Are Women More Resistant to Corruption? A Formal Theory and Evidence from a Laboratory Experiment*

Preliminary Version: Active development of this paper is underway. Results and conclusions may change as research proceeds.

Justin Esarey
Wake Forest University
Department of Politics and International Affairs
justin@justinesarey.com

Leslie Schwindt-Bayer
Rice University
Department of Political Science
schwindt@rice.edu

January 13, 2019

Abstract

Are women more resistant to corruption than men? More importantly, *why* is gender a moderating influence on corruptibility? In this experiment, we use a micro-level laboratory experiment to investigate whether there is a link between gender and corruption. Our design also allows us to determine whether any such link is attributable to differential risk aversion and/or gender differences in moral attitudes toward corruption. The experiment investigates the propensity of subjects to appropriate common pool resources for individual gain. We systematically manipulate the probability of being detected and penalized for such appropriation, which disproportionately deters more risk-averse subjects. We also measure subjects' risk aversion and moral judgments concerning corruption and determine whether these measures explain gender differences in appropriation rates in our experiment.

*Thanks to Rob Carroll (FSU) for comments and suggestions on our formal theoretical framework in an earlier draft of this paper. We also thank Maggie Emerson (Rice University) for her assistance with creating a zTree program.

Introduction

Research has demonstrated a link between women's representation in national legislatures and reduced country levels of corruption since the early 2000's (Dollar, Fisman and Gatti, 2001; Esarey and Chirillo, 2013; Esarey and Schwindt-Bayer, 2018; Stensöta and WÃd'ngnerud, 2018; Stockemer, 2011; Stensöta, 2018; Sundström and Wängnerud, 2016; Swamy et al., 2001; Watson and Moreland, 2014). Numerous theories exist to explain the relationship but most lack empirical validation. In this paper, we present two theories for why more women's representation in government may lead to less corruptionâĂŤthe risk aversion theory and the morality theoryâĂŤand we test them in a laboratory experiment. The risk aversion theory argues that women in government will be less likely to engage in corruption because they are more risk averse than men. The morality theory suggests that women are less likely to engage in corruption because they are more deterred by the immorality of an act like corruption than are men. Neither theory assumes that any differences in women's and men's risk aversion or moral attitudes is biological; in all likelihood, the differences (if they exist) result from gendered socialization and experiences. But, the key point for this study is not where these attitudes come from but whether they exist and whether they lead to women being less likely to engage in corruption than men.

We test these two theories with a micro-level laboratory experiment that used Rice University students as a subject pool. The experiment assesses how varying degrees of risk, measured as the likelihood of being detected and punished, influence subjects' willingness to appropriate common pool resources for personal gain, a classic definition of corruption. More specifically, the experiment has 10 rounds (run three times) where a pool of tokens exists from which participants can either take tokens (and keep them if not caught) or leave them and get their share of 1.33 times the total remaining tokens in the pool distributed to them at the end of a round. In one treatment, participants have no chance of being detected and punished for taking tokens, and in the other, they do. "Punishment" is simply losing all tokens taken in that round and losing out on what remains in the pool. To test the risk aversion theory, we compare the tokens taken by women and men to determine whether more gender differences exist when there is a risk of being caught and punished for taking tokens. The morality theory is tested by whether or not gender differences exist in the no punishment treatmentâĂŤif they do, then it cannot be explained by risk aversion and instead may result from a differential moral aversion to corruption. We use a post-experiment survey to measure subjects' self-perceptions of their general risk aversion and moral aversion to corruption to probe deeper into the findings.

We find that subjects operating under the treatment of detection and punishment for stealing resources do take fewer tokens than those in the treatment of no detection and punishment. However, there are no gender differences in token-taking in the no punishment treatment, and thus no evidence that women are more morally averse to corruption than men. And, women are not significantly different from men in the number of tokens they take under the punishment treatment, meaning that risk aversion does not appear to drive their behavior. That said, a supplemental structural equation model using the experimental data reveals some indirect evidence of gender differences in risk aversion driving token taking. Limitations of the experiment may be obscuring the relationship in the raw experimental results. Future research with a larger subject pool, a stronger punishment treatment, and a clearer connection between token taking and corruption may produce stronger and more discernable experimental differences.

Risk Aversion and Moral Aversion Theories of Gender and Corruption

Research on women's representation and corruption has demonstrated a link between women in government and corruption numerous times. As is already well-known, Dollar, Fisman and Gatti (2001) and Swamy et al. (2001) introduced the idea by showing that more women in national legislatures correlates with reduced levels of country corruption, and later studies have confirmed this. A key question that lingers over this research area, however, is exactly why this relationship exists. A number of theories have been put forth, and they and their strengths and weaknesses have been summarized in other places (see Stensöta and WÃd'ngnerud, 2018, for a thorough review). Here, we focus on the risk aversion theory and the morality theory.

Risk Aversion Theory

The risk aversion theory builds upon evidence largely from studies in economics that have found that women are more risk averse than men (Byrnes, Miller and Schafer, 1999; Croson and Gneezy, 2009; Eckel and Grossman, 2008*b*,*a*). The explanations for why this is the case vary but most focus on gendered socialization and life experiences. Linking risk aversion to corruption, scholars have argued that women are less likely to be involved in corruption because it is a risky activity, and women are more risk averse than men. Several recent studies have provided evidence that supports this indirectly (Barnes and Beaulieu, 2018; Esarey and Schwindt-Bayer, 2018), but none have directly tested women's and men's risk aversion and how it drives their willingness to engage in corruption.

Despite the large body of evidence finding that women are more risk averse than men, a recent study suggested that the finding is not as solid as sometimes suggested. Nelson (2015)conducts a meta-analysis of thirty-five studies and argues that the results are more mixed than usually acknowledged. She also suggests that the studies sometimes overstate the amount of gender differences in risk aversion. Nelson (2015) then criticizes the essentialist nature of such a claim, suggesting that innate risk aversion delegitimizes women and their power.[1] Stensöta (2018) also urges some caution with the risk aversion explanation for corruption, explicitly noting that findings about gender differences in risk aversion in one context (e.g., financial and economic interactions) do not necessarily mean that they will apply in corruption contexts or in all types of corruption contexts (e.g., both petty and grand corruption).

These concerns emphasize just how important establishing whether risk aversion is a plausible explanation for the link between women's representation in government and corruption actually is. Our goal in this paper is to explore a set of testable hypotheses that links gender and corruption through risk aversion and test it directly. We elaborate the specific theoretical implications of the risk aversion theory in the next section.

Moral Aversion Theory

The morality theory suggests that women will be less likely to engage in corruption than men because they are more morally averse to corruption. This theory has its roots in the early studies conducted on women's representation and corruption. Dollar, Fisman and Gatti (2001) argued that one reason that women may be less likely to engage in corruption than men is that they are more ethical than men, citing evidence from a couple of economic studies of the late 1990's (Glover et al., 1997; Reiss and Mitra, 1998). More recent research has also suggested that, in some contexts, women have more ethical attitudes and behaviors than men. One of these contexts is business. A number of studies have found that female students and executives in business environments demonstrate more ethical attitudes and

[1] We don't argue that risk aversion is innate. It likely results from any number of cultural influences, socialization practices, or gendered life circumstances, and few would argue it is somehow biologically based.

behavior than men (Albaum and Peterson, 2006; Roxas and Stoneback, 2004; Ruegger and King, 1992; Valentine and Rittenburg, 2007; Wang and Calvano, 2015).

Swamy et al. (2001) also linked gender differences in morality to corruption. They provided evidence from the World Values Survey that more women than men say bad behaviors, such as littering, driving drunk, lying, and cheating, are unjustifiable, and more directly related to corruption, they showed that more men condone bribe-taking than women. They suggested that this may be one reason that they also find that women in government is associated with reduced corruption.

Critics of the morality theory identify several problems with it. Some note the essentialist nature of it and argue that it puts women in politics in the wrong light by suggesting that they are not there because of their political value and worth (Goetz, 2007). The theory is also criticized for assuming that women are a single category and lack diverse attitudes, behaviors, and ideas (Goetz, 2007). The theory also is solely about women as agents of non-corrupt behavior and ignores the larger context in which women elected to office operate. It ignores the systemic obstacles that women face being marginalized in office and being kept out of the patronage and corruption networks that benefit men (Alexander and Bågenholm, 2018; Goetz, 2007; Tripp, 2006). These critiques are all relevant, and what we want to determine in this paper is precisely whether the morality theory holds any weight for explaining the relationship between gender and corruption or not.

Theoretical Expectations

The controlled environment of the laboratory allows us to form strong expectations about subject behavior, specifically the extent to which our subjects are willing to appropriate public resources under their control for their private gain (one definition of corruption, similar to the one often used by the World Bank (Gray and Kaufmann, 1998; Bhargava, 2005)). In this section, we first lay out a theoretical model of the relationship between risk aversion, (moral) corruption aversion, punishment, and the appropriation of resources. This model allows us to derive directional expectations for behavior that we expect to apply within our experiment. We are in some cases able to fill in specific parameters from our experiment to derive more specific expectations for the behavior of subjects.

