flipping the coin. The Qualtrics survey included a hidden timer that recorded how long an individual spent on the page that described the coin flipping task and compensation, and asked MTurkers to report the observed number of heads. To determine the minimum amount of time needed to complete this task, we asked 28 students in a large university class to flip a coin that had been provided to them 10 times as quickly as possible, count the number of heads, and enter the result online in the same way MTurkers in the experiment reported their answers. The fastest that any student completed the coin flipping task was 27 seconds, with a mean of 102 seconds. Based on this, we concluded that it was impossible to complete the task in less than 30 seconds (note that in the classroom pilot, students already had a coin available and were prepared to flip before the timer started, whereas for the MTurkers, flipping time also included time spent getting a coin). We define a “quick” response as one that was completed in less than thirty seconds and we label those workers as “shirkers.” A “slow” response was completed in at least thirty seconds and we refer to these MTurkers as “workers”.

Figure 1 shows the distribution of flipping time by treatment (for display purposes, the figure omits those MTurkers who took more than 200 seconds). In the No Oath treatment, 42.6% (n=292) of MTurkers completed the task in less than 30 seconds. Because the task was

done in private, we have no way of knowing whether a “slow” MTurker actually performed the task. Therefore, these “quick” responses represent a lower bound on shirking. The data clearly indicate that, yes, a nontrivial number MTurkers did not flip the coin as instructed, and did shirk.

We define lying as intentionally making a false statement, which in this context means an MTurker misreported the actual number of heads observed. Because all decisions were made in private, it is impossible to detect whether any particular individual lied. However, we can detect lying in the aggregate by comparing the distribution of reported outcomes to the expected distribution if all reports were truthful.

Privacy and anonymity, combined with the bonus payment of 10 cents per reported head, created an incentive for MTurkers to report a larger number of heads than they actually observed since there are no financial or social consequences for lying. Pooling across shirkers and workers, in the No Oath treatment, MTurkers reported an average of 6.33 heads (Table 3) and we reject the null hypothesis that this is less than or equal to the expected mean of five ( $p$ -value = 0.000) if all reporting were truthful. Panel (d) of Figure 2 compares the reported outcomes for the No Oath treatment with the truthful distribution, pooled across workers and shirkers. The modal response ( $n=298$ , 21.8%) was six (a small lie), and 18% of MTurkers ( $n=122$ ) reported flipping ten heads in a row (a big lie). This result is similar to Fischbacher and Föllmi-Heusi (2013) who find that 20% of subjects “lie to the fullest extent possible” in their die-rolling experiment. The binomial probability of observing 10 heads is 0.1%, which implies that we should expect to observe this outcome no more than once if all MTurkers reported truthfully. Even if we were to drop all MTurkers who reported flipping ten heads,

the average number of reported heads flipped is 5.5, which is still statistically different from five ( $p$ -value = 0.000). Similarly, Panel F of Table 3 shows that for “slow” MTurkers, who had enough time to complete the task, the mean number of heads is 5.76, with 9.4% reporting 10 heads. We therefore conclude that, yes, on average MTurk workers lie. These lies come in two primary forms. Some of these lies are plausible (i.e. reporting 6) and others are implausible “big” lies that maximize the worker’s earnings (reporting 10).

The conclusion that MTurkers lie is robust across both the “shirkers” (i.e., MTurkers who completed the task in under 30 seconds) and the “workers” for whom the time spent on the flipping task was sufficient for them to have possibly done the task (Table 3). Not surprisingly, in the No Oath treatment the shirkers reported more heads than the slow workers (6.79 vs 5.98,  $p=0.000$ ). Shirkers are also significantly more likely to report observing 10 heads (29.1% vs 9.4%,  $p=0.000$ ). Panels b and c of Figure 2 show that while the modal responses for shirkers were five and ten, for workers the mode was six. This finding is consistent with the idea that slow workers were engaged in a more sophisticated form of deception that was more time consuming and exerted more cognitive effort.

### **3.2 Does an oath reduce shirking & lying?**

Next we examine whether agreeing to a solemn oath increased the intrinsic cost of lying. Table 3 shows the unconditional results. The average number of heads reported flipped by MTurkers in the Oath treatment was 6.05, which is significantly less than the number reported flipped by No Oath MTurkers (6.33,  $p$ -value = 0.008). That the mean exceeded 5 ( $p$ -value = 0.000) indicates that the oath is not a panacea for truth telling. The oath also worked to reduce the

number of MTurkers who reported flipping ten heads in a row ( $p$ -value = 0.006). Eighty eight (12.9%) MTurkers in the Oath treatment reported flipping ten heads in a row where as 122 (17.8%) of No Oath MTurkers did so.

The first column of Figure 2 gives the distribution of heads flipped for Oath and No Oath treatments pooled across shirkers and workers. The “truthful distribution” is provided for comparison purposes. A one-tailed Kolmogorov-Smirnov test confirms that the distribution in the Oath treatment is significantly different than that for the No Oath treatment (at the 10% confidence level). A test of normality (Shapiro-Wilk test) confirms the two distributions are significantly different from the normal distribution. However, dropping MTurkers that reported flipping ten heads, we cannot reject the null hypothesis that the oath had no effect. This implies that the oath largely worked by decreasing the number of MTurkers that told big, obvious lies—like reporting to have flipped ten heads in a row and is consistent with the idea that telling big lies is more costly than telling small lies (Fischbacher and Föllmi-Heusi, 2008; Suri and Mason, 2011).

Table 3 also shows that the oath had little effect on the time MTurkers spent answering the coin-flipping question. Further, the oath had no effect on the probability an MTurker shirks (responds in to the coin-flipping task in less than thirty seconds). However, we do observe that the oath induced MTurkers to spend approximately 30 additional seconds filling out the survey (net of the time spent on the coin-flipping task). See Figure 3 for the distribution of response times by Oath and No Oath MTurkers. This amounts to roughly a  $30/214=14\%$  increase in survey duration. One speculative interpretation for these contrasting findings is that workers view their responses to survey questions as potentially consequential; their

answers may directly influence any conclusions drawn from the study. In contrast, the coin-flipping task may be viewed as a time-consuming random number generator that can be costlessly avoided by strategically picking a number between zero and ten.<sup>9</sup>

Shirkers—those MTurkers that spent less than thirty seconds on the coin-flipping task—certainly did not carry out the task as requested whereas workers may have carried it out. We now examine the effect of the oath separately for these two groups of people. Table 3 shows that the oath was similarly effective at reducing the number of heads reported flipped by both shirkers and workers. Figure 2 shows the distributions of heads flipped for shirkers and workers. The oath reduced the probability a shirker reports ten heads, and increased the probability of reporting five heads. The oath had a similar effect for workers, but for this group the distribution is less bi-modal. However, according to Kolmogorov-Smirnov tests, we cannot reject the null hypothesis that the oath had no effect on the distribution of heads flipped for either shirkers or workers. According to Shapiro-Wilk tests, we also reject the null hypothesis that each of the distributions in Figure 2 are similar to a normal distribution.

