Gender and Citizen Responses to Corruption among Politicians: The U.S. and Brazil

Leslie Schwindt-Bayer
Justin Esarey
Erika Schumacher
Rice University

Book Chapter in *Gender and Corruption: Revising relationships and establishing new avenues for research*
Helena Stensöta and Lena Wängnerud eds.

Abstract

Recent research has found that a relationship exists between women's representation and corruption, but is limited to democracies with higher levels of electoral accountability. Research has not established definitively why this relationship exists, however. One possible explanation is rooted in voter perceptions of male and female politicians' corruptibility. This explanation suggests that voters may perceive the corruptibility of female and male politicians differently, with women being less likely to be viewed as corrupt than men. If that is the case, then voters may punish corrupt women more harshly than corrupt men because their behavior does not conform to gender stereotypes. Women in office will recognize this and become less likely to engage in corruption than men because of the harsher punishment they are likely to receive if caught. Yet very little research has explored this argument empirically. In this chapter, we examine this theory with an analysis of whether voters perceive of comparable male and female candidates differently in terms of how likely they are to be involved in a corruption scandal and punish them differently when they are involved in corruption. We conduct survey experiments in two countries, the United States (with high electoral accountability) and Brazil (with moderate to low electoral accountability). If differential treatment is the causal mechanism linking women's representation and corruption, we should find (a) evidence of differential perception of corruptibility in the United States but not Brazil, and (b) evidence of differential punishment for corruption in the United States. We find weak and statistically uncertain evidence that women are perceived as less corruptible than men in both countries, and we find no evidence that women will be punished more harshly than men for corruption scandals.

Gender and Citizen Responses to Corruption among Politicians: The U.S. and Brazil

As this book has made clear already, research on the relationship between women's representation and corruption has established an important correlation between the two—greater women's representation is related to reduced corruption (Dollar, Fisman, and Gatti 2001; Swamy et al. 2001). Yet, studies also show that the relationship is context-dependent: it exists in some countries but not others (Alatas et al. 2009; Esarey and Chirillo 2013; Schwindt-Bayer 2016). One explanation for this is that the relationship is conditional upon electoral accountability, whereby the link is stronger when electoral accountability is high and weaker when it is low (Esarey and Schwindt-Bayer 2016). However, research has not established exactly why this would be the case.

One possible explanation is rooted in how politician gender shapes voter perceptions of politician corruptibility and whether voters punish corrupt male and female politicians differently. Specifically, voters may perceive of the corruptibility of female and male politicians differently, with women being less likely to be viewed as corrupt than men. If that is the case, then voters may punish corrupt women more harshly than corrupt men because their behavior does not conform to gender stereotypes of women being more honest, trustworthy, and less corrupt than men. Women in office should recognize this and be less likely to engage in corruption than men. This would only occur, however, in settings of high accountability where voters have the ability to punish elected officials directly at the voting booth. It should be less likely to occur in settings of low accountability. Unfortunately, this causal mechanism has not been thoroughly explored by quantitative empirical research.

In this chapter, we conduct a pair of survey experiments to determine whether empirical support for this "differential treatment" explanation exists. We conduct one experiment in a

country with high electoral accountability—the United States—and the other experiment in a country with moderate to low electoral accountability—Brazil.[1] We ask citizens in both countries to evaluate the corruptibility of a hypothetical governor. In the U.S., where voters can and do exercise electoral accountability, we also ask them whether they would vote for a corrupt governor. The treatment in both questions is the sex of the governor, with half of the survey respondents evaluating a female governor and the other half evaluating a male governor. If the theory just described is correct, we should see that (a) voters in the high accountability context perceive women as less corrupt, but those in the low accountability context do not, and (b) voters in the high accountability context are less likely to vote for a corrupt female governor compared to a corrupt male.

We find some statistically uncertain evidence that respondents *in both countries* are more likely to think male governors will be embroiled in a corruption scandal during their term in office than female governors. The substantive magnitude of this gender difference is similar in both countries; but, in both countries, the difference is at or just beyond conventional thresholds for statistical significance. In addition, the differences in Brazil are concentrated in one demographic group—women—and in one treatment condition. That finding is inconsistent with the theory we were testing, and we are unsure of the explanation for the finding. We also find that voters in the United States do not differ in their punishment of corrupt male and female governors. Our overall conclusion is that this evidence leans against differential perception and punishment as the causal mechanism for a context-dependent relationship between women's representation and corruption relationship.

[1] Both experiments received human subjects approval from the Rice University Institutional Research Board (IRB). U.S. experiment: study number IRB-FY2017-332; Brazilian experiment: study number IRB-FY2016-607.

The Differential Treatment Theory of Gender and Corruption

The differential treatment theory of gender and corruption has two key parts. First, it argues that voters perceive of the corruptibility of male and female elected officials differently, with women being less likely to be viewed as corrupt than men. This idea is rooted in traditional stereotype literature that links women in office with more feminine stereotypes of honesty and trustworthiness. If voters view women as more honest and trustworthy than men, then they may view women in elected office as less corruptible than their male counterparts, as well. Second, the theory presumes that those views of corruptibility will translate into voting behavior that more harshly punishes women than men when women deviate from the gendered stereotype of being less corruptible. In other words, if women engage in corruption, then voters will be more likely to vote against them than they would a man who engaged in corruption. As Dolan (2010, 70) writes, "Gender stereotypes about the abilities and traits of political women and men are clear and well documented and could easily serve to shape an individual's evaluations about the appropriate level and place for women in office."

Much research exists that shows that gendered stereotypes of male and female politicians exist (Alexander and Andersen 1993; Dolan 2004; Dolan 2010, 2014; Huddy and Terkildsen 1993; Murray 2010; Sanbonmatsu 2002), and more specifically, citizens tend to view women as more honest than men (Alexander and Andersen 1993; Dolan 2004; Dolan 2014; McDermott 1998). Studies have also found that stereotypes can translate into the political attitudes and behaviors of voters (Alexander and Andersen 1993; Dolan 2010; Fox and Smith 1998; Sanbonmatsu 2002). Sanbonmatsu (2002, 31), for example, found that "Voters' gender schemas give rise to a baseline preference to support either male or female candidates." The honesty stereotype itself has been found to contribute to more willingness to support female candidates

than male ones. McDermott (1998) found in a California study that individuals who thought ethics were an important problem in government were more likely to vote for a female candidate for governor.

Studies evaluating how gender stereotypes influence views of women and men in office suggest some potential differential treatment around scandals, both corruption-based and non-corruption-based. Funk (1996) finds that voters are more likely to punish officials for scandalous behavior when those officials are viewed as warm. That study creates two officials: an official that is strong and competent, and an official that is warm and charismatic. Subjects of the study evaluated these officials in one of two randomized conditions: either the official has undergone a marriage scandal or a tax evasion scandal. Their study showed that voters on average viewed the tax evasion scandal as more severe for both candidates. However, candidates viewed as "competent" were punished mildly for a marriage scandal, whereas candidates viewed as "warm" were punished almost as severely for a marriage scandal as they were for a tax scandal.