While this section develops a formal model from which we deduce our hypotheses, we begin by summarizing these hypotheses. First, if women are more morally averse to corruption than men on average, then we expect the average appropriation of resources by women to be less than that by men when corruption is not punished. Risk aversion will not be an influence on behavior under these conditions; thus, if women are more risk-averse than men but no more morally averse to corruption, we expect to see no difference in their appropriation rates in this case.

Hypothesis 1. *If women are more morally averse to corruption than men, we expect them to appropriate fewer public resources on average (compared to men) when there is no probability that appropriation will be punished.*

Second, when corruption *is* punished with some probability, then risk aversion will be an influence on behavior. Specifically, if women are more risk averse than men, we expect that women will react more strongly than men to the imposition of punishment for corruption. Specifically, we can compare the difference in appropriation rates when punishment is possible versus impossible by gender. If women are more risk-averse than men, we expect this difference to be larger for women than for men regardless of their intrinsic aversion to corruption.

Hypothesis 2. *If women are more risk averse than men, we expect the average difference in appropriation between punishment and no-punishment treatments to be larger for women compared to men.*

Finally, our theory and experimental design allows us to attribute changes in the strength of the gender-corruption link to corruption aversion or risk aversion without directly measuring these concepts. However, the experimental setting also gives us the opportunity to directly measure corruption aversion and risk aversion in our subjects, and to empirically determine whether any gender-corruption relationships that we observe in the experiment are fully explained once these measures are included in a multivariate empirical model of subject behavior. We are thus able to pose two additional hypotheses:

Hypothesis 3. *If women are more morally averse to corruption than men, then:*

a) we expect corruption aversion to be negatively associated with appropriation (regardless of the probability of punishment); and

b) we expect no systematic difference in appropriation between men and women once corruption aversion is held constant in treatments without a probability of punishment.

Hypothesis 4. *If women are more risk averse than men, then:*

a) we expect risk aversion to be negatively associated with appropriation in the treatment with a positive probability of punishment; and

b) we expect no systematic difference in appropriation between men and women in treatments with a probability of punishment once risk aversion and corruption aversion are held constant.

The rest of this section establishes the assumptions and logic from which these hypotheses are derived.

Payoff Function

Our model assumes an environment where each subject $j \in 1, 2, ..., n$ has control over an equal part of a common pool resource r_j such that $\sum_{j=1}^{n} r_j = n * r = R$. The subject can choose to appropriate a proportion $p_j \in [0, 1]$ of that resource for their personal gain; their individual takings are $t_j = p_j * r_j$. However, there is a chance $\pi(t_j)$ that a subject will be caught appropriating resources; if they are caught, then the subject receives no payoff (as a punishment).[2] The untaken proportion of the resource $(1 - p_j)$ creates externalities for all the subjects in the experiment at a rate m. Every subject enjoys an equal share of the untaken common pool resource times the externality rate m.

Thus, the expected payoff for an individual subject j given his/her choice to appropriate t_j public resources for private gain is is:

$$y(t_j) = \left(t_j + \left(R - \sum_{i=1}^{n} t_i \right) \frac{m}{n} \right) (1 - \pi(t_j))$$

We write the total utility that subject j derives from his/her decision as $u_j(t_j)$, where utility can be a function of payoffs $y(t_j)$ but also of other intrinsic or extrinsic incentives related to the individual's choice t_j.

No Punishment Condition

To begin, consider the case where $\pi(t_j) = 0$ (i.e., there is no exogenous punishment for corruption). In a one-shot, non-repeated interaction, every subject has the incentive to take the maximum share of the resource so long as:

$$\frac{\partial y(t_j)}{\partial t_j} = 1 - \frac{m}{n} > 0$$

[2] Note that the subject's payoff is destroyed, not returned to other subjects, if the subject is caught; in addition to simplifying the analysis, this feature represents administrative losses from detecting and prosecuting corruption.

Our experiment sets $n = 4$; therefore subjects have an incentive to take the maximum amount as long as $m < 4$. Thus, as long as $m < 4$, the unique Nash equilibrium for a one-shot interaction is for all subjects to choose $t_j^\star = r$ for all $j \in 1...n$. In a finitely repeated version of this game with a known endpoint, backward induction ensures that the unique subgame-perfect Nash equilibrium (SPNE) is for subjects to play the single-stage Nash equilibrium strategy in every period (i.e., for all subjects j to play $t_j = r$ in every period).

However, we may still observe that some subjects do not set $t_j = r$, even after becoming familiar with the game and its incentives. This may happen if subjects bring their own intrinsic aversion to corruption to the experiment. We presume that this aversion is a strictly convex[3] function of takings, $a_j(t_j)$, where $a(0) = 0$. Then each subject's utility (with $\pi(t_j) = 0$) is:

$$u_1(t_j) = \left(t_j - a_j(t_j) + \left(R - \sum_{i=1}^{n} t_i \right) \frac{m}{n} \right) \tag{1}$$

Utility maximizing subjects in the no-punishment condition will set $t_j^\star$ to solve:

$$\begin{aligned} \frac{\partial u_1(t_j)}{\partial t_j} &= 1 - \frac{\partial a_j(t_j)}{\partial t_j} - \frac{m}{n} = 0 \\ \frac{\partial a_j(t_j)}{\partial t_j} &= 1 - \frac{m}{n} \end{aligned} \tag{2}$$

noting that marginal utility will only be zeroed out in an interior solution $t_j^\star \in (0, r)$; corner solutions $t_j^\star \in \{0, r\}$ typically reflect subjects' latent desire to take less or more, respectively, than the limit.[4] That is, a subject chooses their level of takings $t_j^\star$ to balance their guilt with their reward. The quantity $(1 - m/n)$ measures the objective marginal gain for taking

[3]This condition ensures strict concavity of $u_1(t_j)$ and a unique utility maximizing choice $t_j^\star$.

[4]Subjects may choose $t_j = 0$ in the event that $\frac{\partial a_j(t_j)}{\partial t_j} > 1 - \frac{m}{n}$ at $t_j = 0$, i.e., when subjects value intrinsic corruption more highly at the margin than extrinsic gains. Similarly, subjects may still choose $t_j = r$ when $\frac{\partial a_j(t_j)}{\partial t_j} < 1 - \frac{m}{n}$ at $t_j = r$, i.e., when subjects are not sufficiently averse to corruption to counterweight their financial gain. We thank Rob Carroll for bringing our attention to these corner conditions.

each token. The subject appropriates resources until the marginal cost (in guilt) equals the marginal benefit, and then stops. Again, t_i for $i \neq j$ does not enter the decision, and so the Nash equilibrium is simply determined by each subject's personal aversion function.

Some prior research argues that women have a systematically greater aversion to corruption than men (Dollar, Fisman and Gatti, 2001). Our experiment allows us to separate greater intrinsic moral aversion (embodied in a_j) from the effects of differential punishment by gender, social desirability biases, and the like: the subjects know that our experiment implements punishment on a provably gender-neutral basis, and they do *not* know the identities of the people with whom they are interacting (as assignment to a group is random and subjects interact through computer terminals in private carrels).

Prior research also argues that gender differences in corruption behavior may not attributable to intrinsic moral aversion, but instead to greater aversion (on average) to *risk* among women (Esarey and Chirillo, 2013; Esarey and Schwindt-Bayer, 2017). However, greater risk aversion cannot explain gender difference in behavior in the no punishment condition. For a continuous utility function, risk aversion is synonymous with strict concavity:

$$u_2(\pi y_1 + (1-\pi)y_2) > \pi u_2(y_1) + (1-\pi)u_2(y_2) \qquad (3)$$

where u_2 is a risk-averse utility function and y_1 and y_2 are payoffs achieved with probabilities π and $(1-\pi)$ respectively; in our application, these payoffs are the values of equation 1 when punishment is applied or not. A subject with this utility function strictly prefers a certain outcome with payoff equivalent to the expected value of a lottery (the right hand side of equation 3) compared to the lottery itself (the left hand side). We further assume that the transformation u_2 is monotonic; that is, $y_1 > y_2 \rightarrow u_2(y_1) > u_2(y_2)$; this property implies that the preference ordering of non-random (certain) payoffs is identical regardless of risk aversion.

Under these conditions, the optimal choice $t_j^\star$ under the risk-neutral utility function $u_1(t_j)$ will also solve the maximization problem for a concave (risk-averse) transformation $g\left(u_1(t_j)\right)$, as monotonicity directly implies that $u_1(t_j^\star) > u_1(t_j') \; \forall \; t' \neq t^\star \rightarrow g(u_1(t_j^\star)) > g(u_1(t_j')) \; \forall \; t' \neq t^\star$.