At the end of the survey, we asked Oath MTurkers: “Did swearing upon your honor to tell the truth affect the number of heads you reported flipping”. We estimate the probability of answering “yes” as a function of age, gender, race, the number of heads flipped, the indicator for flipping ten heads in a row, and the indicator for shirking. We find that both black and white MTurkers were less likely to say that the oath mattered, whereas we find the opposite for Asian MTurkers.<sup>10</sup> This is consistent with the earlier findings that the effect of the oath was

---

<sup>9</sup>A potentially useful variant of this study would be to ask subjects to carry out consequential tasks under oath.

<sup>10</sup>These results are available from the authors upon request.

statistically insignificant for both White and Black MTurkers. We also find that, conditional on reporting to have flipped ten heads, heads flipped enters positively. This implies that, the oath may have been most effective for people that told small lies (that may have otherwise told large lies). Taken together, there is some evidence that subjects were aware of the effect of the oath.

### 3.3 Further Analysis

### 3.4 Heterogeneous Effects

We test for heterogeneous oath effects by regressing the number of heads reported flipped on the oath, exogenous MTurker characteristics, and the interaction of the two. Table 4 shows that being young or male both increase the number of heads reported flipped. Interestingly, while the oath enters negatively, the interaction term  $\text{White} \times \text{Oath}$  enters positively. A Wald test confirms that  $\text{Oath} + \text{White} \times \text{Oath}$  is not statistically different from zero ( $p$ -value = 0.78), implying the oath had no effect on white MTurkers. However, a similar test confirms that the oath was weakly significant for male MTurkers ( $p$ -value = 0.08). This result is not surprising given that we have already shown that males tend to tell bigger lies more often than females. Similar results are found by examining the probability that an MTurker reported flipping ten heads in a row. For this outcome, in the second column of Table 4, the oath enters negatively and is significant at the 10.4% confidence level. Wald-tests confirms that the oath was only effective for male MTurkers ( $p$ -value = 0.04).

We explore other potential sources of heterogeneity in Table 5. However, these results



should be viewed with some added caution as the subject characteristics considered here are potentially endogenous and this makes causal inference more challenging. Each row gives the results of separate regressions of the number of heads flipped on the oath, the respective subject characteristic, and the interaction of the two. The “Direct” effect of each characteristic is given in column 1 and the interaction of each characteristic and the oath is given in the second column. For example, believing that it is often justified to claim public benefits that one is not legally entitled to is positively associated with the number of heads reported flipped (coefficient=0.160,  $p$ -value = 0.000) and the oath did not uniquely affect this group of people (coefficient on interaction=-.033,  $p$ -value =.502).

For nearly all subject characteristics, we find no evidence of heterogeneous effects of the oath. However, we do find that the oath is more effective for people making less than \$30,000. We also find some evidence that the oath is less effective for people who report going to church once a week, but the other measures of religiosity do not interact with the oath in a meaningful or significant way. Turning to the direct effects, we observe that people who think it is often justified to cheat, steal, bribe, or fail to pay due taxes are more likely to report a high number of heads. We similarly find that people who trust others are less likely to report a high number of heads. For example, a person that believes the majority of foodstamp recipients are “acting fraudulently or otherwise cheating the U.S. welfare system” report flipping 0.465 more heads. Given the average number of heads reported flipped among Oath MTurkers is 6.05, this amounts to a 7.6% increase in the number of heads reported flipped. These specific results reinforce the idea that dishonest people view others as being similarly dishonest.

### 3.4.1 Ex-Post Oath

Immediately after answering the coin-flipping question, No Oath MTurkers were asked if they would “swear upon their honor that the number of heads they reported flipped was accurate”. We estimate the relationship between the likelihood of agreeing to this ex-post oath and both reported heads flipped and response time.

Table 6 shows that MTurkers who reported flipping a large number of heads were less likely to agree to the ex-post oath. This result is statistically significant and robust to conditioning on observed MTurker heterogeneity. We also find that MTurkers who reported flipping ten heads in a row were less likely to agree to the ex-post oath (columns 3 and 4). Interestingly, the last column of Table 6 shows that the effect of heads flipped remains negative after conditioning its effect on the indicator for flipping ten heads as well as the indicator for shirking “Quick”. This implies that even MTurkers that lied a little (did not report flipping ten heads) were less likely to agree to the ex-post oath than people who reported more honest answers. Also, conditional on heads reported flipped, shirkers (Quick) MTurkers were less likely to agree to the ex-post oath. This suggests that MTurkers who did not carry out the coin-flipping task may have viewed their behavior as dishonest, regardless of the answer they gave.

Dropping the 70 MTurkers from the No Oath treatment ( $N = 685$ ) that did not agree to the ex-post oath, the average MTurker in the No-Oath treatment reported flipping 6.08 heads. While this number is significantly greater than 5 ( $p=0.000$ ), it is also significantly less than the number of heads reported flipped by MTurkers that did not agree to the ex-post oath (8.5) in this treatment. Taken together, this result suggests that asking MTurkers to swear on their honor following the completion of a task may help identify shirkers and liars.

## 4 Conclusion

We test whether MTurk workers lie and shirk, and explore whether a solemn oath to be honest can reduce the prevalence of both. We asked roughly 1,400 MTurkers to flip a coin ten times and report the number of heads they flipped. MTurkers were paid a bonus of ten cents for each head reported flipped. In this environment, there is a clear and direct cost associated with telling the truth. Although we cannot tell whether individual MTurkers told the truth, we can observe whether groups of people lied on average because we know the underlying truthful distribution. Using a hidden timer, we are also able to identify shirkers—those MTurkers that answered the coin-flipping question too quickly to have actually carried out the task.

We find that MTurkers both lie and shirk and that the oath reduces lying, but has little effect on shirking. Whereas no-oath MTurkers reported to have flipped 6.33 heads on average, MTurkers under oath reported just 6.05 heads (a statistically significant difference). We also find that the oath increased the time MTurkers spent answering a set of post-coin flipping survey questions, which is consistent with the idea that agreeing to a solemn oath caused MTurkers to answer survey questions more carefully. Examining MTurker characteristics, being young and male significantly increases the likelihood of reporting a high number of heads flipped. Measures of religiosity (such as regularly attending church or believing in God) had no effect on either the number of heads reported flipped, or the effectiveness of the oath.<sup>11</sup> We do however find that MTurkers who think other people are liars and cheaters are more likely to lie themselves. We also find that MTurkers who shirked and lied were less likely to

---

<sup>11</sup>This result is consistent with existing literature which finds that religiosity is not significantly correlated with honesty (see for example Huelsman, Piroch, and Wasieleski, 2006).

agree to an ex-post honesty oath. Taken together, these results suggest that oaths may be an effective tool to induce honesty, and identify dishonest workers in online, crowd sourced environments.