Focusing on corruption, more specifically, Żemojtel-Piotrowska (2016) examined how women were treated in response to corruption scandals and found some evidence of less positive evaluations of female politicians who had been associated with a corruption scandal. Barnes and Beaulieu (2014) also examined corruption with a national survey experiment in the U.S. and found that the presence of women in government reduced suspicions of fraud. Then, in a related study, they and Saxton (2017) found that citizens think that women will be less corrupt police officers than men. However, they found that the most compelling explanations for this were the fact that people think women are more risk averse than men and more likely to be political outsiders. Viewing women as more honest than men did not lead to any stronger views that women will be less corrupt police officers.

Much of this research suggests support for a differential treatment explanation for why women's representation might lead to lower levels of corruption. But, it has not been tested comparably in high and low accountability political contexts, where recent research suggests the relationship between women's representation and corruption varies. If differential treatment explains the relationship between women's representation and corruption, then we should expect to find empirical support for the following two hypotheses in high accountability political systems:

Hypothesis 1: Fewer voters will view a female politician as corruptible than a male politician.

Hypothesis 2: More voters will vote against a corrupt female politician than a corrupt male politician.

In low accountability systems, there are two possible outcomes consistent with a theory of differential treatment by gender. Voters might view female and male politicians as equally corruptible, and therefore have no differential expectation of their behavior on which to base disproportionate punishment (and thus women in office will be no less corrupt than men). It could also be the case, however, that voters view women as less corruptible than men, but nevertheless do not behave differently by disproportionately punishing them for corruption (and therefore providing little incentive for women to be less corrupt).

A Test in the United States and Brazil

We analyze the differential treatment theory in the high accountability political system of the United States and the moderate to low accountability Brazilian context. Esarey and Chirillo (2013) suggest that the relationship between women's representation and corruption is conditional upon the level of democracy in countries (more democratic = more accountability), and Esarey and Schwindt-Bayer (2016) identify four additional indicators of electoral accountability that they find moderate the women's representation and corruption relationship—the absence of corruption norms, a parliamentary system of government, freedom of the press, and the personalistic nature of electoral rules.

According to these criteria, the U.S. is a high accountability context. Although it is a presidential system, it is a strong democracy with significant freedom of the press, a general absence of corruption norms, and personalistic electoral rules in the form of single-member districts for the House of Representatives and two-member districts for the Senate with party primaries.[2] In contrast, Brazil is in the moderate to low range on many of these dimensions. It did score 8 out of 10 on the Polity Index and 2.0 on the Freedom House Index in 2015, placing it in the "democracy" and "free" categories of each organization, respectively. However, it has corruption norms, with a corruption score of 43 out of 100 (100 is "clean," on the Transparency International Corruption Perceptions Index) in 2012 and 2014. This ranked it 76 out of 167 countries and placed it just ahead of countries such as India, Thailand, and China. It scored a 45 on the 2015 Freedom House Freedom of the Press index (on a scale of 1 to 100, with a higher score being less free) and was considered having only a "partly free" press. It is also a

[2] The U.S. scored 22 on the 2015 Freedom House Freedom of the Press ranking indicating that its press is "free." It scored a 74 out of 100 (100 = clean) on the Transparency International Corruption Perceptions Index in 2014, ranking it the 16th cleanest government out of 167 countries. On Johnson and Wallack's (2005) personalism index, it scored a 10 out of 13, with 13 being the most personalistic.

presidential system, rather than a parliamentary one, which indicates less accountability on the form of government dimension. It does have an open-list proportional representation electoral system in the lower house of the national parliament, which is highly personalistic, but it scored only a 7 on the Johnson and Wallack (2005) personalism scale (range is 1 to 13, with 13 being the most personalistic).

The U.S. Experiment

We conducted a survey experiment in the United States to explore whether citizens perceive of the corruptibility of female and male elected officials differently and whether or not they would vote for a corrupt politician. We focused on a hypothetical elected governor for this study and asked respondents to imagine themselves in a neighborhood like their own but in a different state to minimize bias from an experience with an actual governor in their own state (Schwindt-Bayer and Tavits 2016; Winters and Weitz-Shapiro 2013). We provided every participant in the survey experiment with a short description of this governor that varied only on the sex of the governor:

> *Imagine you live in a neighborhood like yours, but in a different state. In that state, a [man/woman] from your party was just elected governor. The new governor promises to create jobs, improve access to healthcare and education and fight crime and corruption. [His/her] approval ratings are fairly high, and [he/she] has strong support from many citizens in the state.*

We then asked respondents to answer a question about the corruptibility of that governor. Specifically, the question was "How likely do you think this governor would be involved in a corruption scandal at some time during the term?" with answers given on a four-point scale[3] (from "very likely" to "not likely at all"). Our analyses below invert and dichotomize the responses into *not likely* (=0) and *likely* (=1) for ease of analysis and presentation.

After answering this question, respondents were told that the governor they had just read about had been accused of corruption recently. The prompt follows, and again, it varied only on the sex of the governor, with the respondents assigned a female governor treatment in the previous prompt also assigned a female governor treatment in this prompt and those having read about a male governor again assigned the male governor treatment:

> *Now, suppose the governor you just read about became embroiled in a corruption scandal while serving in office. Specifically, a well-respected newspaper has reported that [he/she] illegally accepted campaign contributions for [his/her] upcoming reelection campaign in exchange for awarding government contracts to donors.*

We then asked respondents if they would consider electing this person again. Specifically, we asked "Would you vote for this governor if he/she was running for re-election in an upcoming election?" and allowed for yes (=1) or no (=0) responses. This treatment and question allows us to test whether citizens punish male and female governors differently when they engage in corruption. Overall, then, we are exploring two parts of the theory of differential

[3] Respondents in this survey could choose to leave this or any other question blank.

treatment of male and female politicians: whether voter attitudes towards the likelihood of governors engaging in corruption differ depending on the sex of a governor, and whether female governors are less likely than male governors to be re-elected after engaging in corruption.

We used Qualtrics for the survey platform and sampling. The Qualtrics sample is a convenience sample, a common sampling technique for these kinds of survey experiments. A convenience sample is not necessarily a representative sample, but the Qualtrics sample does contain significant diversity in its respondent characteristics. Our sample included a total of 422 respondents: 210 were men and 210 were women (with two respondents not answering this question about their gender).[4] Half of the sample received the male governor treatment and the other half received the female governor treatment.

Additionally, we asked questions at the beginning of the survey about respondents' demographics and political interest. Specifically, we asked about respondents' gender (male/female), age in years, education level,[5] race/ethnicity,[6] region of residence,[7] and political interest.[8] The treatment conditions appear well-balanced on these covariates: we found no evidence of statistically significant differences in any of these contextual variables when comparing the subjects in each treatment condition.

[4] Drawing from the language used in the National Election Study, the question asked for a respondent's "gender," not "sex." Thus, we use "gender" to discuss this question and the findings in this section. The Brazilian experiment, by contrast, asked for a respondent's "sex."