We can see a more direct demonstration of this if we assume our subjects have constant relative risk aversion:

$$u_2(t_j) = \frac{u_1(t_j)^{(1-\eta)} - 1}{1-\eta} \tag{4}$$

where u_1 is defined in equation 1 and $\eta > 0$. The optimal choice for t_j solves:

$$\frac{\partial u_2}{\partial t_j} = \left[t_j - a_j(t_j) + \left(R - \sum_{i=1}^{N} t_i\right)\frac{m}{n}\right]^{-\eta}\left[1 - \frac{\partial a_j(t_j)}{\partial t_j} - \frac{m}{n}\right] = 0$$

where the second bracketed quantity is the same as the maximization problem in equation 2. If this quantity is zero, as it will be when $t_j = t_j^\star$ for a non-corner solution, then the derivative is zero and utility is maximized.[5]

Therefore, if women have a systematically greater intrinsic moral aversion to corruption than men that is activated by our experiment, we should be able to see this in systematically lower takings by women in the no-punishment condition. Moreover, any gender differences we see in this condition cannot be attributed to differential punishment, risk aversion, or social desirability biases.

With Punishment

When subjects who appropriate resources face a chance of losing all their payoffs, they naturally appropriate fewer resources. We can write a subject's utility with non-zero punishments

[5]The conditions for a corner solution are the same as those in footnote 4.

as:

$$u_3(t_j) = u_1(t_j)\,(1 - \pi(t_j))$$

We assume that $\pi(t_j)$ makes $u_3(t_j)$ strictly concave, so that any solution $t_j^\star$ is unique. This will be true when (writing $p(t_j) = 1 - \pi(t_j)$ for convenience):

$$\frac{\partial^2 u_3(t_j)}{\partial t_j^2} = \frac{\partial^2 u_1(t_j)}{\partial t_j^2} p(t_j) + 2\frac{\partial u_1(t_j)}{\partial t_j}\frac{\partial p(t_j)}{\partial t_j} + \frac{\partial^2 p(t_j)}{\partial t_j^2} u_1(t_j) < 0 \tag{5}$$

The first term of equation (5) is negative because of strict concavity of $u_1(t_j)$ as long as punishment $\pi(t_j) \in [0, 1)$.[6] The last term is always negative or zero as long as punishment is weakly increasing in t_j and utility is scaled to be positive in $t_j \in [0, r]$ (that is, corruption aversion never outweighs profit inside of the choice set). The middle term (specifically $\partial u_1(t_j)/\partial t_j$) can change sign depending on t_j: when $t_j^\star$ for maximizing u_1 is in $(0, r)$, then this term will be negative for $t_j > t_j^\star$ and thus the full middle term will be positive. The condition requires that punishment does not cause utility improvement (i.e., declines in utility from corruption aversion are mitigated by increasing probability of punishment) that could push someone to a local maximum at $t_j = r$. We could guarantee this by making disutility from corruption aversion not subject to the penalty:

$$u_{1a}(t_j) = \left(t_j + \left(R - \sum_{i=1}^{n} t_i\right)\frac{m}{n}\right)$$
$$u_{3a}(t_j) = u_{1a}(t_j)\,(1 - \pi(t_j)) - a_j(t_j)$$

[6] We assume that punishment is never certain to avoid an infinite term in the first order conditions for $t_j^\star$ in maximizing u_3; we further assume that $\pi(t_j = 0) = 0$ (i.e., subjects are not punished when not appropriating resources).

Then strict concavity of $u_{3a}(t_j)$ is guaranteed when:

$$\frac{\partial^2 u_{3a}(t_j)}{\partial t_j^2} = \frac{\partial^2 u_{1a}(t_j)}{\partial t_j^2} p(t_j) + 2\frac{\partial u_{1a}(t_j)}{\partial t_j}\frac{\partial p(t_j)}{\partial t_j} + \frac{\partial^2 p(t_j)}{\partial t_j^2} u_{1a}(t_j) - \frac{\partial a_j(t_j)}{\partial t_j} < 0 \qquad (6)$$

Linearity of $u_{1a}(t_j)$ means that the first term of equation (6) is zero and the second term is negative or zero for weakly increasing punishment. The third term remains negative or zero, and the new fourth term is positive (but subtracted) by assumed convexity of $a_j(t_j)$; therefore strict concavity is guaranteed.

Punishment decreases appropriation relative to an otherwise equivalent case without punishment. The utility maximization problem under punishment solves:

$$\frac{\partial u_3(t_j)}{\partial t_j} = \frac{\partial u_1(t_j)}{\partial t_j} - \pi(t_j)\frac{\partial u_1(t_j)}{\partial t_j} - \frac{\partial \pi(t_j)}{\partial t_j} u_1(t_j) = 0$$

for an interior solution $t_j^\star \in (0, r)$.[7] This can be rearranged:

$$\frac{\partial u_1(t_j)}{\partial t_j} = \frac{\partial \pi(t_j)/\partial t_j}{1 - \pi(t_j)} \times u_1(t_j) \qquad (7)$$

When the subject is choosing t_j without punishment, the left hand side of equation (7) is set equal to zero. For the choice including punishment, the left hand side is positive under the mild assumptions that (a) the probability of punishment increases in takings, i.e. $\partial \pi(t_j)/\partial t_j > 0$, and (b) subject utility is positive under the optimal choice $t_j^\star$, i.e.

[7] When $t_j^\star = 0$, maximization requires:

$$\frac{\partial u_1}{\partial t_j}(t_j = 0) - \frac{\partial \pi}{\partial t_j}(t_j = 0) \times u_1(t_j = 0) < 0$$

When $t_j^\star = r$, maximization requires:

$$\frac{\partial u_1}{\partial t_j}(t_j = r) - \pi(t_j = r)\frac{\partial u_1}{\partial t_j}(t_j = r) - \frac{\partial \pi}{\partial t_j}(t_j = r) \times u_1(t_j = r) > 0$$

We thank Rob Carroll for bringing our attention to these corner conditions.

$u_1(t_j^\star) > 0$. Given that $u_1(t_j)$ is strictly concave, the subject's preferred appropriation level $t_j^\star$ for a subject with punishment must be strictly less than that without punishment.

When punishment is positive, a subject appropriates fewer resources as their risk aversion rises. Moreover, their responsiveness to punishment will increase in their degree of risk aversion. We can write a risk-averse subject's utility function as a transformation of pre-existing functions:

$$u_4 = (1 - \pi(t_j))g(u_1(t_j))$$

Here, $g(x)$ is a strictly concave transformation of $u_1(t_j)$, such as the constant relative risk averse utility function defined in equation (4). Conditions similar to equation (5) or (6) will ensure the strict concavity of $u_4(t_j)$. A subject maximizing their utility chooses $t_j^\star$ such that the first-order condition is satisfied, which yields:

$$\frac{\partial u_1(t_j)}{\partial t_j} = \frac{\partial \pi(t_j)/\partial t_j}{1 - \pi(t_j)} \times \frac{g(u_1(t_j))}{\partial g(u_1(t_j))/\partial u_1} \tag{8}$$

when $t_j \in (0, r)$.[8] Strict concavity of $u_1(t)$ implies that positive increases in the right hand side must imply declining $t^\star$. Thus, increases in either the absolute or relative risk aversion

[8]When $t^\star = 0$, utility maximization entails:

$$\frac{\partial g}{\partial u_1}(u_1(t_j = 0)) \times \frac{\partial u_1}{\partial t_j}(t_j = 0) - g(u_1(t_j = 0)) \times \frac{\partial \pi}{\partial t_j}(t_j = 0) < 0$$

When $t^\star = r$, maximization entails:

$$(1 - \pi(t_j = r)) \times \frac{\partial g}{\partial u_1}(u_1(t_j = r)) \times \frac{\partial u_1}{\partial t_j}(t_j = r) - g(u_1(t_j = r)) \times \frac{\partial \pi}{\partial t_j}(t_j = r) > 0$$

We thank Rob Carroll for bringing our attention to these corner conditions.

coefficient imply declining $t^\star$:

$$\begin{aligned}\frac{g(u_1(t_j))}{\partial g/\partial u_1(t_j)} &= \frac{g(u_1(t_j))}{\partial g/\partial u_1(t_j)} \times \frac{\partial^2 g/\partial u(t_j)^2}{\partial^2 g/\partial u(t_j)^2} \\ &= \frac{\partial^2 g/\partial u(t_j)^2}{\partial g/\partial u_1(t_j)} \times \frac{g(u_1(t_j))}{\partial^2 g/\partial u(t_j)^2}\end{aligned}$$

The first term on the last line is the negative of the Arrow-Pratt coefficient of absolute risk aversion; when the numerator of the second term is included, it becomes the Arrow-Pratt coefficient of relative risk aversion. When either of these risk aversion coefficients rises (and noting that concavity of g implies $\partial^2 g/\partial u(t)^2 < 0$), the $\partial g/\partial u_1(t_j)$ term on the right hand side of equation (8) must rise and consequently $t^\star$ must fall. Moreover, the larger that this term is, the greater than an increase in punishment severity $(\partial\pi(t_j)/\partial t_j)$ will grow the right hand side of equation 8, and thus the more that $t^\star$ must fall.

Therefore, if women are (on average) more risk averse than men, we anticipate that the change in appropriation between the punishment and no punishment conditions will be greater for men that it is for women. We can also predict that women in the punishment condition will (on average) appropriate fewer resources than men. However, this might also be explained by women having greater aversion to corruption (particularly if we observe a gender difference in the no-punishment condition). Thus, we must control for corruption aversion when comparing appropriation between genders in the punishment condition when studying the effect of risk aversion on gender differences by directly comparing appropriation levels in the punishment condition.

