Gender and Voter Punishment of Corruption

Preliminary version: This paper is under active development. The results and conclusions reported herein may change as research continues.

Addison Collins, Justin Esarey,* and Betina Wilkinson

Wake Forest University
Department of Politics and International Affairs

April 9, 2021

Abstract

Why is corruption lower when more women are involved in governance? Using data from a conjoint survey experiment in Latin America conducted by Klašnja, Lupu and Tucker (2020), we examine two explanations. First, we study whether voters hold women incumbents to a higher ethical standard than men, providing an *extrinsic* motivation for female politicians to avoid corruption. Second, we test whether women have a greater *intrinsic* aversion to corruption that is reflected in voters' preferences for female candidates and whether female voters punish corruption more harshly. Our results indicate that all voters, regardless of gender, are more likely to vote against corrupt male politicians compared to equivalent corrupt women. We also find qualified evidence that female voters are less likely to support a corrupt candidate compared to male voters.

*Corresponding author: justin@justinesarey.com

Introduction

Extant research supports the conclusion that, at least in some contexts, greater participation by women in government causes reduced corruption in that government (Jha and Sarangi, 2018; Brollo and Troiano, 2016; Paweenawat, 2018; Bauhr, Charron and Wängnerud, 2018; Esarey and Schwindt-Bayer, 2019). But the reason *why* this causal relationship exists is far from clear. There are many plausible explanations, and most are not mutually exclusive with one another. One potential explanation is that female politicians are held to a higher moral standard than their male counterparts (Batista Pereira, 2020). Another is that women have more other-regarding values than men, and bring these values into office with them (Wängnerud, 2020).

In this study, we look for evidence of these two potential mechanisms linking gender to corruption in data from a conjoint survey experiment conducted by Klašnja, Lupu and Tucker (2020) in three Latin American countries. The experiment asked respondents to choose between two candidates whose characteristics varied on multiple dimensions; most importantly for our purposes, two of these characteristics were gender and alleged corruption. This research design allows us to pose and answer two questions. First, does corruption hurt voter support for female candidates more than it does for equivalent male candidates? If so, we could conclude that (at least in some instances) voters hold women to a higher ethical standard than men. If not, we might instead conclude that voters expect female candidates to be intrinsically more ethical even when suspected of corruption. Second, are female voters less likely than male voters to support candidates involved in corruption? If so, it would support the claim that women bring different values to bear when making political choices; the survey respondents are not professional politicians, but a value gap among voters would be consistent with such a value gap among their representatives.

We find that voters punish male candidates involved in corruption scandals more than equivalent women in the survey experiment. Specifically, the effect of alleged corruption on a woman candidate's vote share is substantially smaller than the effect for a man with

comparable characteristics. We also find qualified evidence that women voters are less likely than men to support corrupt incumbents. We infer that, at least in the Latin American countries we study, the causal linkage from gender to corruption flows (at least in part) through an intrinsic, value-based resistance to corruption. We find no evidence to corroborate Eggers, Vivyan and Wagner's (2018) finding that women voters punish female politicians for corruption more than male voters. We are also hesitant to conclude that internalized value differences are a universally suitable explanation for why more women in government seems to cause lower corruption, although our evidence shows that this mechanism is at work at least at some times and places.

Klašnja, Lupu and Tucker (2020) found that women candidates, on average, were less likely to be supported by respondents in this experiment when averaging over all treatment conditions. They also found that candidate gender does not moderate the effect of corruption on voter behavior. A recent study by Le Foulon and Reyes-Housholder (2021) of the same data argues that in general voters treat corrupt women candidates similar to men, although in Uruguay they receive slightly *more* support from voters than men in some circumstances. We disagree with both of these conclusions. First, Klašnja, Lupu and Tucker (2020) reversed the coding of candidate gender in their data; after correcting this oversight, we find that on average women candidates receive *more* voter support than equivalent men in this experiment.[1] Second, Le Foulon and Reyes-Housholder (2021) analyzes the data from each of the three countries studied in split samples with separate models, resulting in relatively low statistical power to detect treatment effects. AIC and BIC statistics for non-nested model comparison indicate that these models can be pooled with country fixed effects, increasing the power of the analyses. Moreover, the inter-country differences in candidate and respondent gender effects are substantively small in addition to being statistically undetectable; this implies that we can gain significant efficiency at the cost of relatively little bias by consolidating

[1]This miscoding was first discovered by Le Foulon and Reyes-Householder, who corrected the coding as part of their analysis. The mistake was confirmed by Joshua Tucker in e-mail correspondence among the authors of all the papers in question.

the analysis (Friedman et al., 2001, pp. 37-38). Similarly, Klašnja, Lupu and Tucker (2020, Appendix Table A2) detect no moderation effect of gender because of a similarly saturated (and therefore less efficient) model. Finally, Le Foulon and Reyes-Housholder (2021) follows Klašnja, Lupu and Tucker (2020) in analyzing the data monadically: a voter's preference for one candidate is assumed to not be a function of the other candidates characteristics. However, the experiment is intrinsically dyadic: voters choose between two candidates. Moreover, prior evidence indicates that corruption will only influence voter preferences when an anti-corruption candidate is available (Klašnja, Tucker and Deegan-Krause, 2016). Although we find a substantively meaningful and statistically significant effect of candidate gender on voter punishment of corruption in a monadic analysis, we also present a dyadic analysis of the data that provides additional evidence supporting the causal mechanism hypothesized to be at work.

Theory and Background

After an initial debate about spuriousness (Sung, 2003; Goetz, 2007), multiple studies have confirmed the conclusion of Dollar, Fisman and Gatti (2001) and Swamy et al. (2001): greater participation by women in government can reduce corruption in that government, at least in some contexts. For example, instrumental variables models (Correa Martínez and Jetter, 2016; Jha and Sarangi, 2018; Paweenawat, 2018; Esarey and Schwindt-Bayer, 2019) have shown that an increased proportion of women in public life[2] can cause reduced corruption at the country level. A regression discontinuity design among municipalities in Brazil reports the same finding (Brollo and Troiano, 2016). In India, random assignment of a female gender quota to local governments reduced bribery in those governments compared to those without a quota (Beaman et al., 2009). There are a large number of theoretical explanations for this relationship (Wängnerud, 2014; Wängnerud, 2020), all of which which

[2]The precise measure is either of the proportion of women in parliament (Jha and Sarangi, 2018; Paweenawat, 2018; Esarey and Schwindt-Bayer, 2019) or participation by women in the labor force (Correa Martínez and Jetter, 2016).

are difficult to definitively falsify given the contextual sensitivity of the gender-corruption relationship (Esarey and Chirillo, 2013; Esarey and Schwindt-Bayer, 2018).

Consider the possibility that voters punish female politicians more harshly for corruption than their male counterparts. If true, it would explain (at least in part) why greater involvement of women in government reduces corruption but only when politicians are highly accountable to voters (Schwindt-Bayer and Tavits, 2016; Esarey and Schwindt-Bayer, 2018). Moreover, there is empirical evidence to support this proposition. Survey experiments by Batista Pereira (2020) discover that voters in Mexico are less supportive of women politicians accused of corruption compared to men. Eggers, Vivyan and Wagner (2018) conducts similar experiments in the UK, finding that voters in general do not punish women more harshly than men for corruption but *women* voters are less likely to vote for women involved in corruption compared to equivalent male politicians. These findings are possibly attributable to voters having greater moral expectations for women. Barnes and Beaulieu (2014) finds that US survey respondents expect female politicians to be less corrupt than equivalent men; Schwindt-Bayer, Esarey and Schumacher (2018) finds the same expectation in both the US and Brazil. Barnes and Beaulieu (2018) and Barnes, Beaulieu and Saxton (2018) confirms these findings and is able to attribute them to voters' belief that women are more risk-averse.