[5] Respondents could select from one of the following categories: less than high school degree; high school graduate (high school diploma or equivalent including GED); some college but no degree; associate degree in college (2-year); bachelor's degree in college (4-year); master's degree; doctoral degree; professional degree (JD, MD).

[6] Respondents could select among the following categories, including the possibility of selecting multiple options: White/non-Hispanic, Black or African-American, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander, Hispanic, or Other. The number of respondents who reported being Pacific Islander, Native American, or Other was so small that we classified all such responses as being in a combined "Other" category to use in our analyses. Because these are not mutually exclusive categories, Appendix Table A.1 reports coefficients for all categories.

[7] Respondents reported their state of residence. We then classified states as being in one of four regions: North, Midwest, Southeast, and West.

[8] Respondents could select from one of the following four categories: very interested, somewhat interested, not very interest, and not interested at all.

United States Findings

We examine the treatment effects with simple bivariate comparisons of the percentage of respondents who viewed the governors as corruptible and would not vote for them again.[9] Overall, the results show some difference in the perceptions of the corruptibility of male and female governors and who punishes corrupt male and female governors but the differences are only borderline statistically significant. Focusing first on the corruptibility question, we find that the proportion of respondents that answered that a governor was "somewhat likely" or "very likely" to engage in a corruption scandal while in office was somewhat different for female governors and male governors (the rightmost column in Figure 1). Forty-five percent of respondents considered the male governor likely to engage in corruption while in office, while 36% of respondents considered the female governor likely. This is almost a nine percentage point difference (95% confidence interval: [-0.956%, 18.7%]); however, a difference of proportions test reports a two-tailed *p*-value of 0.080, a result at the margins of conventional thresholds for statistical significance. In multivariate logit models (see Column 1 of Table A.1), the governor sex treatment is associated with a two-tailed *p*-value of 0.062, another result at the margins of conventional thresholds for statistical significance, with male governors being perceived as more corruptible than female governors.

We then examine responses to the question about corruptibility according to the sex of the respondent, shown in the leftmost and center columns of Figure 1. Fewer respondents of both sexes thought the female governors would be corruptible compared to male governors, but again, the differences are not statistically significant. Thirty-eight percent of men thought female

[9] Specifically, we employ difference of proportions tests using `prop.test` in R (R Core Team 2017). The chi-square values on which the difference of proportions tests are based use the Yates' continuity correction, which is the default in R.

governors would be corruptible compared to 46% of male governors, and 34% of women thought female governors were corruptible compared to 43% of male governors. However, difference of proportions tests find that both of these differences are statistically insignificant using conventional thresholds (possibly because dividing the sample into male and female respondents reduces the power of the experiment). In multivariate logit models, we find no evidence that male and female respondents differed in their views of the corruptibility of male and female governors (see the interaction term in the model of Column 2 of Appendix Table A.1).

To determine whether corrupt male and female governors are differentially punished by respondents, we conduct similar analyses using the vote question in our survey. We find that similar proportions of respondents would vote for male and female governors who had engaged in corruption while in office (the rightmost column in Figure 2). The proportion of respondents who would vote for a governor after a reported scandal is 22% for the male governor and 20% for the female governor. While this suggests a slightly harsher punishment for the female governor, $p = 0.566$ for a difference of proportions test. In other words, there is no substantively meaningful or statistically significant evidence that respondents are less likely to punish the male governor for corruption compared to the female governor. This conclusion is also supported by a multivariate logit model (Column 3 in Table A.1), which shows no statistically significant treatment effect of being a female governor on the respondent's choice to vote for a corrupt candidate.

Disaggregating by sex of the respondent (the leftmost and center columns in Figure 2), we see larger differences for the proportion of men who would vote for male and female governors. Twenty-six percent of men would re-elect a male governor who had engaged in

corruption compared to only 22% who would re-elect a woman. The gap is smaller for female respondents. Nineteen percent would re-elect a man but only 18% would re-elect a woman. Both of these gaps suggest harsher punishment for female governors than male governors, but the *p*-values for difference of proportions tests indicate that neither of these differences is statistically distinguishable from zero. Multivariate logit models find no statistically significant difference for how men and women respond to the treatment of a female corrupt governor (compared to a male corrupt governor) when choosing to vote (Column 4 in Table A.1).

In sum, the survey experiment in the high electoral accountability context of the United States reveals some substantively meaningful (albeit statistically uncertain) differences in how respondents perceive the corruptibility of male or female governors. However, we find no substantively or statistically meaningful evidence that respondents would treat male and female candidates differently at the ballot box if suspected of engaging in corruption. Taken as a whole, the evidence suggests that it is unlikely that differential treatment of male and female governors explains why corruption levels would be lower in high accountability contexts that have women in public office.

The Brazilian Experiment

For the Brazilian analysis, we use a similar survey experiment that was designed by one of this paper's authors for another project (Schwindt-Bayer and Reyes-Housholder 2017). One of the questions in that experiment allows us to test Hypothesis 1 (concerning gender differences in perceived corruptibility) in a moderate to low accountability context.[10] The Brazilian survey experiment presented respondents with one of four treatment prompts that provided a description of a recently elected governor in a hypothetical state varying on the sex of the governor and the

[10] The Brazilian survey experiment was conducted in Portuguese.

past history with a female governor.[11] In this analysis, we are not concerned with how differences in the previous history with female governors might affect views of corruptibility. We are concerned only with differences that may emerge as a result of varying the sex of the governor. As a result, we analyze the survey experiment results below focusing primarily on governor sex differences. Due to the survey design, however, we make sure to note whether those differences exist where a previous history with female governors existed or not.

Following the prompt, the survey asked a set of questions. One of these questions, a question about how corruptible the respondent thought the governor would be, was the same question described above for the U.S. survey experiment. A question about voting for a corrupt politician was not asked; therefore, we cannot test hypothesis 2 in the Brazilian context. However, we can determine whether Brazilians perceive female politicians to be as corruptible as otherwise identical male politicians. If the Brazilian data show no evidence of a gender gap in perceived corruptibility, that could explain in part why there is no relationship between corruption and women in office in contexts with low electoral accountability. If male and female politicians and men are not differentially viewed as corrupt in Brazil, then women in office have little reason to fear greater retribution at the ballot box and may be just as likely to engage in corruption as men.

[11] Specifically, the prompt was the following: "Imagine you live in a neighborhood like yours, but in a different state. In that state, a [man/woman**]** from a moderate party (neither extreme right or extreme left) was just elected governor. In the past, the state [has never had a female governor/has had a female governor**]**. The new governor promises to create jobs, improve access to healthcare and education and fight crime and corruption. [His/Her**]** approval ratings are fairly high, and [he/she**]** has strong support from many citizens in the state." The prompt described the governor as "moderate" to downplay the significance of party ideology. In Brazil, the main cleavage among parties is not left-right, but whether the party supports the executive party in power, so the experiment aimed to minimize party ideology in the prompts (Samuels and Zucco 2014).