Experimental Design

Our experiment implements the precise environment presumed by our theoretical model. Each of four subjects $j \in \{1, 2, 3, 4\}$ chooses to appropriate some proportion t_j of 100 tokens over which they have control; tokens not appropriated are multiplied by 1.33 and then redistributed equally to all participants. In one treatment condition, the probability of being punished for appropriation is always zero. In another treatment, the punishment probability is given by $0.4 \times (t_j/100)^4$; thus, increasing appropriation leads to a geometrically increasing probability of being caught. Any subject caught appropriating tokens forfeits all payoffs in that period of decision-making; these payoffs are destroyed (not returned to other subjects). This decision is repeated for ten periods; subjects know the specific number of repetitions to be made in advance of beginning the experiment. At the end of the ten periods, subjects are randomly rematched into new groups and the process is repeated for ten more periods. Each experiment involves three rounds of decision-making, each round of which contains ten periods; each experimental session contains twenty subjects who are randomly reassigned to different four-person groups in every round.

At the conclusion of all three rounds, subjects answer a short battery of survey questions. Three questions cover simple demographic factors (gender, race, and ideology).[9] One question asks whether the subject holds a leadership position in the student government; this allows us to assess whether those who select into leadership positions have greater or less moral corruption aversion than those who do not (and whether this moral aversion confounds any relationship between behavior and risk aversion). One question, modeled closely on Dohmen et al. (2011), is designed to measure the risk aversion of our subjects; Dohmen et al. (2011) provide evidence that this measure is valid and tracks well with lottery-based measures of risk aversion sometimes used in the experimental literature (Holt and Laury, 2002).[10] The last four questions on the survey ask subjects about their willingness to accept various forms of corruption on a ten point scale: claiming undeserved government benefits,

[9]The ideology question is based on a question from the American National Elections Survey (American National Election Studies, 2018).

[10]The survey coded this question with higher numbers indicating greater risk acceptance; we recoded this in the analysis so that higher numbers indicated greater risk aversion.

avoiding fares in public transportation, cheating on taxes, and accepting bribes.[11] These questions (measuring intrinsic corruption aversion) come from Wave 6 of the World Values Survey (World Values Survey, 2015). We combine subjects' scores on these questions to form an overall "corruption aversion" index used in some of our models.[12] The full survey is available as an online appendix.

Our subjects are undergraduates at Rice University who signed up to be members of the Behavioral Research Laboratory subject pool. Subjects sign up to participate in one of the sessions. A total of 112 subjects participated in each of our two treatment conditions, with 52 subjects in the no punishment treatment ($M_T = 0$) and 60 subjects in the punishment treatment ($M_T = 0.4$).[13] Participants in each treatment condition participate in a sessions with either 16 or 20 subjects; each treatment is applied in three sessions, and a total of six sessions of the experiment are run (for a total of 112 subjects who participate in the experiment). The experiments take place in a computer laboratory where subjects interacted through private computer terminals in individual carrels using zTree (Fischbacher, 2007); subjects do not know with whom (in the room) they are interacting during the experiment.[14] All subjects receive a show-up fee of $8, and additionally receive one dollar for every 200 experimental tokens (or 0.5 cents per token) they earn in the experiment. Subjects are read a brief instructional script at the beginning of the experiment (included as an Appendix) and were given a table relating the number of tokens taken to the probability of being caught (also included as an Appendix), and have an opportunity to try the software and ask questions

[11]The survey coded these questions with higher numbers indicating greater tolerance for corruption; we recoded them in the analysis so that higher numbers indicated greater corruption aversion.

[12]The index is the sum of the individual measures minus four, giving a theoretical range of scores between 0 and 36. Our sample actually ranges between 9 and 36 on this combined scale.

[13]An additional 16 subjects took part in a session before any of the other six sessions. This session experienced a software error at the beginning of the session, which caused the experimental software to switch from the punishment treatment to the no punishment treatment. A handout detailing the punishment treatment had already been distributed to the subjects, although we adjusted the instructions to correspond to the no punishment treatment. The rest of the session proceeded normally. Because of this glitch, we excluded these subjects from our analysis.

[14]Maggie Emerson assisted with the creation of this zTree program.

before the experiment begins.

Results

Figure 1 shows the average number of tokens per period that subjects appropriated, broken out by treatment condition and subject gender.[15] Two facts are immediately apparent from the figure. First, the punishment treatment lowered the number of tokens that subjects appropriated in the experiment. On average, subjects in the punishment treatment took just over 16 fewer tokens per period than those in the no-punishment treatment (out of 100 possible tokens they could have taken). Second, there is no immediately apparent effect of gender on taking rates, in either the punishment or the no-punishment treatment. Contrary to the SPNE prediction, most subjects did *not* appropriate all tokens in the no punishment treatment (though the median subject took an average of 70 tokens per period).

A closer analysis supports these two basic findings. Table 1 shows linear regressions predicting subjects' average token appropriation (Models 1 and 2) and a random-effects linear model predicting subjects' token appropriation in every period of the experiment (Model 3); Model 2 includes fixed effects for the experimental session. These analyses were preregistered (Esarey and Schwindt-Bayer, 2018).

For the models in Table 1, subject gender did not have a substantive nor statistically significant impact on the number of tokens that a subject took. This is illustrated in Figure 2, which calculates the model-predicted difference between tokens taken by females compared to males in both treatment conditions. Confidence intervals are corrected to ensure an overall 5% false discovery rate among marginal effects calculated in a model (Esarey and Sumner, 2018). In all three models, women take no fewer tokens than men in any treatment. This non-finding exists regardless of whether we control for aversion to corruption or risk (as in Models 2 and 3) or not (as in Model 1). Thus, the evidence is not consistent with the

[15] All figures and tables are created in R 3.5.0 (R Core Team, 2018).

Figure 1: Subject Average Number of Tokens Appropriated per Period, by Treatment and Gender

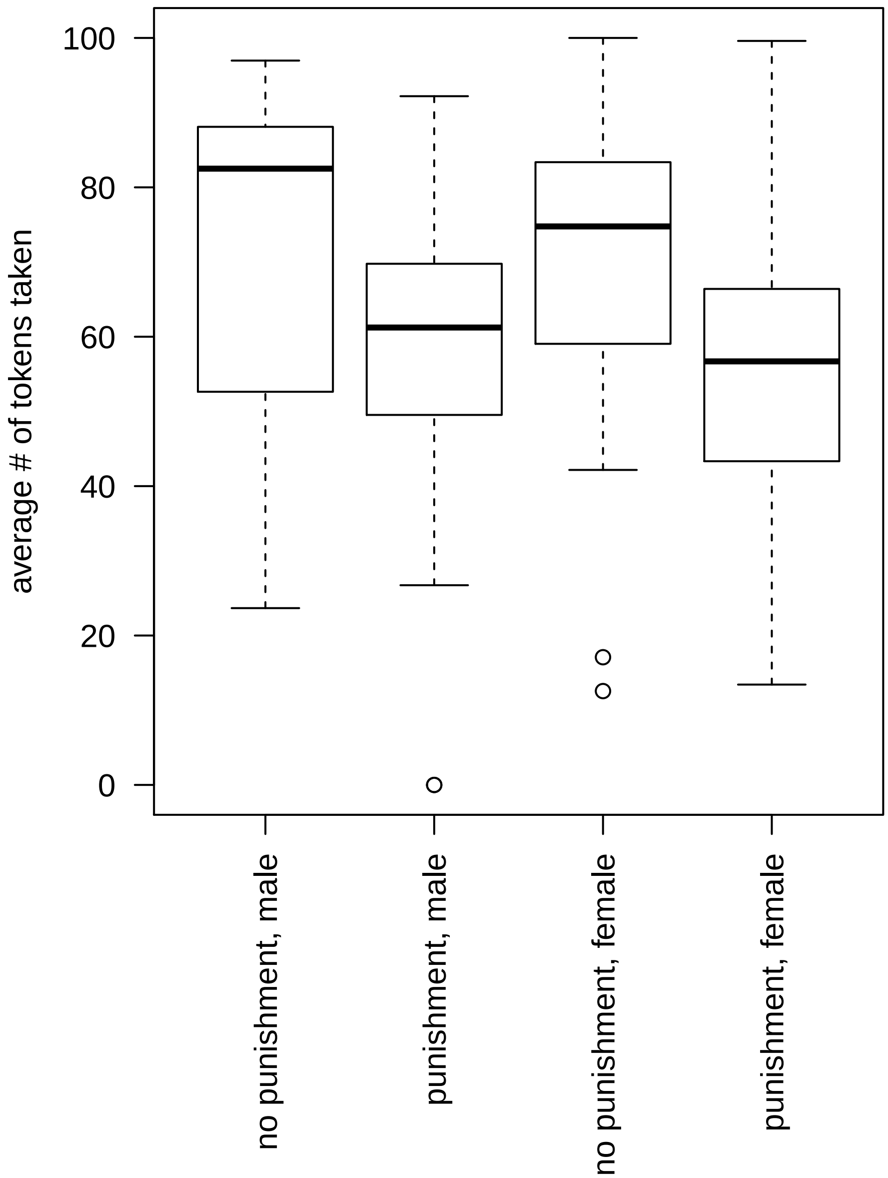


Table 1: Preregistered Analyses of Appropriation

	(1)	(2)	(3)
female	-2.453	-3.412	0.122
	(6.186)	(7.032)	(6.472)
punishment treatment	-14.766^{**}	-10.991	-15.376^{**}
	(6.391)	(9.025)	(6.648)
punishment X female	-0.100	3.744	1.065
	(8.274)	(9.063)	(8.606)
risk aversion		-1.903	-2.062^{*}
		(1.159)	(1.147)
corruption aversion		-0.163	-0.152
		(0.399)	(0.395)
leader		0.924	2.551
		(6.269)	(5.946)
intercept	72.206^{***}	86.179^{***}	64.684^{***}
	(5.075)	(12.402)	(11.886)
Averaged DV	Yes	Yes	No
Individual REs	No	No	Yes
Session REs	No	No	Yes
Session FEs	No	Yes	No
Period FEs	No	No	Yes
Observations	110	104	3,120