It is possible that the failure of the oath to reduce shirking was because workers took an oath to honesty, rather than an oath to task (i.e., a commitment to actually perform the task as described). Future research should test whether an “oath to task” can reduce shirking. In addition, it is possible that one reason we observe a large amount of shirking on the coin-flipping task, but a significant effect of the oath on the amount of time spent on the survey, is because workers perceive the survey as meaningful or consequential, whereas reporting the number of heads flipped is viewed as less so. Future research could explore this conjecture further.

## 5 References

- Abeler, J., Becker, A., and Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, 113, 96-104.
- Alm, J., McClelland, G.H., and Schulze, W.D. (1992). Why do people pay taxes? *Journal of Public Economics*, 48(1), 21-38.
- Akerlof, G. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3), 488-500.
- Andreoni, J., Erard, B., and Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature*, 36(2), 818-860.
- Beck, T., Bührn, C., Frank, B., and Khachatryan, E. (2018). Can honesty oaths, peer interaction, or monitoring mitigate lying? *Journal of Business Ethics*, 1-18.
- Berinsky, A.J., Huber, G.A., and Lenz, G.S. (2012). Evaluating online labor markets for experimental research: Amazon.com’s Mechanical Turk. *Political Analysis*, 20(3), 351-368.
- Bordignon, M. (1993). A fairness approach to income tax evasion. *Journal of Public Economics*, 52(3), 345-362.
- Buccioli, A. and Piovesan, M. (2011). Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology*, 32(1), 73-78.

- Buhrmester, M., Kwang, T., and Gosling, S.D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high quality data? *Perspectives on Psychological Science*, 6(1), 3-5.
- Carlsson, F., Kataria, M., Krupnick, A., Lampi, E., Löfgren, Å., Qin, P., and Sterner, T. (2013). The truth, the whole truth, and nothing but the truth—A multiple country test of an oath script. *Journal of Economic Behavior and Organization*, 89, 105-121.
- Charness, G., and Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3), 817-869.
- Chandler, J., Mueller, P. and Paolacci, G. (2014). Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behavior research methods*, 46(1), 112-130.
- Crump, M.J., McDonnell, J.V. and Gureckis, T.M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PloS one*, 8(3), e57410.
- de-Magistris, T., and Pascucci, S. (2014). The effect of the solemn oath script in hypothetical choice experiment survey: A pilot study. *Economics Letters*, 1123(2), 252-255.
- Dreber, A. and Johannesson, M. (2008). Gender differences in deception. *Economics Letters*, 99(1), 197-199.
- Ellingsen, T. and Johannesson, M. (2004). Promises, Threats, and Fairness. *The Economic Journal*, 114(495), 397-420.

- Fleischer, A., Mead, A.D., and Huang, J. (2015). Inattentive responding in MTurk and other online samples. *Industrial and Organizational Psychology*, 8(2), 196-202.
- Fischbacher, U., and Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11(3), 525-547.
- Ford, J.B. (2017). Amazon’s Mechanical Turk: a comment. *Journal of Advertising*, 46(1), 156-158.
- Garbarino, E., Slonim, R., and Villeval, M.C. (2019). Loss aversion and lying behavior. *Journal of Economic Behavior and Organization*, 158, 379-393.
- Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 95(1), 384-394.
- Goodman, J.K., Cryder, C.E. and Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, 26(3), 213-224.
- Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., and Bigham, J.P. (2018, April). A data-driven analysis of workers’ earnings on Amazon Mechanical Turk. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p 449). ACM.
- Hauser, D. J., and Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior research methods*, 48(1), 400-407.

- Hergueux, J., Jacquemet, J., Luchini, S. and Shogren, J. (2016). Leveraging the honor code: Public goods contributions under oath. Working Paper.
- Horton, J.J., Rand, D.G., and Zeckhauser, R.J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 14(3), 399-425.
- Huelsman, M.A., Piroch, J., and Wasieleski, D. (2006). Relation of religiosity with academic dishonesty in a sample of college students. *Psychological Reports*, 99(3), 739-742.
- Hurkens, S. and Kartik, N. (2008). Would I lie to you? On social preferences and lying aversion. *Experimental Economics*, 12(2), 180-192.
- Jacquemet, N., Joule, R.V., Luchini, S. and Shogren, J.F. (2013). Preference elicitation under oath. *Journal of Environmental Economics and Management*, 65(1), 110-132.
- Jacquemet, N., James, A., Luchini, S. and Shogren, J.F. (2017). Referenda under oath. *Environmental and Resource Economics*, 67(3), 479-504.
- Jacquemet, N., Luchini, S., Shogren, J.F. and Zylbersztejn, A. (2018a). Coordination with communication under oath. *Experimental Economics*, 21(3), 627-649.
- Jacquemet, N., Luchini, S., Rosaz, J. and Shogren, J. (2018b). Truth telling under oath. *Management Science*, 65(1), 426-438.
- Joule, R., and Beauvois, J. (1998). La soumission librement consentie. Presses Universitaires de France, Paris.



- Joule, R.V., Girandola, F., and Bernard, F. (2007). How can people be induced to willingly change their behavior? The path from persuasive communication to binding communication. *Social and Personality Psychology Compass*, 1(1), 493-505.
- Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies*, 1359-1395.
- Kiesler, C., and Sakumura, J. (1996). A test of a model for commitment. *Journal of Personality and Social Psychology*, 3(3), 349-353.
- Laffont, J.J. and Tirole, J. (1988). The dynamics of incentive contracts. *Econometrica: Journal of the Econometric Society*, 1153-1175
- Maskin, E.S. (1999). Nash equilibrium and welfare optimality. *The Review of Economic Studies*, 66(1), 23-38.
- Mason, W. and Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, 44(1), 1-23.
- Paolacci, G. and Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23(3), 184-188.
- Paolacci, G., Chandler, J. and Ipeirotis, P.G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision making*, 5(5), 411-419.
- Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives*, 21(1), 25-48.

- Smith, S.M., Roster, C.A., Golden, L.L., and Albaum, G.S. (2016). A multi-group analysis of online survey respondent data quality: Comparing a regular USA consumer panel to MTurk samples. *Journal of Business Research*, 69(8), 3139-3148.
- Stevens, T.H., Tabatabaei, M. and Lass, D. (2013). Oaths and hypothetical bias. *Journal of Environmental Management*, 127, 135-141.
- Suri, S., Goldstein, D.G., and Mason, W.A. (2011). Honesty in an online labor market. In *Workshops at the twenty-fifth aaai conference on artificial intelligence*.
- Torgler, B. (2006). The importance of faith: Tax morale and religiosity. *Journal of Economic Behavior and Organization*, 61(1), 81-109.
- Tyler, J.E. (1835). *Oaths, their Origin, Nature, and History*. John W. Parker.