In addition to asking about views of corruptibility, the survey contained data on respondent demographics: sex (male/female), age in years, social class,[12] and region.[13] The survey also asked two questions about respondents' race and political interest.[14] To correct for residual imbalances in some covariates,[15] these factors are included as control variables in the multivariate logit models presented in Appendix Tables A.2 and A.3. Respondents were asked two post-treatment manipulation check questions, as well, to determine whether they received the governor sex and past history with a female governor treatments. The first question asked "Is the newly elected governor a man or a woman?" and 84% of respondents answered the question correctly. The second question asked "Has the state already had a female governor?" and 75% of respondents answered this question correctly. Combining the two, 65% of respondents answered both correctly, indicating that they fully received the treatment. In the analyses below, we present results for the full set of respondents and the reduced sample of those who answered both questions correctly. The latter sample offers a stricter test of the hypothesis because it focuses on those respondents who we can be highly confident were paying attention to the survey and comprehended the treatments.

[12] Social class is rated on a six-point scale, with 1 = upper class and 6 = lower class. In our sample, the lowest category had no respondents in it.

[13] Respondents could select from the following categories: north, northeast, southeast, south, and central west.

[14] The race and political interest questions were: 1) Do you consider yourself white, black, brown, indigenous or yellow? (with respondents selecting only one of these categorical options or "other"), and 2) How interested are you in politics? (with answers on a four point ordinal scale from "very interested" to "not interested at all"). Unlike the U.S. experiment, respondents in the Brazilian experiment could choose only one racial category; thus one category (white) is excluded in the analyses of Appendix Tables A.2 and A.3.

[15] The sex distribution across treatment groups was the following (male-female): Male governor, no history: 55.2% - 44.8%; Male governor, history: 44.1% - 55.9%; Female governor, no history: 49.8% - 50.2%; Female governor, history: 51.0% - 49.0%. The chi-square test for independence was statistically significant at conventional levels (p=0.0193) as a result of the sex distributions in the two groups given the male governor treatments not being well balanced. We also found some evidence of imbalance in whether subjects answered both manipulation checks correctly: a chi-square test for independence of treatment and manipulation checks was statistically significant (p<0.001), with the largest difference being an apparent excess of those answering both questions correctly in the "no history, female" treatment. Finally, we found evidence of imbalance in whether subjects were thinking of a specific state or politician (discussed in the next paragraph): a chi-square test for independence between treatment and this question was statistically significant (p=0.007), with respondents in the female governor treatments being more likely to be thinking of a specific state or politician than men.

For robustness, the survey asked respondents a follow-up question after all other questions were asked. The question was "Were you thinking of a specific state in Brazil or a specific politician when you responded to these questions?" If the respondent answered yes, then he/she was prompted to select a state from a drop-down list and/or write in the name of the politician they were thinking of. We use a binary measure of whether a respondent was thinking of a specific state or politician as a control in our multivariate analysis.

Netquest, in São Paulo, Brazil, fielded the survey experiment. They used a convenience sample of Brazilians (Boas 2014; Boas 2015; Samuels and Zucco 2014), but their panel does include Brazilians from every major region and features a fairly balanced dispersion in terms of social class.[16] The sample was block randomized by sex of the respondent to allow comparison of differences in treatment effects for men and women. The sample included a total of 1,600 individuals, 800 men and 800 women, aiming for approximately 200 male and 200 female respondents per treatment group (Boas 2014; Samuels and Zucco 2014).

Brazilian Findings

On the whole, the results of the survey experiment do not strongly support the conclusion that Brazilian respondents perceive of the corruptibility of male and female governors differently. The proportion of respondents who believed that the male governors they were asked to evaluate were "very likely" or "somewhat likely" to be involved in a corruption scandal during their time in office was larger than the proportion of respondents who believed that female governors were likely to be embroiled in corruption; however, this was only true for female respondents in the treatment with a history of female governors in the models where the

[16] For details on Netquest's Brazilian panel characteristics, see http://www.netquest.com/papers/panelbook_en.pdf, page 3.

correct manipulation check restriction was not in place. Other differences were weakly or not at all statistically significant.

Figure 3 presents the results of a simple bivariate comparison of responses to the corruption question by treatment. Almost the same percentage of survey respondents who evaluated a female governor and a male governor in the context of no history with women in office thought the hypothetical governor was likely to become involved in a corruption scandal—52%. In the context of a history with female governors, a six-percentage point difference emerged (95% confidence interval of the difference: [-1.25%, 13.2%]); fifty-three percent of those who evaluated a male governor thought he was likely to become embroiled in a corruption scandal, but only 47% of those who evaluated a female governor thought she was likely to be corruptible. This difference is in the direction we would expect with more respondents thinking men are likely to be corruptible than women and is similar in magnitude to the gender difference in corruptibility we found in the U.S. experiment, but the *p*-value for a difference in proportions test is slightly above any conventional threshold for statistical significance ($p = 0.109$, two-tailed).

When we examine the responses to the corruption question by sex of the respondent, we find that the difference in views of male and female governors in the treatment where a female governor has previously held office comes primarily from female respondents. Only 42% of the women who evaluated a female governor thought she was likely to engage in corruption compared to 56% of the women who evaluated a male governor. A difference of proportions test inside of this subgroup was statistically significant at conventional levels ($p = 0.010$, two-tailed). There was no statistically significant difference in the evaluation of a female governor's corruptibility compared to a male governor's for male respondents in this condition or for either

sex in the treatment condition with no history of a female governor. Thus, the main difference in evaluations of the corruptibility of governors in the Brazilian experiment is that more women think male governors are corruptible than female governors when women have held office previously in the hypothetical state.

Figure 4, however, suggests caution in overstating those findings. Figure 4 presents a bivariate comparison of corruptibility in the male and female governor treatments for only those respondents who correctly responded to both of the survey manipulation checks. This narrows the sample to just those who were paying close attention to the survey and demonstrated that they correctly received the treatment about both the sex of the governor in the prompt and whether the state had a history with at least one female governor. The figures reveal very similar proportions of female respondents, male respondents, and all respondents thinking male and female governors are likely to be involved in a corruption scandal while in office.[17] None of the comparisons of treatment groups are near typical thresholds for statistical significance. The largest substantive difference is, again, among female respondents' views of female and male governors in hypothetical contexts where women have held office, but the difference is only three percentage points and not statistically significant. These results suggest that the previous finding is not as strong as it first appeared. However, it is important to keep in mind that while these comparisons benefit from focusing on respondents that we are confident received the experimental treatment, they suffer from lower power resulting from a reduced sample size (n = 1032 instead of n = 1592), particularly in the models assessing respondent sex differences (there are between 102 and 167 subjects in each of the columns in Figure 8).

[17] Note that there is no statistically significant gender difference in whether respondents got the manipulation questions correct. Sixty-six percent of women got both questions correct as did 64% of men (chi-squared test for difference of proportions p=0.285).