Note: *p<0.1; **p<0.05; ***p<0.01

Models (1) and (2): unstandardized coefficients for OLS regression of average individual token taking. Model (3): unstandardized coefficients for random-effects linear regression of individual token taking. Standard errors in parentheses.

prediction of hypothesis 1.

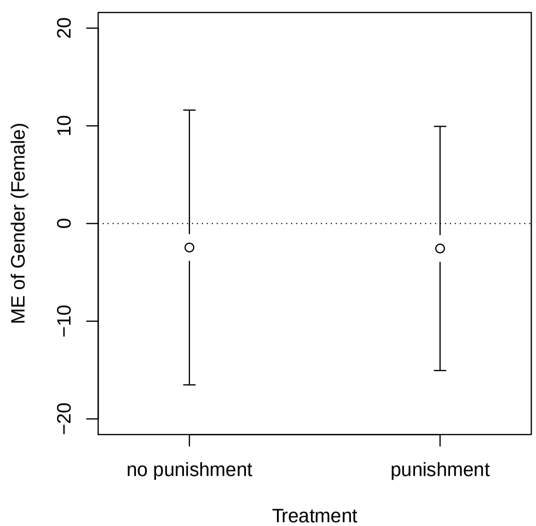
By contrast, the possibility of punishment for taking tokens *does* appear to have reduced the taking of tokens. The marginal effect of punishment in the models of Table 1 is assessed in Figure 3a. In all three models, the estimated effect of the punishment treatment is between -7 and -15 tokens, meaning that those in the punishment treatment take (on average) 7 and 15 fewer tokens than those in the no punishment treatment. In Models 1 and 3, this effect is statistically significant. In all cases, there is no substantively meaningful difference in the treatment effect between men and women; this contradicts the prediction of hypothesis 2.

Importantly, risk aversion *does* have a substantively meaningful and statistically significant ($\alpha = 0.1$, two-tailed) negative effect on token taking in Model 3 of Table 1; this evidence supports hypothesis 4a. In this model, every additional point of risk aversion (on a ten point self-rating scale) is associated with taking just over two fewer tokens. As our subjects all rated themselves between a 2 and a 9 on the scale (that is, neither endpoint was used), that means that the most risk-averse subjects are predicted to take (on average) 14 fewer tokens than the least risk-averse subjects. This effect of risk aversion is substantively comparable in magnitude to the effect of the punishment treatment. By contrast, our corruption aversion index has no statistically detectable relationship to the number of tokens taken; this is counter to the prediction of hypothesis 3a. Hypotheses 3b and 4b are supported by our results, but only trivially because there is never a systematic difference in appropriation between men and women.

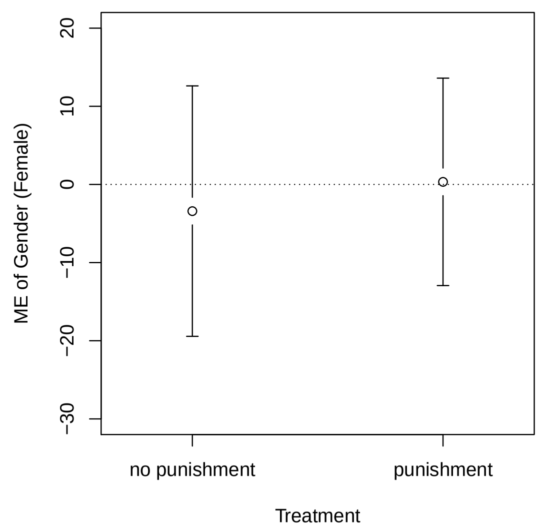
Discussion

The fact that risk aversion has an apparent influence on behavior, but gender does not, raises questions about the overall connection between gender, risk aversion, corruption aversion, and appropriation behavior in our experiment. To what extent is the theoretical basis for

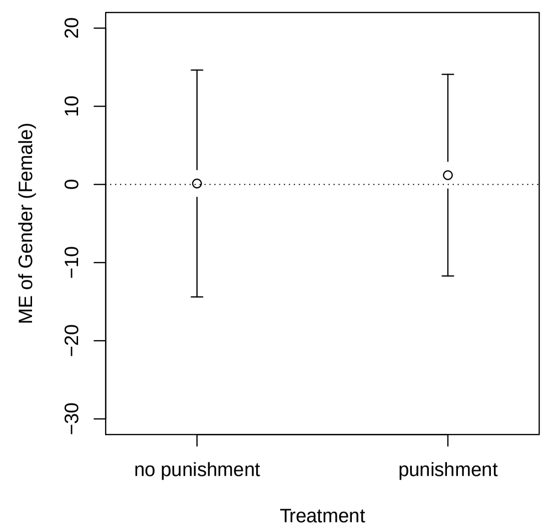
Figure 2: Marginal Effect of Female Gender on Token Taking based on Models in Table 1, with 95% Confidence Intervals



(a) Model 1 in Table 1



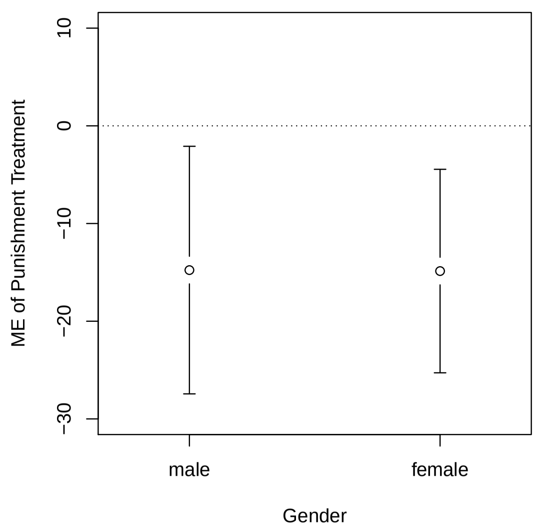
(b) Model 2 in Table 1



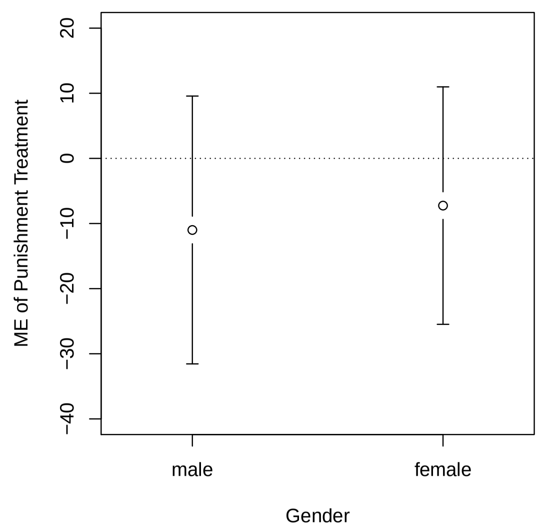
(c) Model 3 in Table 1

Confidence intervals are corrected to ensure an overall 5% false discovery rate among marginal effects calculated in a model (Esarey and Sumner, 2018).

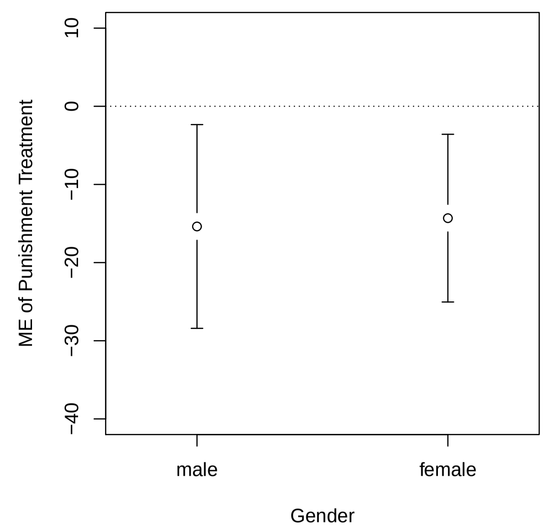
Figure 3: Marginal Effect of Punishment on Token Taking based on Models in Table 1, with 95% Confidence Intervals



(a) Model 1 in Table 1



(b) Model 2 in Table 1



(c) Model 3 in Table 1

Confidence intervals are corrected to ensure an overall 5% false discovery rate among marginal effects calculated in a model (Esarey and Sumner, 2018).

this experiment, and for prior arguments in this literature, contradicted by the present results? Using a structural equation model, we find that the predicted link from gender to risk aversion to appropriation *does* exist in the data, but is too attenuated to be visible as a direct effect of gender on behavior as predicted in hypotheses 1 and 2.