But there is also a substantial body of empirical evidence that voters treat men and women involved with corruption the same way. Survey experiments described in Schwindt-Bayer, Esarey and Schumacher (2018) find no evidence for differential punishment of corruption by voters in the United States. The aforementioned study by Batista Pereira (2020) also finds no such evidence among voters in Brazil. Most recently, Le Foulon and Reyes-Housholder (2021) studies data from a survey experiment originally conducted by Klašnja, Lupu and Tucker (2020) and finds that Uruguayans actually *prefer* allegedly corrupt women politicians to equivalent men while voters in Argentina and Chile do not treat corrupt politicians differently by gender.

Thus, the overall picture that emerges from previous work is quite confusing. There are

sound reasons to believe that voters expect women to be less corrupt than men in a variety of contexts. We even find some evidence for harsher punishment of corrupt women by voters in some studies. But in others, we find either mixed evidence or positive refutation of no such differential punishment. It is unclear *why* we find such inconsistent evidence; there are many possible explanations. It may be that the effect only exists in some institutional contexts (Esarey and Chirillo, 2013). Perhaps the effect is relatively small, and thus only detectable in a highly powered study with a large sample. Conversely, perhaps the positive results are anomalous and an artifact of publication bias (Scargle, 2000).

There is a similarly confusing pattern of support for the thesis that women politicians bring distinct values or ethical commitments to public office that displace corruption. Dollar, Fisman and Gatti (2001) and Swamy et al. (2001) offered this as an explanation for their early findings. Later studies by Esarey and Chirillo (2013), Barnes and Beaulieu (2018), and many others claimed to rule out this explanation by finding that the gender-corruption relationship is sensitive to context. But Wängnerud (2020) re-advances the value-based argument based on evidence from Wängnerud, Solevid and Djerf-Pierre (2019) and a re-interpretation of prior evidence. Le Foulon and Reyes-Housholder (2021), for example, argues for this explanation.

This paper examines the possibility that differential treatment of corrupt women by voters may be a relatively small effect that is easily missed in a low-powered statistical analysis, but large enough to be politically meaningful. If voter support for corrupt female politicians is 1-2 percentage points different compared to that for corrupt men, the effect could be difficult to detect in a small sample but large enough to decide an election outcome. Differential punishment of this magnitude would not be enough to explain the entire connection between gender and corruption observed, but a set of small effects like it could cumulate to this larger aggregate relationship. Similarly, if women politicians bring slightly different values to public office, this could explain part of the gender-corruption connection even if it cannot explain it all.

Our survey of the literature leads us to look for two possible relationships in a survey

conjoint experiment of voters in Latin America:

Hypothesis 1. *Extrinsic incentive differences: respondents are less likely to vote for female candidates implicated in corruption compared to equivalent male counterparts.*

Hypothesis 2. *Intrinsic value differences: Female respondents are less likely to vote for candidates implicated in corruption compared to male respondents.*

Hypothesis 1 examines whether voters punish women politicians more severely for corruption than men. If true, this would suggest an *extrinsic* motivation for female politicians to avoid corruption. Hypothesis 2 looks for evidence that women bring different value priorities to governance than men: if this is true, we should see some evidence that women voters are more opposed to corruption than men and therefore less likely to support politicians implicated in corruption. This is an *intrinsic* motivation for women to resist corruption.

There are also two hypotheses that come directly from previous empirical work:

Hypothesis 3. *Female respondents are less likely to vote for women candidates implicated in corruption compared to male respondents.*

Hypothesis 4. *Respondents are more likely to vote for female candidates implicated in corruption compared to equivalent male counterparts.*

Hypothesis 3 revisits the finding of Eggers, Vivyan and Wagner (2018) that women voters more harshly punish women politicians, implying that there are gender-specific double standards for corrupt politicians. This idea is consistent with an extrinsic incentive for women politicians to avoid corruption, but conditioned by different preferences of male and female voters. Hypothesis 4 comes directly from Le Foulon and Reyes-Housholder (2021, p. 3), based on the "essentialist and structuralist reasoning" that women are more resistant to corruption than men and thus preferable to voters.

Data

Our study uses data collected by Klašnja, Lupu and Tucker (2020). They conducted nationally representative surveys in Argentina, Chile, and Uruguay. Survey respondents were presented with a vignette concerning two mayoral candidates. The characteristics of those candidates were randomly varied by subject; this procedure allows calculation of an Average Marginal Causal Effect (AMCE) using simple regression analysis with each treatment as a binary independent variable (Hainmueller, Hopkins and Yamamoto, 2014).

	Mean	SD	Sum
Vote for Candidate	0.288	0.453	2499
Candidate: Female	0.510	0.500	4757
Respondent: Female	0.511	0.500	4772
Argentina	0.327	0.469	3056
Chile	0.348	0.476	3250
Uruguay	0.324	0.468	3028
Treatment: Fought bribery	0.247	0.431	2303
Treatment: Bribes	0.259	0.438	2418
Treatment: Bribes common	0.247	0.431	2303
Treatment: Bribes but jobs	0.247	0.431	2304
N	9334		

Table 1: Summary statistics from the survey experiment.

Most pertinent to this study, the gender of each candidate and their involvement with corruption was randomly assigned. Specifically, each mayoral candidate was described as someone who (a) fought bribes, (b) accepted bribes, (c) accepted bribes when doing so was a common and widespread practice, and (d) accepted bribes that brought construction jobs to the community. The subjects were then asked "if you had to choose between these two candidates, for whom would you vote?"

Summary statistics for the study are shown in Table 1. As indicated in the table, roughly 1/3 of the sample comes from Argentina, Chile, and Uruguay and roughly 1/4 of the sample is randomly assigned to each of the four bribery treatment conditions. Just over half of the candidate descriptions were of a female candidate, and just over half of the respondents were

female. As discussed by Klašnja, Lupu and Tucker (2020), subjects' support for incumbents is relatively low in this experiment because so many of them are described as being involved in corruption.

Analysis and Results

We begin with a monadic OLS regression of subject vote choice. In this type of analysis, the same approach taken by Klašnja, Lupu and Tucker (2020), each subject has two entries in the data set corresponding to the candidates they see; therefore standard errors are clustered by subject. Appendix Table 2 shows fixed-effect models (by country) of binary vote choice with a variable for candidate gender and interactions between treatment condition and candidate gender. We estimate models for the full sample as well as separately for female and male survey respondent subpopulations. The AIC and BIC selection criteria favor a model with a single interaction between candidate gender and any treatment involving bribery (as opposed to separate interaction terms for each of the three bribery treatment variants), so we present the findings from these simplified models in Figure 1. Except where noted, all such figures use a Bonferroni correction for multiple comparisons based on the number of effects displayed in the figure.