To explore the robustness of the models, we ran several additional analyses. First, we estimated multivariate logit models that included the controls described previously (see Appendix Tables A.2 and A.3). Given that the samples are reasonably well balanced on demographics, it is not surprising that the multivariate results are similar to what we presented in Figures 5-8. The treatment effect (of being a female governor on corruptibility) is only statistically significant at conventional levels in the context of a previous history with female governors (Appendix Table A.2, Columns 1 and 2). There, citizens are less likely to think women are corruptible than they are to think that men are corruptible. This finding is conditional upon the sex of the respondent with the treatment having no effect on men but being statistically significant for women (with $p < 0.01$ for the sum of associated logit coefficients), similar to what the bivariate comparisons showed. When we narrow the analysis just to those respondents who answered the manipulation questions correctly (Appendix Table A.3), those statistically significant differences again disappear.

We also examined whether the results were biased by respondents thinking of specific states and politicians, despite asking them to think only of a hypothetical scenario. About 55% of respondents said they were thinking of a specific state or politician. The multivariate models find that those who were thinking of a specific state or politician were more likely to think the hypothetical governor would be embroiled in a corruption scandal. Bivariate analyses (not shown) of the survey responses by whether or not respondents had someone or some state in mind reveals mild evidence that the results could be biased by this—the previous finding that more respondents think male governors will become corrupt than will female governors in a context where women have been governor previously is only statistically significant (p=0.0748) for those respondents who were thinking of a specific state or politician. Any “Dilma effect” is

negligible, however. Respondents who said they were thinking of Dilma Rousseff specifically (*n* = 115) had no different views of the corruptibility of male and female governors than those who did not report thinking of her.[18]

In sum, our analysis of the survey experiment testing differences in respondents' views of the corruptibility of male and female governors in Brazil reveals that women leaders are thought to be less corruptible than men only among female respondents in a case where subjects are told that there is a history of prior female governors and when those who failed the manipulation checks were included in the analysis. Perceptions of corruptibility do not strongly vary by the sex of a governor in our models. As we expected in a low accountability context, differential treatment of women and men elected officials was limited.

Conclusion

This chapter explores whether a theory of differential treatment of women in public office explains why women's representation is related to reduced corruption. We conduct two survey experiments to do this. One experiment is in the United States, a country with high electoral accountability where, if differential treatment is the correct explanation for the women's representation and corruption relationship, we would expect significant gender stereotyping of female officeholders as less corruptible and greater punishment of women in office who do engage in corruption. The other experiment is conducted in Brazil, a country with lower electoral accountability and where we would expect to see gender stereotyping without differential punishment or no gender stereotyping at all. There is some evidence, albeit uncertain

[18] Specifically, in a linear regression predicting corruptibility, a multiplicative interaction term between (a) a dummy identifying subjects who were thinking of Dilma Rousseff and (b) the female governor treatment is statistically insignificant for all subjects, subjects in the condition with a history of a female governor, and subjects in the condition with no history of a female governor.

and conditional evidence, that women are perceived as less corrupt in both countries, at least in some contexts. However, we find no evidence that voters in the U.S. disproportionately punish women at the ballot box for engaging in corruption. As a whole, these findings lean against the theory that differential treatment of female politicians explains why women's representation leads to less corruption in contexts of high electoral accountability but has little to no relationship to corruption in low accountability contexts.

These findings are suggestive and important but not conclusive. The limited external validity of survey experiments such as these, the fact that we test this in only two countries, and the low power of our design inside of certain subgroups means that more research is necessary to definitively determine if, when, and where voters may perceive of women in office as less corrupt than men and punish them more strongly when they deviate from their anti-corrupt gender stereotypes.

If differential treatment is not the explanation for women's representation reducing corruption levels, then what is? Esarey and Schwindt-Bayer (2016) provided another line of reasoning to support the observational evidence that the relationship between women's representation and corruption is stronger in democracies with high electoral accountability than democracies with low electoral accountability. We argued that it could be linked to differential risk aversion between women and men. Significant evidence exists that women are more risk averse than men, and in high accountability contexts, that risk aversion would be triggered and would reduce the incentive to engage in corruption activities, thereby reducing overall levels of corruption. This explanation has not been subjected to empirical testing either, but the findings of this chapter—that a differential treatment theory is not empirically supported—underscore the

need for that testing. Understanding *why* women's representation may cause reduced corruption in governments continues to be a critically important part of research on gender and corruption.

References

Alatas, Vivi, Lisa Cameron, Ananish Chaudhuri, Nisvan Erkal, and Lata Gangadharan. 2009. Gender, Culture, and Corruption: Insights from an Experimental Analysis. *Southern Economic Journal* 75 (3): 663-680.

Alexander, Deborah, and Kristi Andersen. 1993. Gender as a Factor in the Attribution of Leadership Traits. *Political Research Quarterly* 46: 527-545.

Barnes, Tiffany D, and Emily Beaulieu. 2014. Gender Stereotypes and Corruption: How Candidates Affect Perceptions of Election Fraud.

Barnes, Tiffany D., Emily Beaulieu, and Gregory W. Saxton. 2017. Restoring trust in the police: Why female officers reduce suspicions of corruption. *Governance*: n/a-n/a.

Boas, Taylor C. 2015. Pastors for Pinochet: Authoritarian Stereotypes and Voting for Evangelicals in Chile. *Journal of Experimental Political Science* 00: 1-9.

Boas, Taylor Chase. 2014. Pastor Paulo vs. Doctor Carlos: Professional Titles as Voting Heuristics in Brazil. *Journal of Politics in Latin America* 6: 39–72.

Dolan, Kathleen. 2010. The Impact of Gender Stereotyped Evaluations on Support for Women Candidates. *Political Behavior* 32: 69-88.

———. 2014. *When Does Gender Matter?: Women Candidates and Gender Stereotypes in American Elections*. New York: Oxford University Press.

Dolan, Kathy. 2004. *Voting For Women: How The Public Evaluates Women Candidates*. Boulder, Colo: Westview Press.

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: 423-429.

Esarey, Justin, and Gina Chirillo. 2013. "Fairer Sex" or Purity Myth? Corruption, Gender, and Institutional Context. *Politics & Gender* 9 (04): 361-389.

Esarey, Justin, and Leslie Schwindt-Bayer. 2016. Women's Representation, Accountability, and Corruption in Democracies. *British Journal of Political Science* Forthcoming.

Fox, Richard L., and Eric R.A.N. Smith. 1998. The Role of Candidate Sex in Voter Decision-Making. *Political Psychology* 19: 405-419.

Funk, Carolyn L. 1996. The impact of scandal on candidate evaluations: An experimental test of the role of candidate traits. *Political Behavior* 18: 1-24.

Huddy, Leonie, and Nayda Terkildsen. 1993. Gender stereotypes and the perception of male and female candidates. *American Journal of Political Science* 37 (1): 119-147.

Johnson, Joel W., and Jessica S. Wallack. 2005. Electoral Systems and the Personal Vote. http://polisci2.ucsd.edu/jwjohnson/espv.htm.