Our linear structural equation model has the following form:

$$
\begin{aligned}
\text{tokens taken} &= \beta_{10} + \beta_{11} * \text{punishment} + \beta_{12} * \text{risk aversion} + \\
&\qquad \beta_{13} * \text{female} + \beta_{14} * \text{corruption aversion} + \beta_{15} * \text{leader}^{\star} + \varepsilon_1 \\
\text{corruption aversion} &= \beta_{20} + \beta_{21} * \text{risk aversion} + \beta_{22} * \text{female} + \varepsilon_2 \\
\text{risk aversion} &= \beta_{30} + \beta_{31} * \text{female} + \varepsilon_3 \\
\text{leader}^{\star} &= \beta_{40} + \beta_{41} * \text{risk aversion} + \beta_{42} * \text{female} + \\
&\qquad \beta_{42} * \text{corruption aversion} + \varepsilon_4 \\
\text{leader} &= 1 \text{ if leader}^{\star} > \tau,\ 0 \text{ otherwise};\ \boldsymbol{\varepsilon} \sim \Phi(0, \Sigma)
\end{aligned}
$$

In this model, *leader* is a binary variable and therefore we model it as a function of a normally-distributed latent variable (leader*) where the binary outcome is determined by whether the latent variable exceeds some threshold τ. Σ is estimated using correlations among the variables (where the latent variable is used for binary observed variables), and β is fitted using generalized least squares with the diagonal elements of Σ informing the error weight matrix (Muthén, 1984; Muthén et al., 1997).

This model allows us to directly asses potential indirect pathways from gender through risk and corruption aversion to appropriation in our experiment. The results of this model are shown in Table 2, but a more visually understandable presentation of these results is given in Figure 4 using a path diagram of relationships (using standardized regression coefficients). Directional arrows indicate the effect of an independent variable (from which the arrow emanates) on a dependent variable (to which the arrow points). The width of the arrow is proportional to the size of the standardized regression coefficient (i.e., incorporating both the magnitude and certainty of the estimate).

Table 2: Structural Equation Model of Experimental Data

dependent var.	independent var.	coef.	se	p-value
tokens taken	punishment	-14.826	4.068	0.0003
tokens taken	risk aversion	-2.252	1.398	0.107
tokens taken	female	-0.437	4.743	0.927
tokens taken	corruption aversion	-0.097	0.412	0.814
tokens taken	leader	1.593	5.067	0.753
corruption aversion	risk aversion	0.849	0.284	0.003
corruption aversion	female	1.232	1.071	0.250
risk aversion	female	0.822	0.388	0.034
leader	corruption aversion	0.013	0.038	0.727
leader	female	0.064	0.338	0.849
leader	risk aversion	-0.134	0.085	0.115

χ^2 fit-statistic = 2.226, df = 2.817, p-value = 0.492
RMSEA = 0, RMSEA 5% p-value = 0.608

Unstandardized coefficients for structural equation model; parameters are for model of dependent variable listed in left column. Model fitted in `lavaan` (Rosseel, 2012) via WLSMV with pairwise deletion of missing data (N = 110 with 4 patterns of missing data). Robust χ^2 fit statistics reported.

What Figure 4 shows is that there *is* indeed an indirect pathway from gender through risk and corruption aversion to appropriation. Indeed, as Table 2 shows, women are on average 0.822 points higher on the risk aversion scale compared to men; given that our respondents used only values between 2 and 9 on this scale, this difference is equivalent to just under 12% of the variation on this scale. This difference is statistically significant ($\alpha = 0.05$, two-tailed). Women were also on average 1.2 points higher on the corruption aversion scale compared to men, though this difference isn't statistically significant at conventional levels. Risk aversion exerts a total effect (or "natural direct effect", as defined in VanderWeele 2015, pp. 56-60; see also VanderWeele 2016) of -2.548 fewer tokens taken per one point increase in risk aversion (standard error = 1.539); this effect is statistically significant at conventional levels (two-tailed p-value = 0.098).

Figure 4: Path Diagram of Relationships in Experimental Data

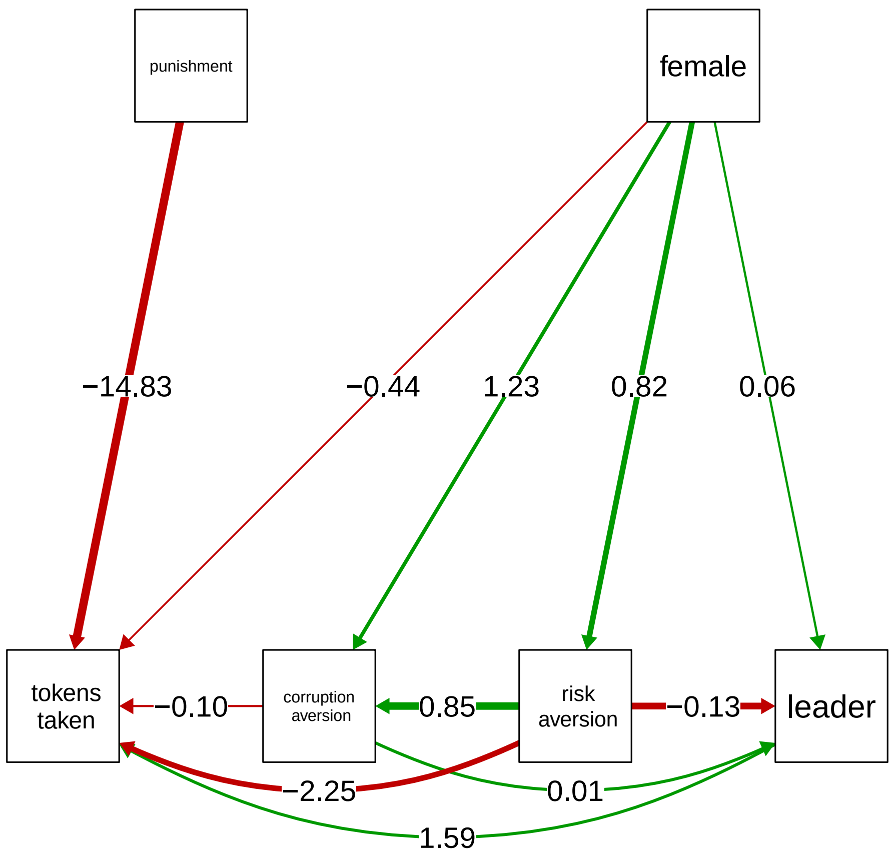


Diagram constructed using `semPlot` (Epskamp and Stuber, 2017). Edge labels indicate unstandardized coefficients from linear regression models in Table 2. Arrows point toward dependent variables and away from independent variables. Edges are sized proportional to the standardized coefficient (i.e., incorporating both magnitude and certainty of the estimate).

Thus, there *is* a link between gender and risk aversion in our experiment, and risk aversion *is* associated with reduced token-taking. There is also a (statistically insignificant) relationship between gender and corruption aversion, although the direct effect of corruption aversion on token-taking is small and statistically insignificant. These linkages are consistent with our theory. However, when the natural direct effect of gender on token taking is assessed, as in Model 1 in Table 1, there is no statistically detectable relationship.

Conclusion

At the present time, our findings do not provide a clear answer to whether gender differences in risk aversion or moral attitudes toward corruption lead people to engage in corruption less. Our estimate of the natural direct effect of gender on behavior suggests that they do not. No discernable gender differences in corruption behavior appear when the risk of engaging in corruption is triggered with detection and punishment (rejecting the risk aversion hypotheses) or when no detection and punishment exists (rejecting moral aversion hypotheses). However, supplemental modeling using direct measures of moral corruption aversion and risk aversion finds some indirect evidence of a link between gender, risk aversion, and corruption, suggesting that a relationship might exist but be undetectable in the experimental conditions we provided. In order to confirm this link, we need to perform more experimental research that changes some of the experimental conditions under which we operated.

Specifically, we think this experiment may not have presented sufficiently large risks, a sufficiently large perception of corruption to activate the indirect linkages strongly enough, and/or a large enough sample size for a natural direct effect of gender on appropriation to be substantively large and statistically meaningful. We are planning a new experiment that operate online instead of in a lab and in doing so allow us to generate a subject pool closer to 500 participants. The experiment will also present stronger punishments to more strongly trigger risk aversion if it exists in a subject. Finally, we will frame the taking of tokens in such a way that it is more clearly linked to corruption. We think these changes will address the possible weaknesses in the experiment just presented and help us accept/reject the theories of whether risk aversion and moral attitudes reduce women's engagement in corruption more firmly.