Figure 1 shows no evidence for gender discrimination against ordinary politicians: there is a near-zero point estimate for the effect of candidate gender in the non-bribery treatment. However, voters are about 3 percentage points more likely to support a female candidate implicated in bribery compared to an equivalent male candidate. This effect is small, but statistically significant, in the full sample ($p < 0.05$, two-tailed with Bonferroni correction). Moreover, a 3 percentage point difference in support can be very substantively significant in a close election: it can be the difference between victory and defeat. The estimated effects in the subsamples of female and male respondents are similar in sign and magnitude, albeit not as statistically certain. We therefore conclude that there is no support for the differential

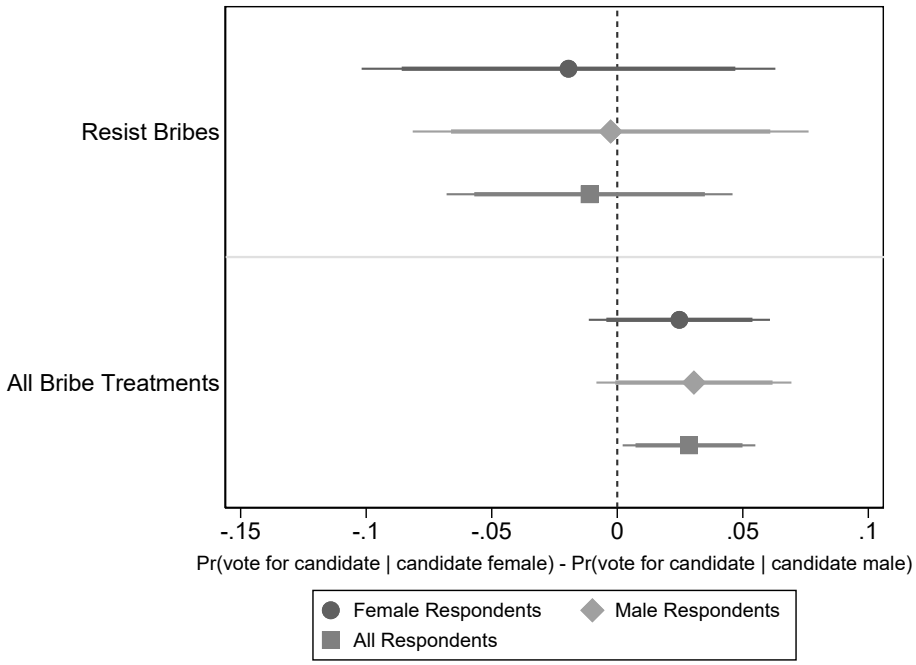


Figure 1: Effect of candidate gender on voting based on Models 2, 4, and 6 from Appendix Table 2. Lines are Bonferroni-corrected 95% and 80% confidence intervals based on standard errors clustered by subject.

punishment hypothesis (H1) in this data set. There is, however, support for Hypothesis 4: voters seem to be more willing to vote for female candidates accused of corruption compared to equivalent men.

Our results for the effect of respondent gender, shown in Figure 2, are not as clear-cut. As before, the AIC and BIC favor a model that consolidates all bribery treatments; we therefore present results from these simplified models. On one hand, female voters seem to punish candidates for involvement for corruption more than male voters; specifically, women are about 3 percentage points less likely to support a candidate accused of bribery compared to male voters ($p < 0.05$, two-tailed with Bonferroni correction). On the other hand, although there is no statistically detectable effect of respondent gender on candidate support in the "resist bribes" treatment, the estimated magnitudes are very close to the estimated magnitudes in the bribe treatments. We therefore have reason to doubt whether there is a real difference in respondent behavior when comparing the treatment and control conditions.

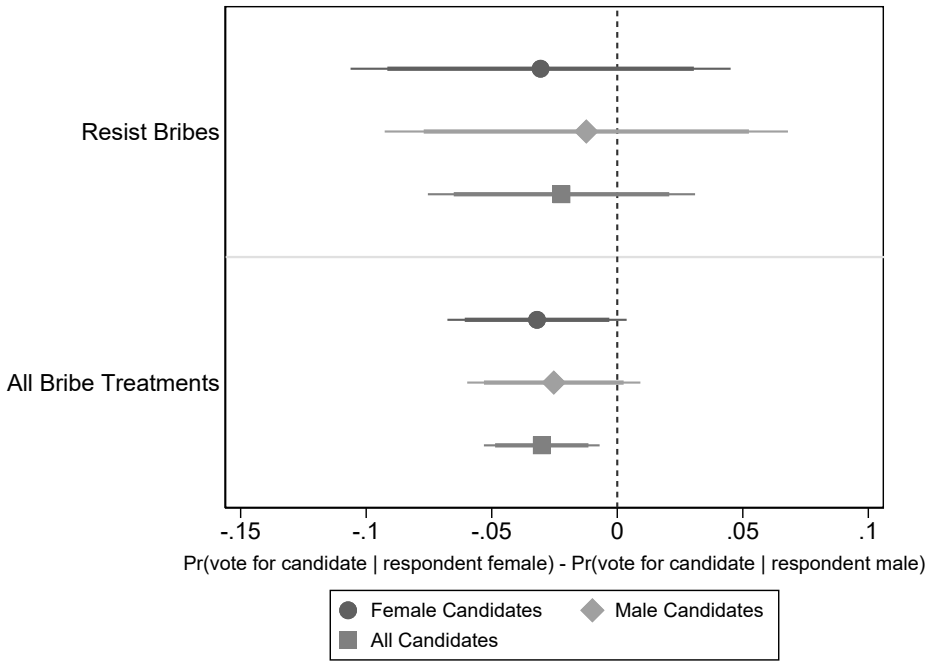


Figure 2: Effect of respondent gender on voting based on Models 2, 4, and 6 in Appendix Table 3. Lines are Bonferroni-corrected 95% and 80% confidence intervals based on standard errors clustered by subject.

Thus, we believe that there is qualified support for Hypothesis 2 in the data set but interpret this finding cautiously. Finally, when comparing the female and male respondents, we see no substantively meaningful difference in support for women candidates implicated in bribery as compared to male candidates; consequently we firmly reject Hypothesis 3.

Dyadic Analysis

While Klašnja, Lupu and Tucker (2020) used a monadic analysis, there is good reason to believe that subjects' vote choice depends not just on each candidate's characteristics but how they compare to one another. We therefore reshaped the data set so that each respondent corresponded to one observation, vote choice indicated whether they would vote for the incumbent (= 1) or challenger (= 0), and both incumbent and challenger characteristics were used predictors of this choice.

Figure 3 shows the predicted difference in support for female and male incumbents in each

possible combination of treatment conditions based on an OLS model shown in Appendix Table 4. Counter to Hypothesis 1, women incumbents receive around 5 percentage points more support from respondents compared to equivalent men when they are alleged to accept bribes. However, this effect may depend on the gender of the challenger as well as the incumbent.

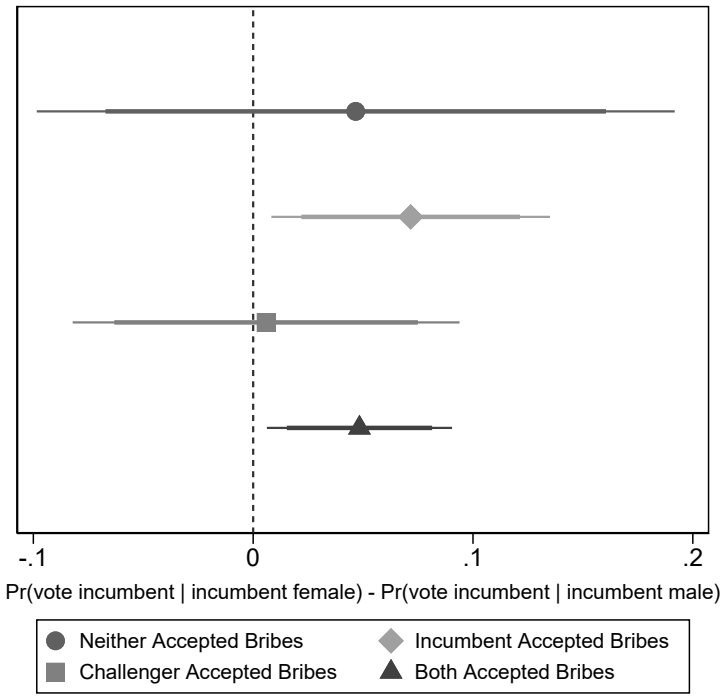


Figure 3: Effect of incumbent gender on voting under all treatment combinations, based on Model 2 in Appendix Table 4. Lines are Bonferroni-corrected 95% and 80% confidence intervals.