McDermott, Monika L. 1998. Race and gender cues in low-information elections. *Political Research Quarterly* 51 (4): 895-919.

Murray, Rainbow, ed. 2010. *Cracking the Highest Glass Ceiling: A Global Comparison of Women's Campaigns for Executive Office*. Santa Barbara, CA: Praeger.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Samuels, David, and Cesar Zucco. 2014. The Power of Partisanship in Brazil: Evidence from Survey Experiments. *American Journal of Political Science* 58: 212-225.

Sanbonmatsu, Kira. 2002. Gender Stereotypes and Vote Choice. *American Journal of Political Science* 46 (1): 20.

Schwindt-Bayer, Leslie A. 2016. *Does the Presence of Women in Politics Reduce Corruption in Latin America?* Issue Brief, 7/29/16. Houston, TX: James A. Baker III Institute for Public Policy at Rice University, Latin America Initiative. http://www.bakerinstitute.org/media/files/files/226840a1/BI-Brief-072916-LAI_Corruption.pdf.

Schwindt-Bayer, Leslie A., and Catherine Reyes-Housholder. 2017. Citizen Responses to Female Executives: Is it Sex, Novelty, or Both? *Politics, Groups and Identities* 5 (3): 373-398.

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

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

Winters, Matthew S., and Rebecca Weitz-Shapiro. 2013. Lacking Information or Condoning Corruption: When Do Voters Support Corrupt Politicians? *Comparative Politics* 45 (4): 418-436.

Żemojtel-Piotrowska, Magdalena Anna, Alison Marganski, Tomasz Baran, and Jaros\law Piotrowski. 2016. Corruption and Sexual Scandal: The Importance of Politician Gender. *Anales de Psicología/Annals of Psychology* 33: 133–141.

Appendix

Table A.1: Multivariate logit models for all respondents in the United States

	Dependent variable:			
	corruptibility		vote support	
treatment: female governor	-0.389*	-0.359	-0.161	-0.299
	(0.209)	(0.295)	(0.251)	(0.342)
female subject	-0.135	-0.106	-0.333	-0.478
	(0.210)	(0.290)	(0.254)	(0.354)
female gov. X female subject		-0.061		0.302
		(0.418)		(0.507)
age	-0.017**	-0.017**	-0.015	-0.015
	(0.008)	(0.008)	(0.010)	(0.010)
education	-0.056	-0.056	0.022	0.022
	(0.070)	(0.070)	(0.083)	(0.083)
political interest	0.111	0.111	0.027	0.033
	(0.133)	(0.133)	(0.167)	(0.167)
race: white	-0.594	-0.598	1.134	1.167
	(0.744)	(0.744)	(0.838)	(0.840)
race: black	-0.648	-0.652	0.940	0.971
	(0.770)	(0.770)	(0.850)	(0.853)
race: asian	-1.329	-1.338	1.726	1.784*
	(1.024)	(1.026)	(1.075)	(1.082)
race: hispanic	-0.404	-0.409	0.459	0.495
	(0.784)	(0.784)	(0.878)	(0.880)
race: other race	-0.461	-0.470	-0.970	-0.920
	(0.794)	(0.796)	(1.175)	(1.177)
region: southeast	-0.151	-0.153	0.842**	0.850**
	(0.318)	(0.318)	(0.387)	(0.388)
region: midwest	-0.037	-0.038	0.393	0.399
	(0.278)	(0.278)	(0.365)	(0.365)
region: west	0.065	0.061	0.614	0.631
	(0.328)	(0.329)	(0.412)	(0.414)
constant	1.125	1.119	-2.135*	-2.127*
	(0.950)	(0.951)	(1.122)	(1.122)
Observations	406	406	406	406
Akaike Inf. Crit.	558.350	560.329	426.942	428.587
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.2: Multivariate logit models for all respondents in Brazil

	Dependent variable:			
	corruptibility			
	no history	history	no history	history
treatment: female governor	0.046	-0.259*	0.022	0.049
	(0.149)	(0.147)	(0.202)	(0.212)
female subject	-0.003	-0.129	-0.029	0.158
	(0.153)	(0.153)	(0.213)	(0.210)
female gov. X female subject			0.050	-0.586**
			(0.291)	(0.293)
age	-0.009*	-0.007	-0.009*	-0.007
	(0.006)	(0.006)	(0.006)	(0.006)
social class	-0.002	-0.032	-0.002	-0.034
	(0.064)	(0.062)	(0.064)	(0.062)
race: black	-0.324	0.031	-0.321	0.040
	(0.279)	(0.286)	(0.280)	(0.287)
race: parda	-0.212	-0.045	-0.212	-0.041
	(0.176)	(0.171)	(0.176)	(0.171)
race: indigenous	0.831	1.193	0.834	1.101
	(0.836)	(1.182)	(0.836)	(1.176)
race: yellow	0.282	0.361	0.280	0.294
	(0.504)	(0.496)	(0.504)	(0.499)
race: no response	0.276	0.353	0.278	0.374
	(0.611)	(0.658)	(0.611)	(0.661)
region: northeast	0.310	0.020	0.310	0.011
	(0.293)	(0.281)	(0.293)	(0.281)
region: southeast	0.215	0.128	0.214	0.138
	(0.294)	(0.279)	(0.294)	(0.279)
region: south	-0.023	0.004	-0.026	-0.006
	(0.307)	(0.305)	(0.308)	(0.305)
region: central west	0.289	-0.004	0.287	-0.017
	(0.313)	(0.314)	(0.313)	(0.315)
political interest	-0.234***	-0.137*	-0.234***	-0.136*
	(0.080)	(0.079)	(0.080)	(0.079)
correct manipulation responses	-0.446***	-0.166	-0.447***	-0.147
	(0.168)	(0.149)	(0.168)	(0.150)
thinking of state or politician	0.332**	0.278*	0.331**	0.283*
	(0.151)	(0.149)	(0.151)	(0.149)
constant	1.142**	0.854*	1.156**	0.693
	(0.516)	(0.511)	(0.522)	(0.517)
Observations	800	781	800	781
Akaike Inf. Crit.	1,111.667	1,101.587	1,113.637	1,099.585
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.3: Multivariate logit models for respondents in Brazil who answered both manipulation checks correctly