## 6 Appendix

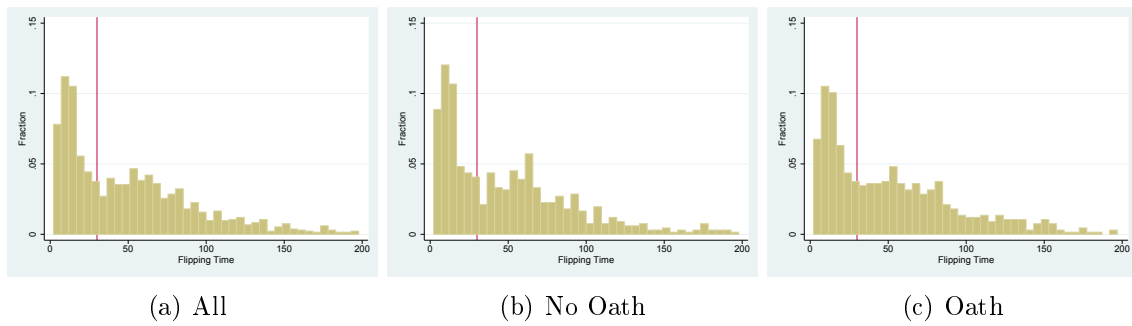
### 6.1 Tables & Figures

Table 1: Variable Definitions

Outcome	Survey Question #	Definition
Heads Flipped	Q3	Number of heads reported flipped. Ranges from 0 to 10
Male	Q8	= 1 for male workers
Age	Q19	Reported age of worker
White	Q28	= 1 for White workers
Black	Q28	= 1 for Black workers
Hispanic	Q28	= 1 for Hispanic workers
Asian	Q28	= 1 for Asian workers
Other	Q28	=1 for non White, Black, Hispanic, Asian.
USA	Q25	= 1 for workers living in the U.S.A.
Justified Benefits	Q57	Is it ever justified to claim government benefits you are not entitled to? Ranges from 0 (never justified) to 10 (always justified).
Justified Transport	Q57	Is it ever justified to avoid paying for public transport? Ranges from 0 (never justified) to 10 (always justified).
Justified Steal	Q57	Is it ever justified to steal? Ranges from 0 (never justified) to 10 (always justified).
Justified Taxes	Q57	Is it ever justified to cheat on taxes? Ranges from 0 (never justified) to 10 (always justified).
Justified Bribe	Q57	Is it ever justified to accept a bribe? Ranges from 0 (never justified) to 10 (always justified).
Trust People	Q41	Dummy variable =1 for workers who think "Most people can be trusted" and 0 for workers that think you "need to be very careful" in dealing with people.
SNAP Cheat	Q29	=1 if more than 50% of foodstamp recipients are believed to be "acting fraudulently, or otherwise "cheating the U.S. welfare system"
Heads Guessed	Q30	How many heads will the average worker report to have flipped? Variable ranges from 0 to 10.
God	Q51	= 1 for workers that believe in god.
Hell	Q52	= 1 for workers that believe in hell.
No Church	Q49	= 1 for workers that report going to church "Never, practically never".
Low Church	Q49	= 1 for workers that report going to church "Once a month".
Med Church	Q49	= 1 for workers that report going to church "Once a week".
High Church	Q49	= 1 for workers that report going to church "More than once a week".
No Religion	Q26	= 1 for workers that "Do not belong to a denomination".
Hindu	Q26	= 1 for workers that are Hindu.
Muslim	Q26	= 1 for workers that are Muslim.
Jew	Q26	= 1 for workers that are Jewish.
Catholic	Q26	= 1 for workers that are Catholic.
Protestant	Q26	= 1 for workers that are Protestant.
High Income	Q21	=1 for workers that report a household income greater than \$99,000.
Low Income	Q21	= 1 for workers that report a household income less than \$30,000.
Satisfied Financial	Q43	Ranges from 1 (not satisfied) to 10 (satisfied) "with the financial situation of your household".
Satisfied Generally	Q39	Ranges from 1 (not satisfied) to 10 (satisfied) "with your life as a whole these days".
Conservative	Q22	=1 for workers that said they were a 9 or 10 on a 10 point scale where 1=liberal and 10=conservative.
Liberal	Q22	=1 for workers that said they were a 1 or 2 on a 10 point scale where 1=liberal and 10=conservative.
Flipping Time	NA	Time to answer the coin flipping question.
Duration	NA	Total time to complete the survey, less flipping time.
2 + 3 Correct	Q60	Answer to the question, "What is the number two plus the number three equal to?"

**Note:** See the online Appendix for a copy of the survey that was administered. Survey question numbers listed in the second column correspond to the oath treatment. "Other" includes Hawaiian and Pacific Islander, and non White, Hispanic, Black, and Asian.

Figure 1: Flipping Time: Distributions by Oath



Note: 35 workers that spent more than 200 seconds on the coin-flipping question were dropped to construct these figures.

Table 2: Summary Statistics

Outcome	No Oath			Oath		
	Mean	Min	Max	Mean	Min	Max
Heads Flipped	6.331	1	10	6.058	0	10
Male	.627	0	1	.593	0	1
Age	35.04	18	82	35.12	18	84
White	.630	0	1	.627	0	1
Black	.056	0	1	.061	0	1
Hispanic	0	0	0	.047	0	1
Asian	.055	0	1	.057	0	1
Other	.256	0	1	.207	0	1
USA	.820	0	1	.825	0	1
Justified Benefits	2.023	0	10	2.02	0	10
Justified Transport	2.420	0	10	2.505	0	10
Justified Steal	1.394	0	10	1.350	0	10
Justified Taxes	1.943	0	10	1.998	0	10
Justified Bribe	1.643	0	10	1.625	0	10
Trust People	.505	0	1	.484	0	1
SNAP Cheat	.290	0	1	.289	0	1
Heads Guessed	6.44	1	10	6.35	0	10
God	.582	0	1	.574	0	1
Hell	.398	0	1	.402	0	1
No Church	.550	0	1	.552	0	1
Low Church	.099	0	1	.098	0	1
Med Church	.109	0	1	.117	0	1
High Church	.035	0	1	.036	0	1
No Religion	.490	0	1	.496	0	1
Hindu	.124	0	1	.117	0	1
Muslim	.016	0	1	.014	0	1
Jew	.011	0	1	.019	0	1
Catholic	.127	0	1	.126	0	1
Protestant	.162	0	1	.165	0	1
High Income	.207	0	1	.234	0	1
Low Income	.405	0	1	.417	0	1
Satisfied Financial	6.501	1	10	6.393	1	10
Satisfied Generally	5.357	1	10	5.177	1	10
Conservative	.091	0	1	.096	0	1
Liberal	.211	0	1	.245	0	1
Flipping Time	56.87	2.25	971	57.9	2.179	987.057
Duration	214	53.7	2722	243	49	2057
2 + 3 Correct	97.44	0	1	96.46	0	1
<i>N</i>	685			681		

**Note:** See Table 1 for variable definitions.

Table 3: Effect of the Oath

Variable	Avg Heads	Flipped 10(%)	Quick	Flipping Time	Duration
<u>All Workers</u>					
Oath	6.06	.129	.400	57.90	243.14
No Oath	6.33	.178	.426	56.87	214.23
<i>p</i> -value	.008	.006	.170	.389	.002
<i>N</i>	1,366	1,366	1,366	1,366	1366
<u>Shirkers</u>					
Oath	6.49	.241	-	14.14	221
No Oath	6.79	.291	-	13.62	203
<i>p</i> -value	.072	.093	-	.196	.138
<i>N</i>	565	565	-	565	565
<u>Workers</u>					
Oath	5.76	.053	-	87.17	257
No Oath	5.98	.094	-	89.00	222
<i>p</i> -value	.041	.014	-	.637	.003
<i>N</i>	801	801	-	801	801

**Note:** *p*-value corresponds to a one-tailed t-test of equality between oath and no oath values. “Shirkers” reported the number of heads flipped in less than 30 seconds. The “Workers” reported this information in more than 30 seconds.