Figure 4 shows the difference in predicted support for female and male incumbents accused of accepting bribes depending on whether the challenger accepted bribes as well as the challenger's gender; these predictions are based on OLS models in Appendix Table 5. The biggest difference in support for corrupt women politicians (relative to male counterparts) is in the condition where a male challenger fought bribes; this is consistent with the supposition of Hypothesis 4 that allegations of corruption are less plausible for women (and allegations of corruption resistance are less plausible for men). The magnitude of this effect is large,

with corrupt incumbent women receiving well over ten percentage points more support than corrupt incumbent men. There is a smaller, marginally statistically significant advantage for corrupt women when a male challenger is also accused of corruption; this could reflect voters' greater skepticism of corruption allegations applied to women compared to men. Finally, there is no statistically detectable difference in support for corrupt women politicians compared to men when the challenger is female.

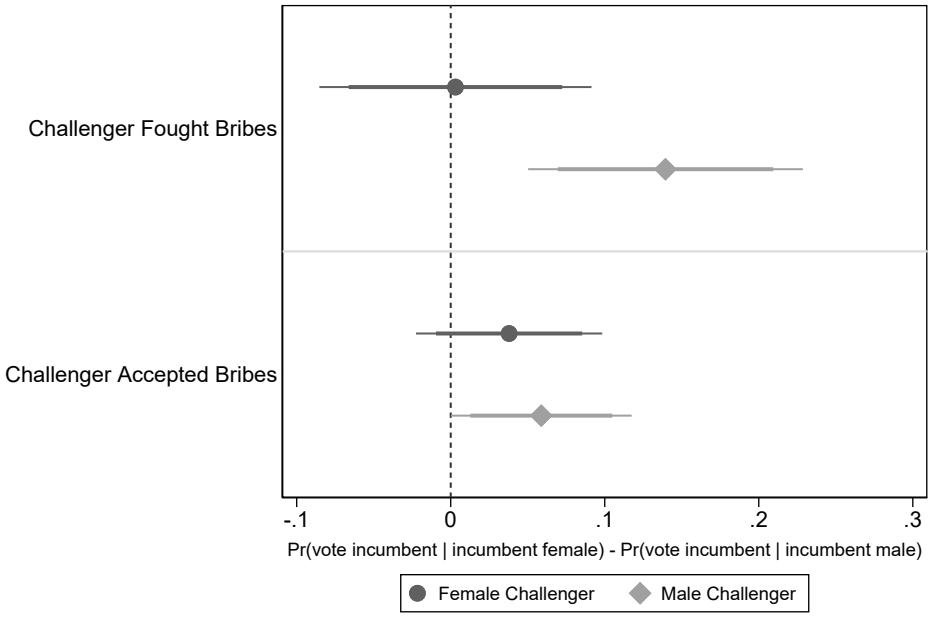


Figure 4: Effect of incumbent gender on voting when the incumbent is bribed, based on Appendix Table 5. Lines are Bonferroni-corrected 95% and 80% confidence intervals.

As in the monadic analysis, there is qualified support for Hypothesis 2 in the data. Figure 5 shows the difference in Pr(vote for incumbent) comparing female to male respondents. In Figure 5b, this quantity is estimated separately for each of the possible combinations of treatment conditions (consolidating all bribery treatments) using the estimates from OLS model 2 in Appendix Table 6. The figure shows that women respondents are slightly less likely than men to vote for incumbents accused of bribery; however, none of these differences is statistically significant at conventional levels. If we further consolidate the analysis to compare all treatments where incumbents fought bribes to all treatments where they didn't,

as in Figure 5a, we find that women respondents are about 3.8 percentage points less likely than men to support an incumbent accused of bribery, whilst there is no difference in support for incumbents who fought bribes. However, the AIC and BIC model selection criteria prefer the more complex model of Figure 5b. Consequently, support for Hypothesis 2 should be interpreted cautiously.

Finally, and as in the monadic analysis, we find no support for Hypothesis 3 in the dyadic data set. Figure 6 shows the estimated disproportionate propensity to vote for women incumbents over men under the four treatment conditions in the survey experiment, estimated separately for male and female respondents, based on Models 1 and 2 in Appendix Table 7. It is immediately apparent that there is no substantively meaningful difference in the treatment of female and male candidates by respondent gender. These findings are contrary to Hypothesis 3.

Comparison with past analysis

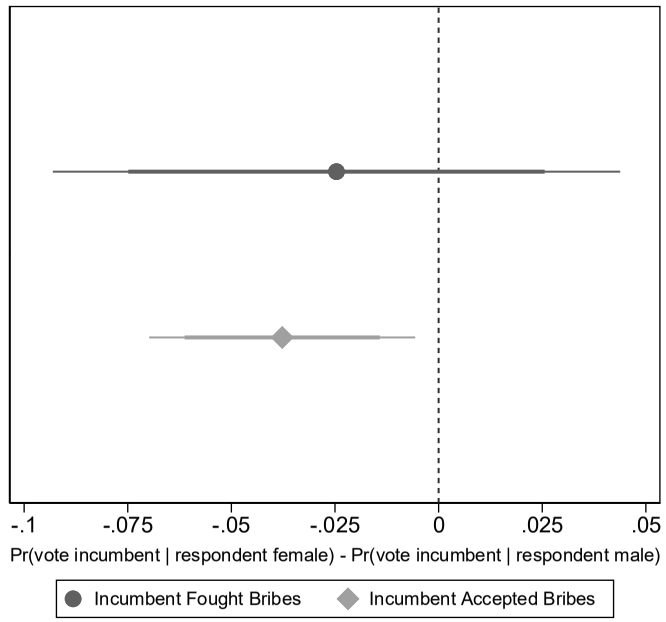
In their analysis of the same data set, both Klašnja, Lupu and Tucker (2020) and Le Foulon and Reyes-Housholder (2021) come to conclusions much different than ours. Klašnja, Lupu and Tucker (2020, p. 9) says:

> We do not find heterogeneous effects [of corruption on vote choice] based on the source of the corruption allegation, or the gender or party affiliation of the candidate (see Supplementary Figure A5, Supplementary Table A2, and Supplementary Figure A10).

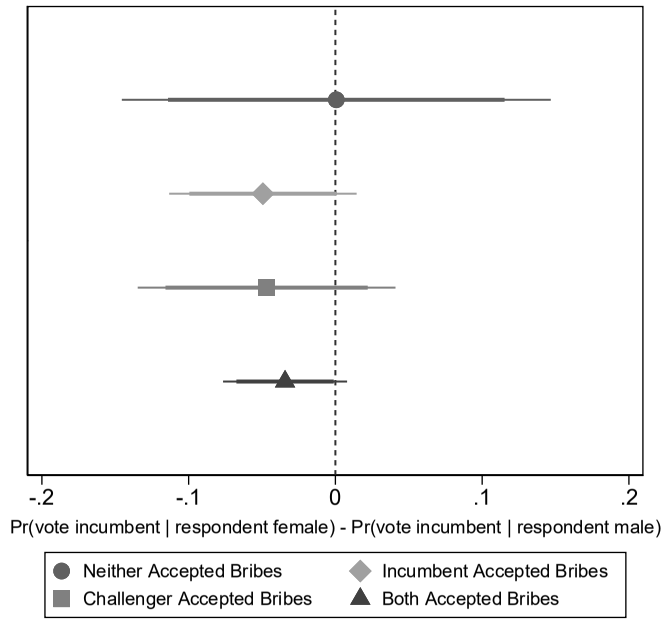
Furthermore, Figure 1 in Klašnja, Lupu and Tucker (2020, p. 6) shows slightly lower support for female candidates compared to men when averaging over all treatment conditions.