Of course, this experiment and the subsequent one do not directly analyze women (or men) in government. Getting access to women and men in government for an experimental study is much more difficult than getting access to average citizens. Although not ideal, we do look for any differences that may exist between student leaders and non-leaders and find little evidence that leadership confounds attitudes or behavior. This coupled with studies that have found few differences in risk-taking between the two populations (Donahue, Eckel and Wilson, 2014) increases our confidence that our experimental results would be valid for government officials too. Although our experimental results do not allow us to offer clear conclusions in this paper, we hope that with an additional experiment we can draw stronger conclusions about whether women's representation in government reduces corruption and why.

References

Albaum, Gerald and Robert A. Peterson. 2006. "Ethical attitudes of future business leaders: Do they vary by gender and religiosity?" *Business & Society* 45(3):300–321.

Alexander, Amy C. and Andreas Bågenholm. 2018. Does gender matter? Female politicians' engagement in anti-corruption efforts. In *Gender and Corruption*, ed. Helena Stensöta and Lena Wängnerud. Cham: Springer pp. 171–189.

American National Election Studies. 2018. "The ANES Guide to Public Opinion and Electoral Behavior." Online. URL: `http://www.electionstudies.org/nesguide/toptable/tab3_1.htm` accessed 3/19/2018.

Barnes, Tiffany D. and Emily Beaulieu. 2018. "Women Politicians, Institutions, and Perceptions of Corruption." *Comparative Political Studies* FirstView:0010414018774355. URL: `https://doi.org/10.1177/F0010414018774355`.

Bhargava, Vinay. 2005. "The Cancer of Corruption." Online. URL: `http://siteresources.worldbank.org/EXTABOUTUS/Resources/Corruption.pdf` accessed 4/9/2018.

Byrnes, James P., David C. Miller and William D. Schafer. 1999. "Gender differences in risk taking: a meta-analysis." *Psychological bulletin* 125(3):367.

Croson, Rachel and Uri Gneezy. 2009. "Gender differences in preferences." *Journal of Economic Literature* 47(2):448–74.

Dohmen, Thomas, David Huffman, Jürgen Schupp, Armin Falk, Uwe Sunde and Gert G. Wagner. 2011. "Individual Risk Atittudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association* 9(3):522–550.

Dollar, David, Raymond Fisman and Roberta Gatti. 2001. "Are women really the 'fairer' sex? Corruption and women in government." *Journal of Economic Behavior & Organization* 46(4):423–429.

Donahue, Amy K., Catherine C. Eckel and Rick K. Wilson. 2014. "Ready or not? How citizens and public officials perceive risk and preparedness." *The American Review of Public Administration* 44(4_suppl):89S–111S.

Eckel, Catherine C. and Philip J. Grossman. 2008*a*. Differences in the economic decisions of men and women: Experimental evidence. In *Handbook of experimental economics result*, ed. Vernon L. Smith Charles R. Plott. Elsevier chapter 57, pp. 509–519.

Eckel, Catherine C. and Philip J. Grossman. 2008*b*. "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices." *Journal of Economic Behavior & Organization* 68(1):1–17.

Epskamp, Sacha and Simon Stuber. 2017. *semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output*. R package version 1.1.
URL: *https://CRAN.R-project.org/package=semPlot*

Esarey, Justin and Gina Chirillo. 2013. "'Fairer Sex' or Purity Myth? Corruption, Gender, and Institutional Context." *Politics and Gender* 9(4):390–413.

Esarey, Justin and Jane Lawrence Sumner. 2018. "Marginal Effects in Interaction Models: Determining and Controlling the False Positive Rate." *Comparative Political Studies* 51(9):1144–1176. URL: `http://dx.doi.org/10.1177/0010414017730080`.
URL: *https://doi.org/10.1177/0010414017730080*

Esarey, Justin and Leslie A. Schwindt-Bayer. 2017. "Women's Representation, Accountability and Corruption in Democracies." *British Journal of Political Science* pp. 1–32.

Esarey, Justin and Leslie A. Schwindt-Bayer. 2018. "Preregistration Plan: Are Women More Resistant to Corruption? Evidence from a Laboratory Experiment." Open Science Foundation. URL: `http://dx.doi.org/10.17605/osf.io/k2gxd`.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10(2):171–178. URL: `https://doi.org/10.1007/s10683-006-9159-4`.

Glover, Saundra H., Minnette A. Bumpus, John E. Logan and James R. Ciesla. 1997. Re-examining the influence of individual values on ethical decision making. In *From the universities to the marketplace: the business ethics journey*, ed. Marilynn Fleckenstein, Mary Maury, Laura Pincus and Patrick Primeaux. Dordrecht: Springer Netherlands pp. 109–119.

Goetz, Anne Marie. 2007. "Political Cleaners: Women as the New Anti-Corruption Force?" *Development and Change* 38(1):87–105.

Gray, Cheryl W. and Daniel Kaufmann. 1998. "Corruption and Development." *Finance and Development* 35(1):7–10.

Holt, Charles A. and Susan K. Laury. 2002. "Risk aversion and incentive effects." *American Economic Review* 92(5):1644–1655.

Muthén, Bengt. 1984. "A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators." *Psychometrika* 49(1):115–132.

Muthén, du Toit Bengt, Stephen H.C. and Damir Spisic. 1997. "Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes." *Unpublished manuscript* . URL: `http://www.statmodel.com/bmuthen/articles/Article_075.pdf`.

Nelson, Julie A. 2015. "Are Women Really More Risk-Averse Than Men? A Re-Analysis of the Literature Using Expanded Methods." *Journal of Economic Surveys* 29(3):566–585.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: `https://www.R-project.org/`.

Reiss, Michelle C. and Kaushik Mitra. 1998. "The effects of individual difference factors on the acceptability of ethical and unethical workplace behaviors." *Journal of Business Ethics* 17(14):1581–1593.

Rosseel, Yves. 2012. "lavaan: An R Package for Structural Equation Modeling." *Journal of Statistical Software* 48(2):1–36. URL: `http://www.jstatsoft.org/v48/i02/`.

Roxas, Maria L. and Jane Y. Stoneback. 2004. "The Importance of gender Across Cultures in Ethical Decision-making." *Journal of Business Ethics* 50(2):149–165.

Ruegger, Durwood and Ernest W. King. 1992. "A study of the effect of age and gender upon student business ethics." *Journal of Business Ethics* 11(3):179–186.

Stensöta, Helena. 2018. *Final Thoughts: Taking Stock and Reflections on Ways Forward*. Cham: Springer International Publishing chapter 14, pp. 275–287.

Stensöta, Helena and Lena WÃd'ngnerud. 2018. *Gender and Corruption*. Cham: Springer International Publishing.

Stockemer, Daniel. 2011. "Women's Parliamentary Representation in Africa: The Impact of Democracy and Corruption on the Number of Female Deputies in National Parliaments." *Political Studies* 59(3):693–712.

Sundström, Aksel and Lena Wängnerud. 2016. "Corruption as an obstacle to women's political representation: Evidence from local councils in 18 European countries." *Party Politics* 22:354–369.

Swamy, Anand, Stephen Knack, Young Lee and Omar Azfar. 2001. "Gender and Corruption." *Journal of Development Economics* 64:25–55.

Tripp, Aili Mari. 2006. *Uganda: Agents of Change for Women's Advancement?* Boulder: Lynne Rienner chapter 5, pp. 111–132.

Valentine, Sean R. and Terri L. Rittenburg. 2007. "The Ethical Decision Making of Men and Women Executives in International Business Situations." *Journal of Business Ethics* 71(2):125–134.

VanderWeele, Tyler. 2015. *Explanation in causal inference: methods for mediation and interaction*. New York: Oxford University Press.

VanderWeele, Tyler J. 2016. "Mediation analysis: a practitioner's guide." *Annual review of public health* 37:17–32.

Wang, Liz C. and Lisa Calvano. 2015. "Is Business Ethics Education Effective? An Analysis of Gender, Personal Ethical Perspectives, and Moral Judgment." *Journal of Business Ethics* 126(4):591–602.

Watson, David and Amy Moreland. 2014. "Perceptions of Corruption and the Dynamics of Women's Representation." *Politics & Gender* 10(3):392–412.

World Values Survey. 2015. "Wave 6 2010-2014 Official Aggregate v.20150418." Online. URL: `http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp` accessed 3/10/2018. Aggregate File Producer: Asep/JDS, Madrid SPAIN.

Experimental Instructions

April 10, 2018

Introduction

Welcome, and thank you for participating in today's experiment. Please listen to the following instructions carefully.

In today's experiment, you will be asked to answer questions and make decisions at your computer terminal. No one participating in the experiment today will be able to match your decisions with your identity.

In addition to the $8 payment you receive for showing up today, you will receive payments based on the decisions that you and other participants make in the experiment. Payments will be allocated in tokens. At the end of the experiment, we will award you $1.00 for every 200 tokens that you have at the end of the experiment.

The experiment should last about one hour.