Table 4: Oath Effects: Interactions

Variable	Heads Flipped	Flipped 10(%)	Quick	Flipping Time	Duration
Oath	-.674 (.475)	-.638† (.392)	-.272 (.290)	5.76 (14.54)	10.37 (39.85)
Age	-.028*** (.007)	-.024*** (.007)	-.018*** (.004)	.690*** (.210)	-.352 (.515)
Male	.362** (.166)	.254** (.123)	.171* (.101)	-4.88 (5.25)	-15.51 (14.80)
White	.041 (.214)	-.226* (.134)	-.094 (.116)	-1.23 (7.29)	-68.44*** (18.38)
Black	.089 (.413)	-.109 (.253)	-.243 (.224)	2.19 (9.77)	-50.86* (26.27)
Asian	.221 (.390)	-.206 (.259)	.038 (.225)	.415 (13.22)	-73.83*** (22.78)
Hispanic	.349 (.389)	-.057 (.348)	-.154 (.253)	3.23 (8.67)	-34.94 (45.74)
Age×Oath	.003 (.011)	.007 (.010)	.006 (.007)	-14.15 (.298)	.483 (.719)
Male×Oath	-.093 (.230)	-.089 (.178)	-.011 (.143)	-7.28 (7.87)	5.99 (21.40)
White × Oath	.543* (.304)	.376* (.213)	.018 (.170)	4.16 (9.27)	-7.64 (26.27)
Black×Oath	-.177 (.562)	.143 (.384)	-.092 (.321)	15.27 (17.42)	13.06 (41.54)
Asian×Oath	-.115 (.538)	.081 (.399)	-.303 (.321)	10.61 (16.50)	56.72 (43.03)
Constant	7.043*** (.323)	-.123 (.260)	.427** (.202)	36.37*** (10.12)	286.47*** (29.38)
N	1,366	1,366	1,366	1,366	1,366
R <sup>2</sup> (Pseudo R <sup>2</sup> )	.035	.039	.020	.018	.033
Total effect of the oath by group:					
Oath + Male×Oath	-.768* (.438)	-.727** (.361)	-.284 (.267)	-1.519 (12.508)	16.371 (34.12)
Oath + White×Oath	-.130 (.476)	-.261 (.420)	-.254 (.291)	9.921 (14.326)	2.735 (38.351)
Oath + Black×Oath	-.852 (.628)	-.495 (.497)	-.364 (.374)	21.036 (18.159)	23.437 (41.458)
Oath + Asian×Oath	-.790 (.607)	-.556 (.509)	-.576 (.378)	16.380 (18.566)	67.099 (46.519)

**Note:** The first columns under the heading “Heads Flipped” gives the results from an OLS regression of the number of heads flipped with robust standard errors. The second column under the heading “Flipped 10” corresponds to a probit regression of the likelihood of reporting ten heads flipped. †, \*, \*\*, \*\*\* correspond to the 10.4%, 10%, 5%, and 1% confidence levels, respectively. The bottom rows give the total effect of the oath by group. Standard errors are computed using Stata’s *lincom* command.

Table 5: Who Lies: Endogenous Covariates

Independent Variable:	Direct Coefficient (Std. Err.)	Interacted Coefficient (Std. Err.)
<b>Self Reported Honesty</b>		
Justified Benefits	.160*** (.036)	-.033 (.050)
Justified Transport	.128*** (.033)	-.005 (.047)
Justified Steal	.074* (.044)	.038 (.063)
Justified Taxes	.093*** (.035)	.059 (.049)
Justified Bribe	.109*** (.040)	.035 (.057)
<b>Trust in Others</b>		
Trust People	-.419*** (.162)	.378* (.227)
SNAP Cheat	.465** (.216)	-.361 (.269)
Heads Guessed	.435*** (.045)	-.012 (.069)
<b>Religiosity</b>		
God	-.154 (.169)	.081 (.229)
Hell	.113 (.169)	-.130 (.233)
No Church	.032 (.169)	.069 (.230)
Low Church	-.659*** (.248)	.760** (.367)
Med Church	-.138 (.283)	.002 (.391)
High Church	-.067 (.475)	.159 (.626)
No Religion	-.152 (.167)	.252 (.226)
Hindu	.375 (.338)	-.471 (.405)
Muslim	-.245 (.786)	.388 (1.016)
Jew	.405 (.801)	.038 (1.072)
Catholic	.412* (.245)	-.268 (.337)
Protestant	-.097 (.210)	-.024 (.289)
<b>Income</b>		
Rich	.148 (.189)	.068 (.262)
Poor	.134 (.167)	-.459** (.230)
Satisfied Financial	.077** (.035)	-.035 (.050)
Satisfied Generally	.037 (.034)	.007 (.047)
<b>Politics</b>		
Conservative	-.269 (.303)	-.023 (.425)
Liberal	.421** (.196)	-.412 (.264)

**Note:** Note that each entry in this table corresponds to a separate regression. Each estimate is conditioned on worker race, gender, age, and the oath indicator. For each regression,  $N=1,366$ . The first column gives the coefficient on each variable. The second column gives the coefficient on each variable interacted with the oath indicator variable. For each regression the outcome variable is the number of heads flipped.