Le Foulon and Reyes-Housholder (2021, p. 1) argue that:

> ...voter belief can yield advantages to hypothetical female candidates as voters sometimes punish them less for bribe accusations. ...We find that only Uruguay-



(a) Model 1



(b) Model 2

Figure 5: Effect of respondent gender on voting based on models in Appendix Table 6. Lines indicate Bonferroni-corrected 95% and 80% confidence intervals.

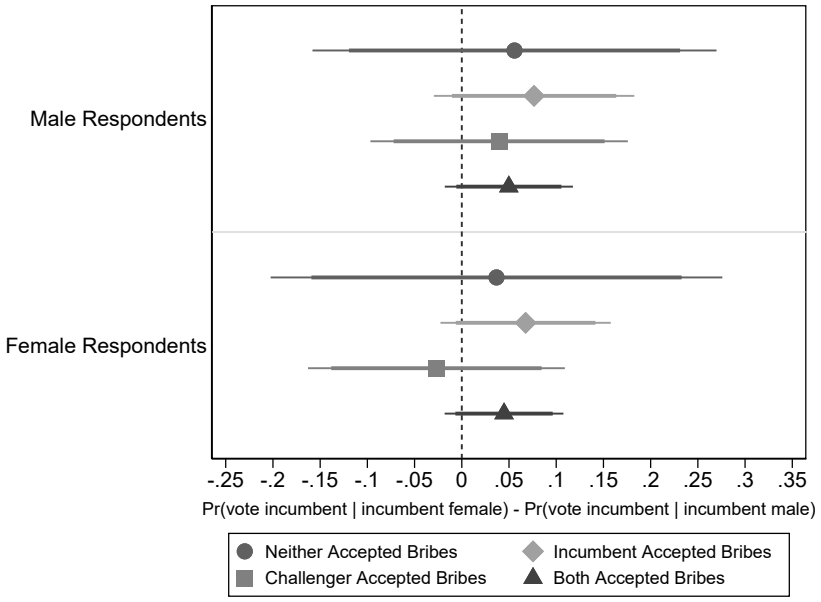


Figure 6: Effect of incumbent and respondent gender on voting, based on Appendix Table 7 models (1) and (2). Lines are Bonferroni-corrected 95% and 80% confidence intervals.

> ans prefer allegedly corrupt female over male candidates, but when Uruguayans are told that hypothetical corruption is widespread, they do not prefer accused females. Moreover, voters in none of the countries prefer females among candidates who have fought corruption.

There are three reasons that we draw dramatically different inferences from the same data; we believe that these reasons tend to support our findings.

Reason 1: Klašnja, Lupu and Tucker (2020) miscodes the variable identifying the gender of the candidate. This explains why they find that male candidates receive, on average, more support from voters than women. We know that our coding of the candidate gender is correct due to correspondence with the study's authors; Le Foulon and Reyes-Housholder (2021) used the correct coding in their analysis.

Reason 2: The country-specific models estimated by Le Foulon and Reyes-Housholder (2021) and saturated interaction models estimated by Klašnja, Lupu

and Tucker (2020) are inefficient. Estimating separate models by country, as Le Foulon and Reyes-Housholder (2021) does, cuts the sample into thirds and drastically decreases the power of the analysis to detect relationships. A similar problem occurs in Klašnja, Lupu and Tucker (2020) Appendix Table A2, where they interact all treatment conditions and country fixed effects with gender. However, the AIC and BIC model selection criteria support pooling the analysis by country (while including country fixed effects). In Appendix Table 8, we separate the samples by treatment condition, estimate a model with full interaction between candidate gender and country in each subsample, estimate a second model with country fixed effects but setting the effect of candidate gender equal in all three countries, then compare the AIC and BIC for these models. In seven out of eight comparisons, the selection criterion prefers the pooled model. When combined with the similar magnitude of country-specific estimates from in Figure 7, a replication of Figure 2 in Le Foulon and Reyes-Housholder (2021), this is strong evidence that a more efficient pooled analysis should be preferred. For models by respondent gender, shown in Appendix Table 9, the BIC always prefers pooling the analysis by country and the AIC prefers pooling in two out of four models. Finally, our analysis in Appendix Table 2 and 3 supports consolidation of all interactions of gender with the three bribery treatments into a single interaction. It may be the case that there are small differences in response to the treatments among countries or differences in the strength of the gender moderator among treatments that are too small to detect in this sample. However, because these differences are so small, it is preferable to pool the sample and obtain a large reduction in variance in exchange for a small amount of bias (Friedman et al., 2001, pp. 37-38).

Reason 3: The results of a dyadic analysis support the theoretical mechanisms underlying Hypotheses 2 and 4. Both Le Foulon and Reyes-Housholder (2021) and Klašnja, Lupu and Tucker (2020) use the monadic form of the data set in their analyses. However, as we showed in Figure 4, a dyadic analysis reveals important additional infor-

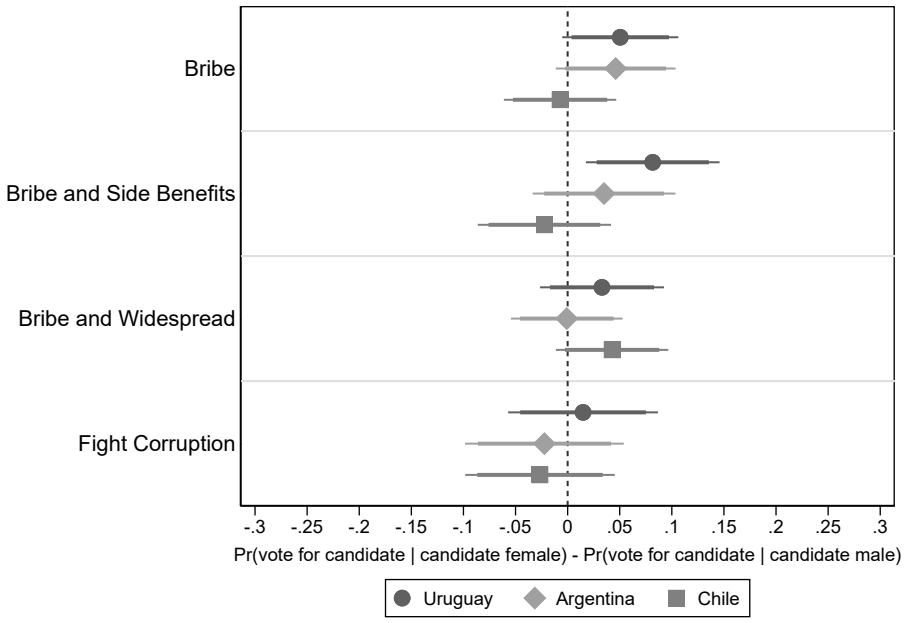


Figure 7: Replication of Figure 2 in Le Foulon and Reyes-Housholder (2021), with corrected gender variable. Point estimates indicate AMCEs for Candidate Gender by Country. Lines show 95% and 90% confidence intervals using standard errors clustered by respondent. To match the original results, no Bonferroni correction is used.

mation. Specifically, female incumbents accused of corruption receive more voter support than equivalent male incumbents when facing a male challenger. There is no advantage for a female incumbent implicated in bribery (compared to an equivalent male incumbent) when the challenger is female. This finding is consistent with the idea that voters believe that women are less susceptible to corruption, and thus are especially skeptical of that a female incumbent will be more corrupt than a male challenger even if she has been accused of bribery.