Please do not talk or communicate with other participants, or look at other participants' screens, during the experiment. If you have a question, you may raise your hand and an experiment supervisor will come over to help you. Not following this or other instructions will lead to your being excused from the experiment (without ill will from the experimenters) with your show-up fee only.

Today's Decision Task - Instructions

Today's experiment involves making decisions in a task. For this task, you will be randomly grouped with three other participants in the room. You will then engage in the decision task for 10 periods with these same persons.

Please look at your computer screen. A practice decision task is shown on your screen now. This is a practice stage only, and your responses will not affect your payoffs or be reported to other participants.

Your group of four participants has been given a group account of 400 tokens. Each person in the group can choose to take between 0 and 100 of these tokens for their personal gain. Any tokens that you take will be placed into your personal account.

Tokens that are not taken will be multiplied by 1.33, and then every person in the group will receive an equal share of the group account. For example, if every person takes 25 tokens, there will be 300 tokens left in the group account. These remaining tokens will be multiplied by 1.33, producing 399 tokens. Then every person in the group will receive an equal share of the group account, or 99.75 tokens; this will be rounded to the nearest whole token, so that in this example every person will receive 100 tokens from the group account.

Your final payoff in this period will be the number of tokens you took, plus your share of the group account. In this example, each person would receive 125 tokens: the 25 tokens that they took, plus their share of the group account.

[READ IN PUNISHMENT TREATMENT ONLY] There is a chance that you will be caught taking tokens. The chance that you are caught is shown on the sheet in front of you. For example, if you take zero tokens, you have a 0% chance of being caught. If you take 50 tokens, you have a 2.5% chance of being caught. If you take 100 tokens, you have a 40% chance of being caught. If you are caught taking tokens, you will forfeit all tokens in this period. If you are caught taking tokens, any tokens you took will be lost; they are *not* returned to the group account.

Are there any questions?

Once you understand everything on this screen, please type the number of tokens you wish to take into the box, then press the Continue button to continue with the experiment. Remember that this is a practice stage only, and your responses will not affect your payoffs or be reported to other participants.

Results

Please look at your computer screen. This screen will show you the results of your decision, including what payoffs you have earned. This is a practice round, and everyone is currently seeing the same set of practice results. Your practice decision on the previous screen is not shown on this practice results screen.

[READ IN PUNISHMENT TREATMENT ONLY] In this practice example, you were not caught taking tokens in this round. This outcome is shown at the top of the results display. Because you were not caught taking tokens in this practice example, the message is colored green. If you are caught taking tokens in a period, the message will be colored red. Below the message indicating that you were not caught taking tokens...

...there is a list of four items indicating the outcome of the last period. The first entry shows the size of the original group account, or 400 tokens. The next entry shows the total number of tokens taken by all participants; in this example, this number has been set to 38 tokens. The third entry shows the group account remaining after tokens have been taken; in this example, there are 362 tokens remaining. The fourth entry shows the tokens in the group account after being multiplied by 1.33; in this example, there are 481 tokens in the group account.

Finally, there is a list of three items indicating your payoffs from the last period. Your share of the group account is shown in the first entry; in this case, this is 120 tokens. The second entry shows that you chose to take 5 tokens. The final entry shows that you earned 125 tokens in this period.

Are there any questions?

Once you understand everything on this screen, please press the Continue button to continue with the experiment.

Begin Experiment

We will now begin the experiment.

Please look at your computer screen. The screen should now show that you have been placed into a group of four participants, with whom you will make decisions for the next ten periods.

Once you understand everything on this screen, press the Continue button to begin making decisions.

Continue Experiment

Please look at your computer screen. You have now finished 10 periods with the same participants. We will now reshuffle the group assignments. You will again be placed into a group of four participants. These participants may be different from the ones with whom you previously participated. You will make decisions with this new group of participants for the next 10 periods.

Once you understand everything on this screen, press the Continue button to begin making decisions with your new group.

Continue Experiment II

Please look at your computer screen. You have now finished 10 periods with the same participants. We will now reshuffle the group assignments. You will again be placed into a group of four participants. These participants may be different from the ones with whom you previously participated. You will make decisions with this new group of participants for the next 10 periods. These will be the last 10 periods of decision-making in the experiment.

Once you understand everything on this screen, press the Continue button to begin making decisions with your new group.

Ending Questionnaire

Thank you for participating in today's experiment! We are now ready to begin payment.

In a moment, you will be asked to enter your name and your Student ID number into the computer so that the supervisors can match your earnings to your name. Once you have pressed the OK button, the computer will also ask you a few survey questions. Your answers will not be matched with your name. Please answer these questions, using the scroll bar at the side of the screen to ensure that you have answered each question. Press the button marked "I'm finished" at the bottom of the survey form to complete the experiment.

After everyone has completed the survey, the supervisor will call you up to the front desk, one by one, to pay you your earnings and show-up fee. You will be asked to sign a receipt verifying that you received payment. Thanks again for participating!

Table: Relationship between Number of Tokens Taken and Probability of Being Caught

Tokens Taken	Probability of Being Caught	Tokens Taken	Probability of Being Caught
0	0.0000000%		
1	0.0000004%	51	2.7060804%
2	0.0000064%	52	2.9246464%
3	0.0000324%	53	3.1561924%
4	0.0001024%	54	3.4012224%
5	0.0002500%	55	3.6602500%
6	0.0005184%	56	3.9337984%
7	0.0009604%	57	4.2224004%
8	0.0016384%	58	4.5265984%
9	0.0026244%	59	4.8469444%
10	0.0040000%	60	5.1840000%
11	0.0058564%	61	5.5383364%
12	0.0082944%	62	5.9105344%
13	0.0114244%	63	6.3011844%
14	0.0153664%	64	6.7108864%
15	0.0202500%	65	7.1402500%
16	0.0262144%	66	7.5898944%
17	0.0334084%	67	8.0604484%
18	0.0419904%	68	8.5525504%
19	0.0521284%	69	9.0668484%
20	0.0640000%	70	9.6040000%
21	0.0777924%	71	10.1646724%
22	0.0937024%	72	10.7495424%
23	0.1119364%	73	11.3592964%
24	0.1327104%	74	11.9946304%
25	0.1562500%	75	12.6562500%
26	0.1827904%	76	13.3448704%
27	0.2125764%	77	14.0612164%
28	0.2458624%	78	14.8060224%
29	0.2829124%	79	15.5800324%
30	0.3240000%	80	16.3840000%
31	0.3694084%	81	17.2186884%
32	0.4194304%	82	18.0848704%
33	0.4743684%	83	18.9833284%
34	0.5345344%	84	19.9148544%
35	0.6002500%	85	20.8802500%
36	0.6718464%	86	21.8803264%
37	0.7496644%	87	22.9159044%
38	0.8340544%	88	23.9878144%
39	0.9253764%	89	25.0968964%
40	1.0240000%	90	26.2440000%
41	1.1303044%	91	27.4299844%
42	1.2446784%	92	28.6557184%
43	1.3675204%	93	29.9220804%
44	1.4992384%	94	31.2299584%
45	1.6402500%	95	32.5802500%
46	1.7909824%	96	33.9738624%
47	1.9518724%	97	35.4117124%
48	2.1233664%	98	36.8947264%
49	2.3059204%	99	38.4238404%
50	2.5000000%	100	40.0000000%

Appendix: Survey Questions

1. Which gender best describes you?

Male . . . . ☐
Female . . . . ☐
Other . . . . ☐

2. Which racial/ethnic category best describes you?

White, non-Hispanic . . . . ☐
African-American / Black . . . . ☐
Asian-American . . . . ☐
Hispanic / Latino . . . . ☐
Native American / American Indian . . . . ☐
Other . . . . ☐

3. When it comes to politics, which of the following best describes how you think of yourself?

extremely liberal . . . . ☐
liberal . . . . ☐
slightly liberal . . . . ☐
moderate . . . . ☐
slightly conservative . . . . ☐
conservative . . . . ☐
extremely conservative . . . . ☐

4. Do you have an elected position (president, representative, committee member, etc.) in the Student Association or in your residential college?

Yes, I have an elected position . . . . ☐
No, I do not have an elected position . . . . ☐

5. Are you generally a person who is willing to take risks, or do you try to avoid taking risks?

Please tick a box on the scale; the value 1 means "not at all willing to take risks" and the value 10 means "very willing to take risks."

not at all willing										**very willing**
	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
	1	2	3	4	5	6	7	8	9	10

6. For each of the following actions, please indicate whether you think it can always be justified, never be justified, or something in between.

Please tick a box on the scale; the value 1 means "never justifiable" and the value 10 means "always justifiable."

a) Claiming goverment benefits to which you are not entitled

never justifiable										always justifiable
	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
	1	2	3	4	5	6	7	8	9	10

b) Avoiding a fare on public transport

never justifiable										always justifiable
	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
	1	2	3	4	5	6	7	8	9	10

c) Cheating on your taxes if you have a chance

never justifiable										always justifiable
	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
	1	2	3	4	5	6	7	8	9	10

d) Someone accepting a bribe in the course of their duties

never justifiable										always justifiable
	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
	1	2	3	4	5	6	7	8	9	10