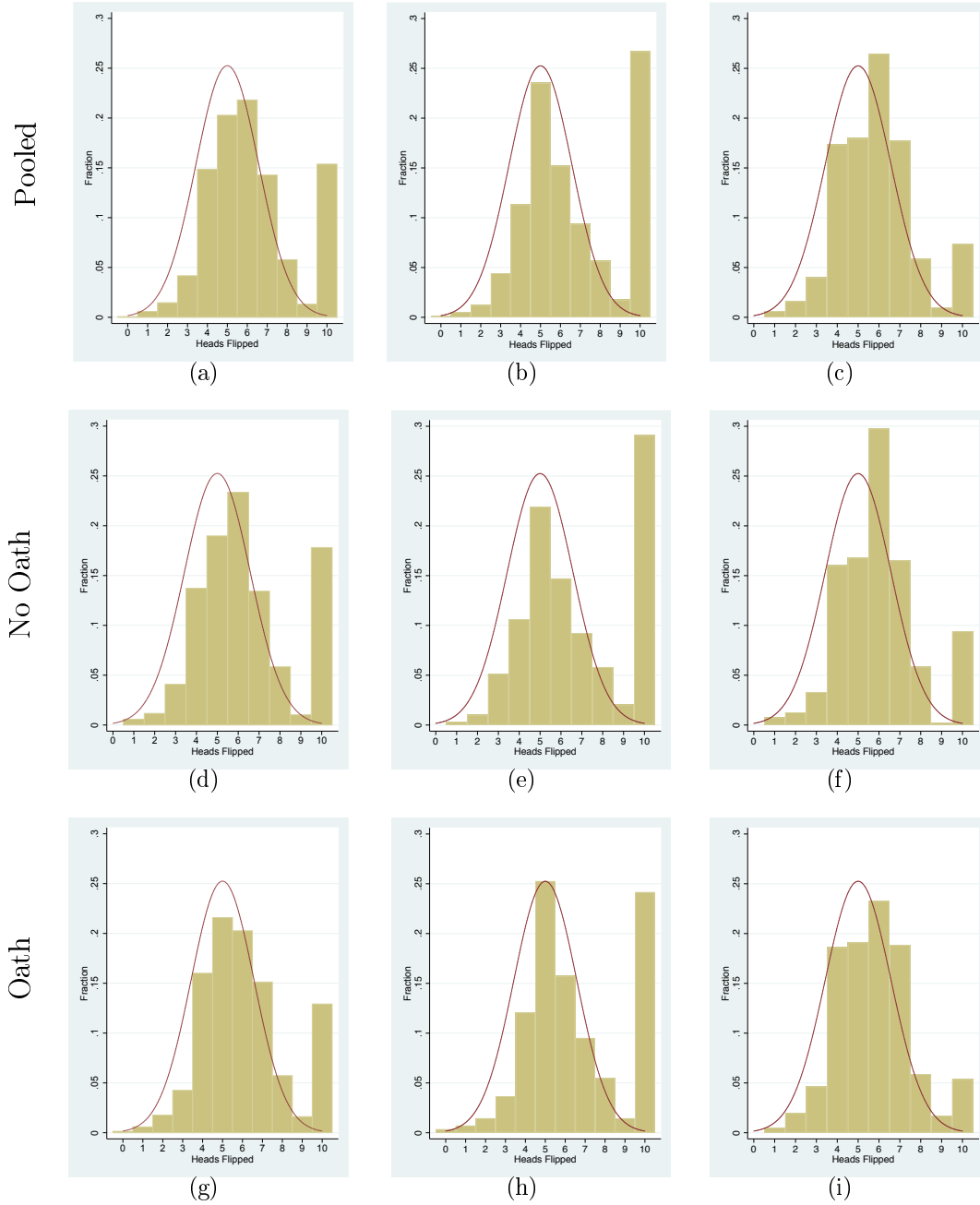


Table 6: Ex-Post Oath

	(1)	(2)	(3)	(4)	(5)	(6)
Heads	-.276*** (.038)	-.280*** (.040)			-.149** (.071)	-.165** (.068)
Flipped 10			-1.28*** (.147)	-1.316*** (.157)	-.688** (.324)	-.489 (.315)
Flipping Time						-.0001 (.0009)
Quick						-.643*** (.176)
Age		-.001 (.006)		-.002 (.006)	-.002 (.006)	-.006 (.006)
Male		.035 (.155)		.045 (.152)	.048 (.154)	.074 (.157)
White		-.080 (.182)		-.110 (.183)	-.100 (.186)	-.078 (.188)
Black		-.059 (.362)		-.030 (.337)	-.053 (.351)	-.067 (.361)
Asian		-.345 (.316)		-.422 (.315)	-.391 (.319)	-.388 (.326)
Constant	3.243*** (.303)	3.384 (.492)	1.664*** (.090)	1.822*** (.327)	2.699*** (.540)	3.21*** (.538)
<i>N</i>	685	685	685	685	685	685
Pseudo $R^2$	.172	.175	.170	.175	.187	.223

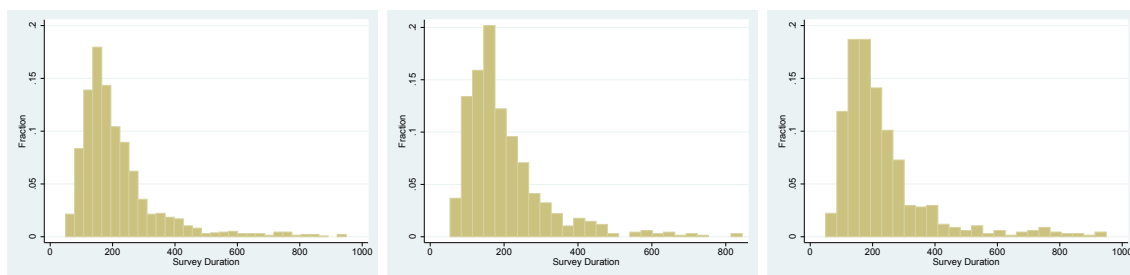
**Note:** This table gives results from probit regressions of the likelihood that a worker in the no-oath treatment agreed to the ex-post oath asking if they swear upon their honor that they answered the coin flipping question honestly. For this analysis, the data is limited to workers not assigned to the ex-ante oath treatment.

Figure 2: Heads Flipped by Flipping Time & Oath  
 Pooled                      Shirkers (Quick)                      Workers (Slow)



Note: Recall that “workers” spent at least thirty seconds answering the coin-flipping task and “shirkers” did not. According to Shapiro-Wilk tests, we reject the null hypothesis that each of the distributions are similar to the normal distribution. For the pooled data, according to Kolmogorov-Smirnov tests, we reject the null hypothesis that the Oath and No Oath distributions are similar ( $p$ -value = 0.10). For shirkers and workers we individually cannot reject the null hypothesis that the Oath distributions are similar to the No Oath distributions.

Figure 3: Survey Duration: Distributions by Oath



(a) All

(b) No Oath

(c) Oath

Note: 13 workers that spent more than 1000 seconds completing the survey (net of flipping time) were dropped to construct these figures.

## 6.2 Survey Instrument

Oath Treatment.

Q7 Welcome and thank you for participating.

STUDY: By proceeding, you have volunteered to participate in a study conducted by the University of Alaska Anchorage Department of Economics. The UAA Principal Investigators are Alexander James and James Murphy.

**VOLUNTARY NATURE OF PARTICIPATION:** Your participation in this study is voluntary. You may stop at any time and you do not have to answer any questions you don't want to. Nothing will happen to you if you choose not to answer any questions or if you decide not to participate.

**PAYMENT:** You will be paid \$1.00 for completing the survey and will have an opportunity to earn additional money during the survey. At the end of the survey you will be given a code to enter into Mechanical Turk to receive this payment. **TIME:** This survey should take no more than 3 minutes to complete.