Conclusion

The balance of our analysis suggests the following substantive inferences about the link between gender and corruption in Latin America:

1. Voters appear less likely to punish women politicians implicated in bribery compared

to their male counterparts, as we infer from the subjects' higher willingness to support corrupt female candidates in the survey experiment. This is consistent with the idea that voters expect women to be intrinsically less corrupt.

2. Women subjects are less willing to support candidates implicated in corruption. This is consistent with the idea that women prioritize corruption as a political issue, perhaps due to different internalized values.

We do *not* find evidence that women voters hold women politicians to a higher standard, as was seen in Eggers, Vivyan and Wagner (2018). Nor do we find evidence for dramatically different behavior within Chile, Argentina, and Uruguay, although small differences might be detectable in a larger sample.

How do our findings speak to the larger question of why gender and corruption are causally linked? We believe that the answer is more complicated than the apparently direct implication that women are intrinsically more averse to corruption than men. Of course, that is true. But we also cannot ignore the weight of prior evidence suggesting no such connections in other contexts.

We are therefore reluctant to conclude that value differences between men and women fully explain the relationship between women in parliament and corruption that we see in cross-national data. Surely they play a role in this link at some places and times, as they seem to do in contemporary Latin America. But the larger history of this research program indicates that there are many potential explanations, that some of these explanations fail in important contexts, and that in other contexts there appears to be evidence for multiple causal mechanisms operating simultaneously (e.g., Esarey and Schwindt-Bayer, 2019). Perhaps the most important lesson to draw from this research going forward is that the push and pull of multiple causal mechanisms is at work in the larger phenomenon of gender and corruption.

References

Barnes, Tiffany D. and Emily Beaulieu. 2014. "Gender Stereotypes and Corruption: How Candidates Affect Perceptions of Election Fraud." *Politics & Gender* 10(03):365–391.

Barnes, Tiffany D. and Emily Beaulieu. 2018. "Women Politicians, Institutions, and Perceptions of Corruption." *Comparative Political Studies* 52(1):134–167.

Barnes, Tiffany D., Emily Beaulieu and Gregory W Saxton. 2018. "Restoring trust in the police: Why female officers reduce suspicions of corruption." *Governance* 31(1):143–161.

Batista Pereira, Frederico. 2020. "Do Female Politicians Face Stronger Backlash for Corruption Allegations? Evidence from Survey-Experiments in Brazil and Mexico." *Political Behavior* (OnlineFirst):1–20. URL: `https://doi.org/10.1007/s11109-020-09602-9`.

Bauhr, Monika, Nicholas Charron and Lena Wängnerud. 2018. "Exclusion or interests? Why females in elected office reduce petty and grand corruption." *European Journal of Political Research* (OnlineFirst). URL: `https://doi.org/10.1111/1475-6765.12300`.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics* 124(4):1497–1540.

Brollo, Fernanda and Ugo Troiano. 2016. "What happens when a woman wins an election? Evidence from close races in Brazil." *Journal of Development Economics* 122:28–45.

Correa Martínez, Wendy and Michael Jetter. 2016. "Isolating causality between gender and corruption: An IV approach." Center for Research in Economics and Finance (CIEF), Working Papers, No. 16-07. Available at SSRN: `https://ssrn.com/abstract=2756794`.

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.

Eggers, Andrew C., Nick Vivyan and Markus Wagner. 2018. "Corruption, Accountability, and Gender: Do Female Politicians Face Higher Standards in Public Life?" *Journal of Politics* 80(1):321–326.

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 Leslie A. Schwindt-Bayer. 2019. "Estimating Causal Relationships Between Women's Representation in Government and Corruption." *Comparative Political Studies* 52(11):1713–1741.

Esarey, Justin and Leslie Schwindt-Bayer. 2018. "Women's Representation, Accountability, and Corruption in Democracies." *British Journal of Political Science* 48(3):659–690.

Friedman, Jerome, Trevor Hastie, Robert Tibshirani et al. 2001. *The elements of statistical learning*. Vol. 1 New York: Springer.

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

Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22(1):1–30.

Jha, Chandan Kumar and Sudipta Sarangi. 2018. "Women and corruption: What positions must they hold to make a difference?" *Journal of Economic Behavior & Organization* 151:219–233.

Klašnja, Marko, Joshua A. Tucker and Kevin Deegan-Krause. 2016. "Pocketbook vs. Sociotropic Corruption Voting." *British Journal of Political Science* 46(1):67–94.

Klašnja, Marko, Noam Lupu and Joshua A. Tucker. 2020. "When Do Voters Sanction Corrupt Politicians?" *Journal of Experimental Political Science* pp. 1–11.

Le Foulon, Carmen and Catherine Reyes-Housholder. 2021. "Candidate sex, corruption and vote choice." *Electoral Studies* 69:102270. URL: `https://www.sciencedirect.com/science/article/pii/S0261379420301487`.

Paweenawat, Sasiwimon W. 2018. "The gender-corruption nexus in Asia." *Asian-Pacific Economic Literature* 32(1):18–28.

Scargle, Jeffrey D. 2000. "Publication Bias: The "File-Drawer" Problem in Scientific Inference." *Journal of Scientific Exploration* 14:91–106.

Schwindt-Bayer, Leslie A., Justin Esarey and Erika Schumacher. 2018. Gender and Citizen Responses to Corruption among Politicians The US and Brazil. In *Gender and Corruption*, ed. Helena Stensöta and Lena Wängnerud. Cham, Switzerland: Palgrave Macmillan chapter 4, pp. 59–82.

Schwindt-Bayer, Leslie A. and Margit Tavits. 2016. *Clarity of Responsibility, Accountability and Corruption*. New York: Cambridge University Press.

Sung, Hung-En. 2003. "Fairer sex or fairer system? Gender and corruption revisited." *Social Forces* 82:703–723.

Swamy, Anand, Stephen Knack, Young Lee and Omar Azfar. 2001. "Gender and corruption." *Journal of Development Economics* 64(1):25–55.

Wängnerud, Lena. 2014. Gender and corruption. In *Routledge Handbook of Political Corruption*, ed. Paul M. Heywood. Routledge chapter 20, pp. 288–298.

Wängnerud, Lena. 2020. "Why Women in Elected Assemblies Reduce Levels of Corruption: The Indirect Approach." Presented to the ECPR General Conference Online, August 24-28, 2020. URL: `https://ecpr.eu/Filestore/paperproposal/20c5788c-29a0-4464-9834-76c888414f7a.pdf` accessed 1/14/2021.

Wängnerud, Lena, Maria Solevid and Monika Djerf-Pierre. 2019. "Moving beyond categorical gender in studies of risk aversion and anxiety." *Politics & Gender* 15(4):826–850.

Appendix A: Vignette Experiment Text

The English translation of the text for the conjoint survey experiment vignette is taken from Section A2, page 4 of the online appendix for Klašnja, Lupu and Tucker (2020). The terms in brackets represent the changes in wording randomly assigned as part of the treatment.

> Imagine that you are voting in an election for mayor with two candidates. The economic conditions of the municipality have [improved/worsened] since the last-election.
>
> [María / Alberto] López is the incumbent [<right party> / <left party> / independent] mayor running for reelection. [The newspaper <left newspaper> / The newspaper <right newspaper> / Judicial officials] [praised López's efforts to punish public employees accepting bribes in exchange for public concessions / accused López of accepting bribes in exchange for public concessions during [her / his] term / accused López of accepting bribes in exchange for public concessions during [her / his] term, a practice that was then common throughout the province / accused López of accepting bribes in exchange for public concessions during [her / his] term, but some suggest that this practice brought construction jobs to the municipality].
>
> The other candidate is [Isabel / Juan] Arias from [<right party> / <leftparty> / independent]. Arias had been the mayor of the municipality before López took office. [The newspaper <left newspaper> / The newspaper <right newspaper> / Judicial officials] [praised Arias's efforts to punish public employees accepting bribes in exchange for public concessions / accused Arias of accepting bribes in exchange for public concessions during [her / his] term / accused Arias of accepting bribes in exchange for public concessions during [her / his] term, a practice that was then common throughout the province / accused Arias of accepting bribes in exchange for public concessions during [her / his] term, but some suggest that this practice brought construction jobs to the municipality].