	Dependent variable:			
	corruptibility			
	no history	history	no history	history
treatment: female governor	-0.138	-0.091	-0.034	0.020
	(0.174)	(0.196)	(0.239)	(0.287)
female subject	-0.031	-0.119	0.091	-0.015
	(0.180)	(0.204)	(0.263)	(0.284)
female gov. X female subject			-0.219	-0.206
			(0.345)	(0.390)
age	-0.009	-0.009	-0.009	-0.009
	(0.007)	(0.007)	(0.007)	(0.007)
social class	-0.019	-0.071	-0.019	-0.074
	(0.075)	(0.082)	(0.075)	(0.083)
race: black	-0.555	-0.123	-0.560	-0.107
	(0.349)	(0.383)	(0.349)	(0.384)
race: parda	-0.354*	-0.192	-0.348*	-0.185
	(0.206)	(0.227)	(0.206)	(0.227)
race: indigenous	0.678	0.633	0.673	0.612
	(0.862)	(1.260)	(0.864)	(1.256)
race: yellow	0.428	-0.841	0.447	-0.832
	(0.593)	(0.902)	(0.594)	(0.903)
race: no response	-0.139	-0.183	-0.142	-0.196
	(0.699)	(1.032)	(0.702)	(1.032)
region: northeast	0.078	-0.323	0.068	-0.316
	(0.368)	(0.360)	(0.368)	(0.360)
region: southeast	-0.171	-0.364	-0.177	-0.357
	(0.363)	(0.353)	(0.363)	(0.353)
region: south	-0.239	-0.467	-0.235	-0.459
	(0.384)	(0.383)	(0.384)	(0.384)
region: central west	0.070	-0.290	0.072	-0.286
	(0.384)	(0.416)	(0.384)	(0.416)
political interest	-0.212**	-0.232**	-0.213**	-0.229**
	(0.097)	(0.109)	(0.097)	(0.110)
thinking of state or politician	0.507***	0.233	0.514***	0.241
	(0.178)	(0.195)	(0.179)	(0.196)
constant	1.025*	1.535**	0.974	1.466**
	(0.622)	(0.676)	(0.627)	(0.688)
Observations	580	449	580	449
Akaike Inf. Crit.	815.679	644.495	817.276	646.216
Note:			*p<0.1; **p<0.05; ***p<0.01	

Chapter Figures

Figure 1: Proportion of US respondents who think male and female governors are likely to engage in a corruption scandal while in office, by respondent sex (with 95% confidence intervals)

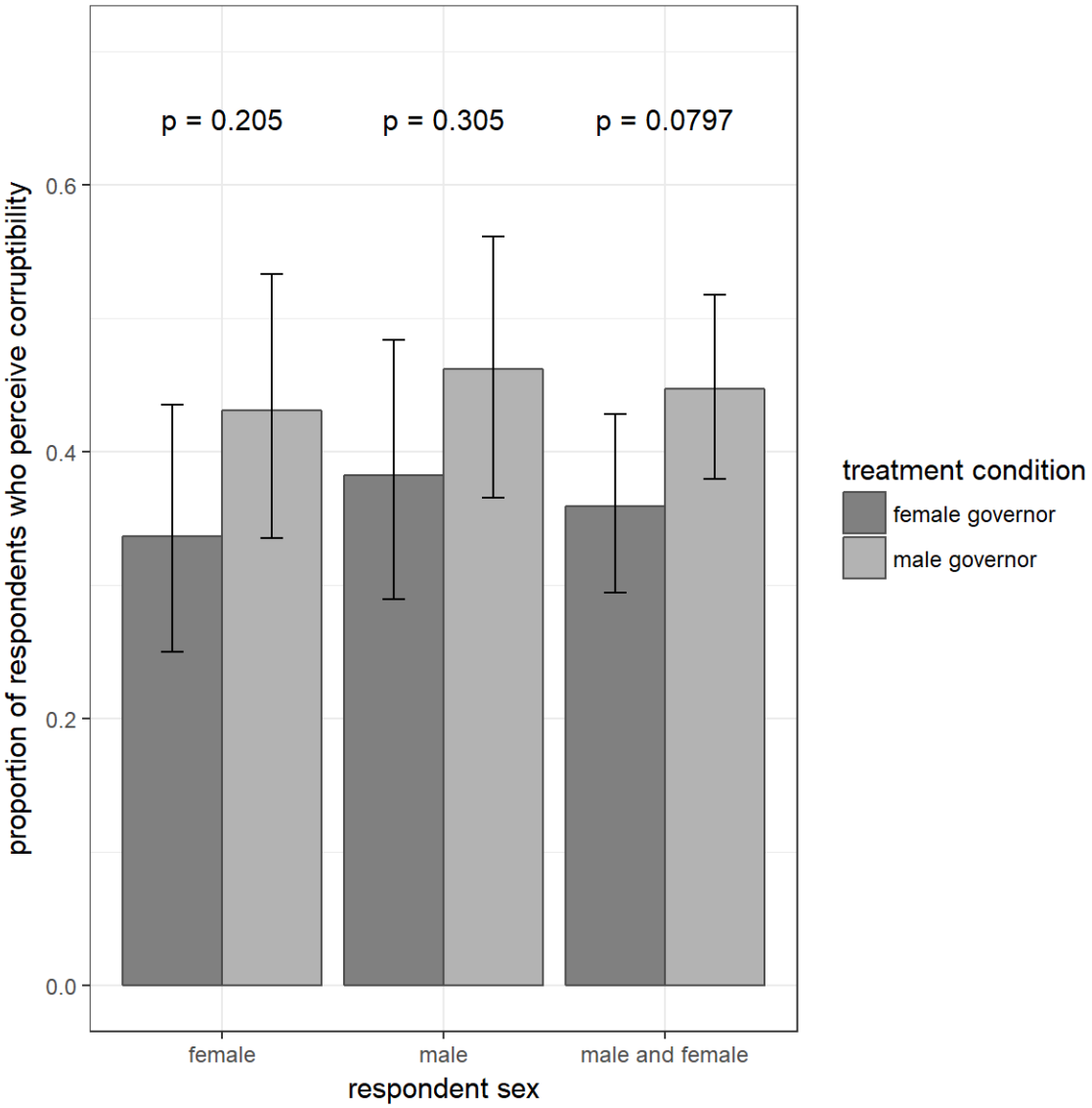

Figure 2: Proportion of US respondents who would vote for male and female governors after they have engaged in a corruption scandal while in office, by respondent sex (with 95% confidence intervals)