**RISKS:** There are no foreseeable risks or discomforts.

**BENEFITS:** There are no direct personal benefits for participation other than the cash payment you receive.

**PARTICIPATION:** You must be 18 or over to participate. Your participation is voluntary, and you may withdraw from the study at any time and for any reason. There are no costs to you or any other party.

**CONFIDENTIALITY:** All of the information we obtain from you will be kept confidential.

**CONTACT:** If you have any questions about the research you can reach Alexander James at [ajames27@alaska.edu](mailto:ajames27@alaska.edu) or James Murphy at [jjmurphy2@alaska.edu](mailto:jjmurphy2@alaska.edu). If you have any questions about your rights as a research participant, please contact the UAA Research Compliance office at 907-786-1099.

**CONSENT:** If you agree to participate in this survey, please click the next button. Otherwise, please close your browser window.

---

Page Break

Q61 Timing  
First Click (1)  
Last Click (2)  
Page Submit (3)  
Click Count (4)

---

Q4

**Solemn Oath**

Before we begin, do you swear upon your honor to answer the following questions truthfully?  
(You will be allowed to continue with this survey regardless of your answer to this question).

Yes (1)

No (2)

---

Page Break

---

Q58 Timing  
First Click (1)  
Last Click (2)  
Page Submit (3)  
Click Count (4)

---

Q3 Flip a coin 10 times. You will be paid a bonus of 10 cents per head that is flipped. How many heads did you flip?

▼ 0 heads (1) ... 10 heads (19)

---

Page Break

---

Q8 What is your gender?

- Male (1)
  - Female (2)
- 

Q19 What is your age in years?

	0	20	40	60	80	100
Years Old ()						

---

Q60 What is your marital status?

- Married (1)
  - Living together as married (2)
  - Divorced (3)
  - Separated (4)
  - Widowed (5)
  - Single (6)
-



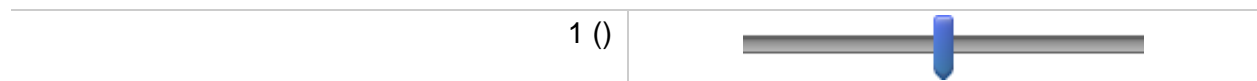
Q62 How many children do you have?

- No children (1)
  - One child (2)
  - Two children (3)
  - Three children (4)
  - Four children (5)
  - Five children (6)
  - Six children (7)
  - Seven children (8)
  - Eight or more children (9)
- 

Q39 All things considered, how satisfied are you with your life as a whole these days? Using the slider bar below on which 1 means you are "completely dissatisfied" and 10 means you are "completely satisfied" where would you put your satisfaction with your life as a whole?

Completely Dissatisfied    Completely Satisfied

1      3      5      6      8      10



Q21 What is your approximate household income?

▼ Less than \$10,000 (1) ... More than \$200,000 (14)

---

Q47 What is your approximate household wealth?

▼ Less than \$10,000 (1) ... More than \$200,000 (14)

Q43 How satisfied are you with the financial situation of your household? Please use this scale to help with your answer where 1 indicates you are "completely dissatisfied" and 10 indicates you are "completely satisfied".

Completely Dissatisfied      Completely Satisfied

1      3      5      6      8      10

1 ( )



Q22 In political matters, people talk of "the left" and "the right". How would you place your views on this scale, generally speaking?

Left

Right

1      3      5      6      8      10

1 ( )



Q23 What is your highest level of education?

- Not a high school graduate (1)
- High school graduate (2)
- Some college (3)
- Associate degree (4)
- Bachelors degree (5)
- Advanced degree (6)



Q25 In which country are you a citizen?

▼ Afghanistan (1) ... Zimbabwe (1357)

---

*Display This Question:*

*If List of Countries = United States of America*

Q53 In which state do you currently reside?

▼ Alabama (1) ... I do not reside in the United States (53)

---

*Display This Question:*

*If List of Countries = United States of America*

Q28 Are you:

- Non-Hispanic White (1)
  - Black or African American (3)
  - American Indian and Alaska Native (4)
  - Asian (5)
  - Native Hawaiian and Other Pacific Islander (6)
  - Hispanic (7)
  - Other (8)
- 

Q60 We want to make sure you are not a robot. What is the number two plus the number three equal to?

- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
-

Q26 Do you belong to a religion or religious denomination? If yes, which one?

- Do not belong to a denomination (1)
  - Roman Catholic (2)
  - Protestant (3)
  - Orthodox (Russian/Greek/etc.) (4)
  - Jew (5)
  - Muslim (6)
  - Hindu (7)
  - Buddhist (8)
  - Other (9) \_\_\_\_\_
- 

Q49 Apart from weddings and funerals, about how often do you attend religious services these days?

- More than once a week (1)
  - Once a week (2)
  - Once a month (3)
  - Only on special holidays (4)
  - Once a year (5)
  - Less often (6)
  - Never, practically never (7)
-

Q50 Apart from weddings and funerals, about how often do you pray?

- Several times a day (1)
  - Once a day (2)
  - Several times each week (3)
  - Only when attending religious services (4)
  - Only on special holidays (5)
  - Once a year (6)
  - Less often (7)
  - Never, practically never (8)
- 

Q51 Do you believe in God?






- Yes (1)
  - No (2)
- 

Q52 Do you believe in hell?

- Yes (1)
  - No (2)
- 

Q57 Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this scale.

Never Justifiable			Always Justifiable		
0	2	4	6	8	10

Claiming government benefits to which you are not entitled ()	
Avoiding a fare on public transport ()	
Stealing property ()	
Cheating on taxes if you have a chance ()	
Someone accepting a bribe in the course of their duties ()	

*Display This Question:*  
*If List of Countries = United States of America*

Q29 What percent of food stamp recipients do you think are acting fraudulently, or otherwise "cheating the U.S. welfare system"?

0 20 40 60 80 100

% of Welfare Cheaters ()	
--------------------------	---

Q41 Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

- Most people can be trusted (1)
- Need to be very careful (2)

Page Break

Q68 Timing  
First Click (1)  
Last Click (2)  
Page Submit (3)  
Click Count (4)

---



Q30 Averaged across all of the responses we received from this MTurk survey, if people were being completely honest, the average number of heads reported flipped should be 5. What do you think will be the average number of heads reported flipped?

▼ 0 heads (1) ... 10 heads (11)

---

Page Break

---



Q31 Did swearing upon your honor to tell the truth affect the number of heads you reported flipping?

- Yes (1)
- No (2)
- I did not swear upon my honor to answer the questions truthfully (3)

End of Block: Default Question Block

---

Start of Block: Block 1



No Oath Treatment.

Q8 Welcome and thank you for participating. **STUDY:** By proceeding, you have volunteered to participate in a study conducted by the University of Alaska Anchorage Department of Economics. The UAA Principal Investigators are Alexander James and James Murphy.