Appendix B: OLS Model Tables

Table 2: OLS Models of vote choice by candidate gender

	All Respondents		Female Respondents		Male Respondents	
	(1)	(2)	(3)	(4)	(5)	(6)
Bribes	-0.374***	-0.373***	-0.363***	-0.366***	-0.385***	-0.379***
	(-19.64)	(-20.60)	(-13.41)	(-14.11)	(-14.31)	(-14.94)
Bribes common	-0.382***	-0.384***	-0.398***	-0.401***	-0.364***	-0.365***
	(-19.82)	(-21.07)	(-14.56)	(-15.34)	(-13.32)	(-14.36)
Bribes but jobs	-0.287***	-0.285***	-0.299***	-0.293***	-0.271***	-0.276***
	(-13.94)	(-15.11)	(-10.37)	(-10.93)	(-9.20)	(-10.31)
Candidate: Female	-0.0110	-0.0110	-0.0194	-0.0194	-0.00263	-0.00263
	(-0.51)	(-0.51)	(-0.62)	(-0.62)	(-0.09)	(-0.09)
Bribes X C: Female	0.0405		0.0370		0.0442	
	(1.51)		(0.96)		(1.18)	
Bribes common X C: Female	0.0360		0.0376		0.0303	
	(1.34)		(1.00)		(0.80)	
Bribes but jobs X C: Female	0.0423		0.0580		0.0243	
	(1.46)		(1.41)		(0.59)	
Any bribe treatment X C: Female		0.0396		0.0442		0.0332
		(1.68)		(1.30)		(1.01)
Chile	-0.00957	-0.00953	-0.00626	-0.00632	-0.0119	-0.0121
	(-1.06)	(-1.06)	(-0.49)	(-0.49)	(-0.94)	(-0.95)
Uruguay	0.0103	0.0103	0.0102	0.0101	0.0122	0.0121
	(1.15)	(1.15)	(0.81)	(0.81)	(0.96)	(0.95)
Constant	0.542***	0.542***	0.534***	0.534***	0.549***	0.549***
	(33.81)	(33.82)	(22.93)	(22.93)	(24.85)	(24.87)
Observations	8668	8668	4408	4408	4260	4260
AIC	9934.5	9930.6	4893.4	4889.9	5041.6	5037.9
BIC	10005.2	9987.1	4957.3	4941.0	5105.2	5088.8

t statistics in parentheses. Standard errors clustered by subject.

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

Figure 8: Replication of the AMCE for candidate gender in Figure 1 of Klašnja, Lupu and Tucker (2020)

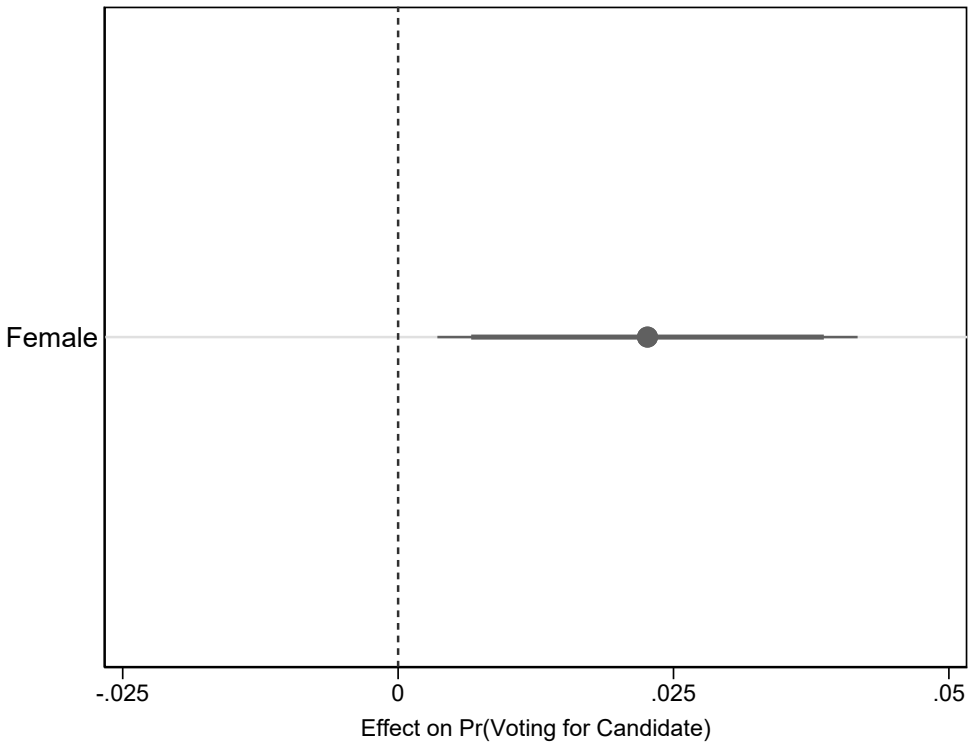


Table 3: OLS Models of vote choice by respondent gender

	All Candidates		Female Candidates		Male Candidates	
	(1)	(2)	(3)	(4)	(5)	(6)
Bribes	-0.362***	-0.349***	-0.341***	-0.333***	-0.385***	-0.366***
	(-19.06)	(-19.62)	(-12.89)	(-13.43)	(-14.31)	(-14.54)
Bribes common	-0.347***	-0.359***	-0.333***	-0.346***	-0.364***	-0.375***
	(-18.27)	(-20.29)	(-12.60)	(-14.00)	(-13.32)	(-14.84)
Bribes but jobs	-0.259***	-0.261***	-0.248***	-0.243***	-0.271***	-0.279***
	(-12.43)	(-13.95)	(-8.54)	(-9.33)	(-9.17)	(-10.57)
Respondent: Female	-0.0222	-0.0222	-0.0305	-0.0305	-0.0123	-0.0123
	(-1.10)	(-1.10)	(-1.06)	(-1.07)	(-0.41)	(-0.40)
Bribes X R: Female	0.0180		0.0160		0.0223	
	(0.67)		(0.42)		(0.58)	
Bribes common X R: Female	-0.0313		-0.0272		-0.0340	
	(-1.16)		(-0.73)		(-0.88)	
Bribes but jobs X R: Female	-0.0115		0.00695		-0.0295	
	(-0.40)		(0.17)		(-0.71)	
Any bribe treatment X R: Female		-0.00785		-0.00143		-0.0130
		(-0.33)		(-0.04)		(-0.38)
Chile	-0.00889	-0.00916	-0.0177	-0.0178	-0.000221	-0.00101
	(-0.99)	(-1.02)	(-1.22)	(-1.23)	(-0.02)	(-0.07)
Uruguay	0.0114	0.0110	0.0268	0.0264	-0.00501	-0.00555
	(1.27)	(1.23)	(1.82)	(1.79)	(-0.34)	(-0.38)
Constant	0.547***	0.547***	0.543***	0.543***	0.550***	0.551***
	(36.95)	(36.97)	(25.00)	(25.02)	(24.37)	(24.40)
Observations	8668	8668	4398	4398	4270	4270
AIC	9929.0	9928.7	5194.1	5191.6	4736.2	4735.3
BIC	9999.7	9985.3	5258.0	5242.7	4799.8	4786.1

t statistics in parentheses. Standard errors clustered by subject.