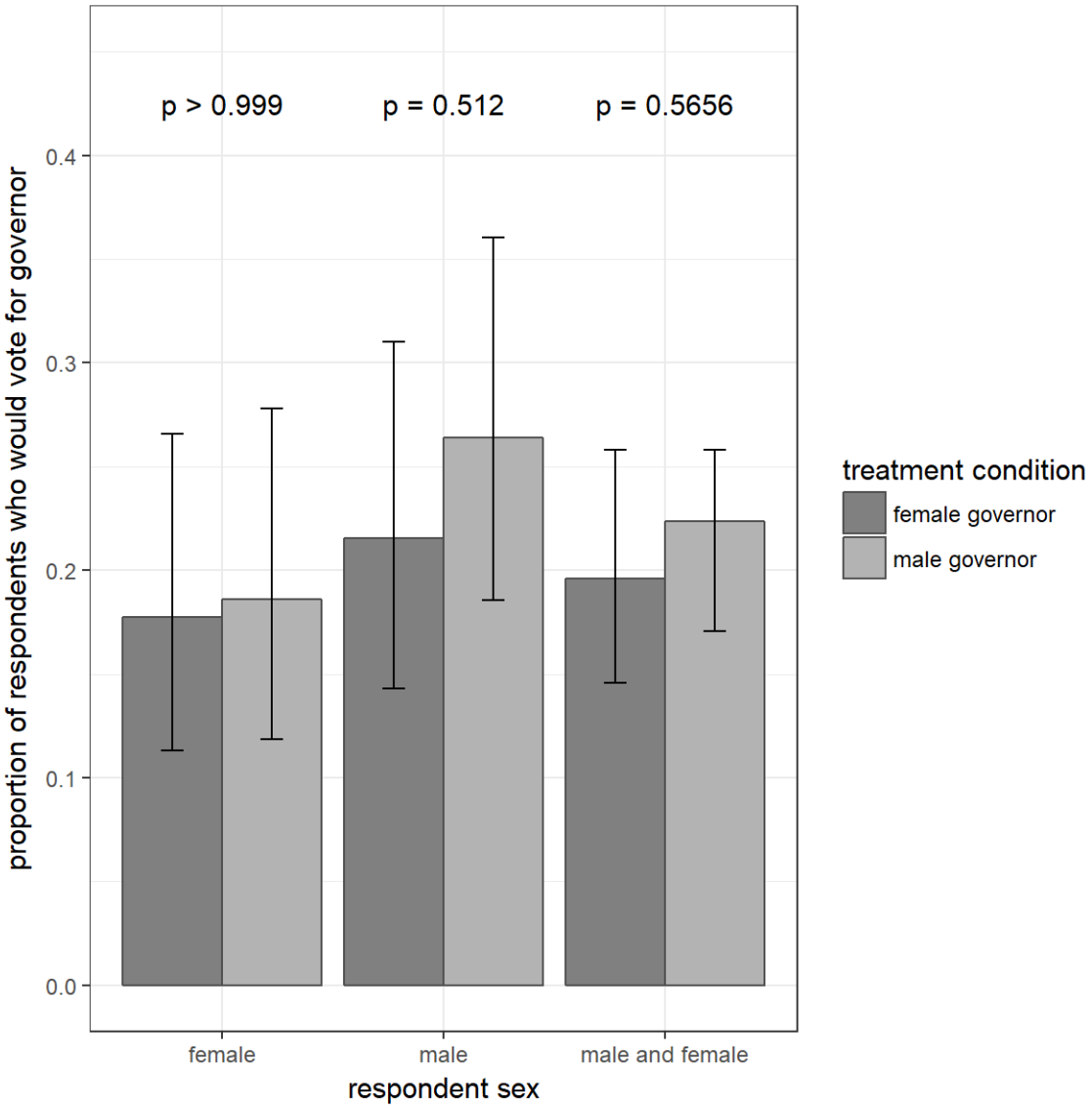

Figure 3: Proportion of Brazilian male and female respondents who think male and female governors in states with and without a history of women in politics will be engaged in a corruption scandal while in office (all respondents, with 95% confidence intervals)

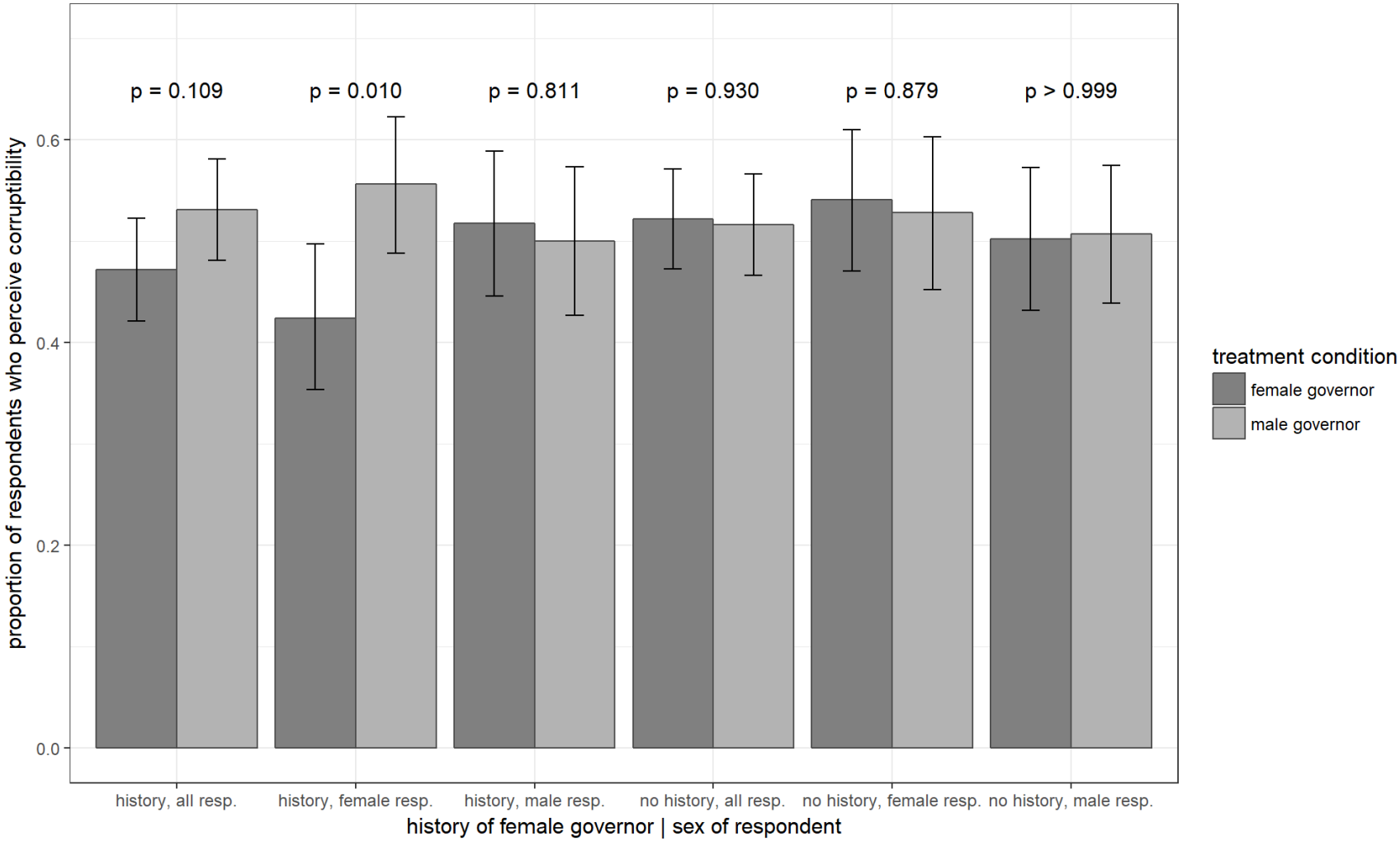

Figure 4: Proportion of male and female Brazilian respondents who think male and female governors in states with and without a history of women in politics will be engaged in a corruption scandal while in office (only respondents who answered both manipulation check questions correctly, with 95% confidence intervals)