**VOLUNTARY NATURE OF PARTICIPATION:** Your participation in this study is voluntary. You may stop at any time and you do not have to answer any questions you don't want to. Nothing will happen to you if you choose not to answer any questions or if you decide not to participate.

**PAYMENT:** You will be paid \$1.00 for completing the survey and will have an opportunity to earn additional money during the survey. At the end of the survey you will be given a code to enter into Mechanical Turk to receive this payment. **TIME:** This survey should take no more than 3 minutes to complete.

**RISKS:** There are no foreseeable risks or discomforts.

**BENEFITS:** There are no direct personal benefits for participation other than the cash payment you receive.

**PARTICIPATION:** You must be 18 or over to participate. Your participation is voluntary, and you may withdraw from the study at any time and for any reason. There are no costs to you or any other party.

**CONFIDENTIALITY:** All of the information we obtain from you will be kept confidential.

**CONTACT:** If you have any questions about the research you can reach Alexander James at [ajames27@alaska.edu](mailto:ajames27@alaska.edu) or James Murphy at [jjmurphy2@alaska.edu](mailto:jjmurphy2@alaska.edu). If you have any questions about your rights as a research participant, please contact the UAA Research Compliance office at 907-786-1099.

**CONSENT:** If you agree to participate in this survey, please click the next button. Otherwise, please close your browser window.

---

Page Break

Q60 Timing  
First Click (1)  
Last Click (2)  
Page Submit (3)  
Click Count (4)

---



Q6 Flip a coin 10 times. You will be paid a bonus of 10 cents per head that is flipped. How many heads did you flip?

▼ 0 heads (1) ... 10 heads (11)

---

Page Break

---

Q62 Timing  
First Click (1)  
Last Click (2)  
Page Submit (3)  
Click Count (4)

---

Q32 Do you swear upon your honor that the number of heads you reported flipping is truthful?  
(You will be paid according to the number of heads you reported flipping regardless of your  
answer to this question).

- Yes (1)
  - No (2)
- 

Page Break

---

Q7 What is your gender?

- Male (1)
- Female (2)

Q33 What is your age in years?



Q61 What is your marital status?

- Married (1)
- Living together as married (2)
- Divorced (3)
- Separated (4)
- Widowed (5)
- Single (6)

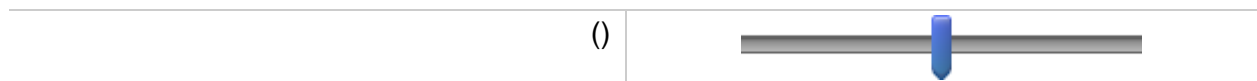
Q63 How many children do you have?

- No Children (1)
  - One child (2)
  - Two children (3)
  - Three children (4)
  - Four children (5)
  - Five children (6)
  - Six children (7)
  - Seven children (8)
  - Eight or more children (9)
- 

Q40 All things considered, how satisfied are you with your life as a whole these days? Using the slider bar below on which 1 means you are "completely dissatisfied" and 10 means you are "completely satisfied" where would you put your satisfaction with your life as a whole?

Completely Dissatisfied    Completely Satisfied

1      3      5      6      8      10



Q49 What is your approximate household income?

▼ Less than \$10,000 (1) ... More than \$200,000 (14)

---

Q50 What is your approximate household wealth?

▼ Less than \$10,000 (1) ... More than \$200,000 (14)

Q45 How satisfied are you with the financial situation of your household? Please use this scale to help with your answer where 1 indicates you are "completely dissatisfied" and 10 indicates you are "completely satisfied".

Completely Dissatisfied    Completely Satisfied

1            3            6            8            10

1 ()



Q37 In political matters, people talk of "the left" and "the right". How would you place your views on this scale, generally speaking?

Left

Right

1            3            5            6            8            10

2 ()





Q36 What is your highest level of education?

- Not a high school graduate (1)
- High school graduate (2)
- Some college (3)
- Associate degree (4)
- Bachelors degree (5)
- Advanced degree (6)



Q39 In which country are you a citizen?

▼ Afghanistan (1) ... Zimbabwe (1357)

---

*Display This Question:*

*If List of Countries = United States of America*



Q59 In which state do you currently reside?

▼ Alabama (1) ... I do not reside in the United States (53)

---

*Display This Question:*

*If List of Countries = United States of America*

Q44 Are you:

- Non-Hispanic White (1)
  - Black or African American (2)
  - American Indian and Alaska Native (3)
  - Asian (4)
  - Native Hawaiian and Other Pacific Islander (5)
  - Hispanic or Latino (6)
  - Other (7)
- 

Q61 We want to make sure you are not a robot. What is the number two plus the number three equal to?

- 1 (1)
  - 2 (2)
  - 3 (3)
  - 4 (4)
  - 5 (5)
-

Q46 Do you belong to a religion or religious denomination? If yes, which one?

- Do not belong to a denomination (1)
  - Roman Catholic (2)
  - Protestant (3)
  - Orthodox (Russian/Greek/etc.) (4)
  - Jew (5)
  - Muslim (6)
  - Hindu (7)
  - Buddhist (8)
  - Other (9) \_\_\_\_\_
- 

Q47 Apart from weddings and funerals, about how often do you attend religious services these days?

- More than once a week (1)
  - Once a week (2)
  - Once a month (3)
  - Only on special holidays (4)
  - Once a year (5)
  - Less often (6)
  - Never, practically never (7)
-

Q48 Apart from weddings and funerals, about how often do you pray?

- Several times a day (1)
  - Once a day (2)
  - Several times each week (3)
  - Only when attending religious services (4)
  - Only on special holidays (5)
  - Once a year (6)
  - Less often (7)
  - Never, practically never (8)
- 

Q53 Do you believe in God?






- Yes (1)
  - No (2)
- 

Q54 Do you believe in hell?

- Yes (1)
  - No (2)
- 

Q56 Please tell me for each of the following actions whether you think it can always be justified, never be justified, or something in between, using this scale.


Never Justifiable			Always Justifiable		
0	2	4	6	8	10

Claiming government benefits to which you are not entitled ()	
Avoiding a fare on public transport ()	
Stealing property ()	
Cheating on taxes if you have a chance ()	
Someone accepting a bribe in the course of their duties ()	

*Display This Question:*  
*If List of Countries = United States of America*

Q40 What percent of food stamp recipients do you think are acting fraudulently, or otherwise "cheating the U.S. welfare system"?

0      20      40      60      80      100

% of Welfare Cheaters ()	
--------------------------	---

Q42 Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

- Most people can be trusted (1)
- Need to be very careful (2)

Page Break

Q69 Timing  
First Click (1)  
Last Click (2)  
Page Submit (3)  
Click Count (4)

---



Q41 Averaged across all of the responses we received from this MTurk survey, if people were being completely honest, the average number of heads reported flipped should be 5. What do you think will be the average number of heads reported flipped?

▼ 0 heads (1) ... 10 heads (11)