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

Table 4: OLS Dyadic Models of vote choice by treatment status

	(1)	(2)
Female Incumbent	0.0167	
	(0.55)	
Incumbent Bribed	-0.342***	
	(-14.27)	
Female Inc. X Incumbent Bribed	0.0372	
	(1.10)	
Male Inc., Incumbent Accepted Bribe		-0.263***
		(-5.94)
Male Inc., Challenger Accepted Bribe		0.209***
		(4.31)
Male Inc., Both Accepted Bribe		-0.166***
		(-3.88)
Female Inc., Neither Accepted Bribe		0.0467
		(0.80)
Female Inc., Incumbent Accepted Bribe		-0.192***
		(-4.20)
Female Inc., Challenger Accepted Bribe		0.215***
		(4.47)
Female Inc., Both Accepted Bribe		-0.118**
		(-2.74)
Chile	0.00675	0.00676
	(0.42)	(0.43)
Uruguay	0.0479**	0.0465**
	(2.96)	(2.89)
Constant	0.508***	0.356***
	(21.21)	(8.38)
N	4334	4334
AIC	5020.0	4964.4
BIC	5058.3	5028.1

t statistics in parentheses

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

Table 5: OLS Dyadic Models of vote choice when incumbent bribed

	Male Challengers	Female Challengers
Female Incumbent	0.00306	0.139***
	(0.09)	(3.91)
Challenger Bribed	0.0681*	0.128***
	(2.27)	(5.07)
Female Inc. X Challenger Bribed	0.0348	-0.0806
	(0.81)	(-1.89)
Chile	0.00836	0.0451
	(0.34)	(1.86)
Uruguay	0.0167	0.0344
	(0.67)	(1.40)
Constant	0.136***	0.0469
	(4.76)	(1.91)
N	1620	1654

t statistics in parentheses

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

Table 6: OLS Dyadic Models of vote choice by respondent gender

	(1)	(2)
Female respondent	-0.0246	
	(-0.81)	
Incumbent bribed	-0.316***	
	(-13.26)	
Female respondent X Incumbent bribed	-0.0131	
	(-0.39)	
Male respondent, Incumbent bribed		-0.225***
		(-5.17)
Male respondent, Challenger bribed		0.212***
		(4.58)
Male respondent, Both bribed		-0.148***
		(-3.61)
Female respondent, Neither bribed		0.000615
		(0.01)
Female respondent, Incumbent bribed		-0.275***
		(-6.48)
Female respondent, Challenger bribed		0.165***
		(3.57)
Female respondent, Both bribed		-0.182***
		(-4.48)
Chile	0.00735	0.00734
	(0.46)	(0.46)
Constant	0.529***	0.379***
	(22.60)	(9.45)
Uruguay	0.0481**	0.0464**
	(2.97)	(2.89)
N	4334	4334
AIC	5026.1	4969.7
BIC	5064.3	5033.5

t statistics in parentheses

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

Table 7: OLS Dyadic Models of vote choice by candidate and respondent gender

	(1) Male Respondents	(2) Female Respondents	(3) All Respondents
Male Inc., Incumbent Accepted Bribe	-0.234***	-0.293***	-0.263***
	(-3.75)	(-4.63)	(-5.94)
Male Inc., Challenger Accepted Bribe	0.222**	0.196**	0.209***
	(3.28)	(2.82)	(4.31)
Male Inc., Both Accepted Bribe	-0.144*	-0.187**	-0.166***
	(-2.41)	(-3.03)	(-3.88)
Female Inc., Neither Accepted Bribe	0.0558	0.0368	0.0467
	(0.71)	(0.42)	(0.80)
Female Inc., Incumbent Accepted Bribe	-0.157*	-0.225***	-0.192***
	(-2.44)	(-3.47)	(-4.20)
Female Inc., Challenger Accepted Bribe	0.261***	0.169*	0.215***
	(3.93)	(2.45)	(4.47)
Female Inc., Both Accepted Bribe	-0.0938	-0.143*	-0.118**
	(-1.57)	(-2.29)	(-2.74)
Chile	0.0150	-0.00159	0.00676
	(0.65)	(-0.07)	(0.43)
Uruguay	0.0633**	0.0311	0.0465**
	(2.68)	(1.43)	(2.89)
Constant	0.341***	0.371***	0.356***
	(5.79)	(6.04)	(8.38)
N	2130	2204	4334

t statistics in parentheses

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

Table 8: Test for Pooling Models by Country, Candidate Gender

	Bribes		Bribes w/ Side Benefits		Widespread Bribes		Fought Bribes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chile	0.0258		0.0195		-0.00422		-0.00422	
	(0.94)		(0.60)		(-0.16)		(-0.16)	
Uruguay	-0.0179		-0.0699*		0.0254		0.0254	
	(-0.67)		(-2.22)		(0.89)		(0.89)	
Candidate: Female		0.0285		0.0298		0.0255		0.0255
		(1.74)		(1.54)		(1.56)		(1.56)
Argentina x Female	0.0461		0.0350		-0.000799		-0.000799	
	(1.57)		(1.00)		(-0.03)		(-0.03)	
Chile x Female	0.0185		-0.00280		0.0384		0.0384	
	(0.68)		(-0.09)		(1.38)		(1.38)	
Uruguay x Female	0.0326		0.0118		0.0584*		0.0584*	
	(1.14)		(0.35)		(2.04)		(2.04)	
Intercept (Baseline: Argentina)	0.166***	0.169***	0.273***	0.256***	0.153***	0.160***	0.153***	0.160***
	(8.49)	(15.38)	(11.78)	(19.63)	(7.81)	(14.21)	(7.81)	(14.21)
N	2260	2260	2142	2142	2151	2151	2151	2151
AIC	2126.0	2121.4	2606.4	2607.8	1925.0	1922.8	1925.0	1922.8
BIC	2160.3	2132.9	2640.5	2619.1	1959.0	1934.2	1959.0	1934.2

t statistics in parentheses. Standard errors clustered by subject.

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

Table 9: Test for Pooling Models by Country, Respondent Gender

	Bribes		Bribes w/ Side Benefits		Widespread Bribes		Fought Bribes	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chile	-0.0326		-0.0182		0.0157		0.0157	
	(-1.13)		(-0.54)		(0.58)		(0.58)	
Uruguay	-0.0796**		-0.0294		0.0765**		0.0765**	
	(-2.78)		(-0.88)		(2.62)		(2.62)	
Respondent: Female		-0.00427		-0.0336		-0.0523***		-0.0523***
		(-0.27)		(-1.81)		(-3.34)		(-3.34)
Argentina x Female	-0.0653*		-0.0304		-0.0365		-0.0365	
	(-2.29)		(-0.91)		(-1.39)		(-1.39)	
Chile x Female	-0.0380		-0.0285		-0.0163		-0.0163	
	(-1.32)		(-0.89)		(-0.62)		(-0.62)	
Uruguay x Female	-0.0232		-0.0924**		-0.0188		-0.0188	
	(-0.79)		(-2.91)		(-0.73)		(-0.73)	
Intercept (Baseline: Argentina)	0.223***	0.185***	0.305***	0.289***	0.170***	0.199***	0.170***	0.199***
	(10.12)	(16.31)	(12.72)	(21.25)	(9.05)	(17.21)	(9.05)	(17.21)
N	2260	2260	2142	2142	2151	2151	2151	2151
AIC	2122.7	2124.4	2609.5	2607.2	1914.6	1915.0	1914.6	1915.0
BIC	2157.0	2135.9	2643.5	2618.5	1948.7	1926.3	1948.7	1926.3

t statistics in parentheses. Standard errors clustered by subject.

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