

Ignoring Female Performance: A Survey Experiment on Policy Implementation in Argentina*

Gustavo Diaz (McMaster University)
Virginia Oliveros (Tulane University)
Rebecca Weitz-Shapiro (Brown University)
Matthew S. Winters (University of Illinois)

Revise and resubmit at the *British Journal of Political Science*

November 2023

Abstract

How does a politician's gender shape citizen responses to his or her performance in office? Existing literature suggests that voters are more likely to punish female politicians for malfeasance, a consequence of higher expectations of women politicians. Using an online survey experiment in Argentina, we randomly assign respondents to receive information about whether the distribution of a government food program in a hypothetical city is biased or unbiased and also randomly assign the gender of the mayor. Contrary to existing literature, we find that respondents are more responsive to performance information—both positive and negative—about male mayors. We explore whether this is a result of different baseline expectations or variation in inferences about program performance across male and female politicians but find little evidence for either explanation. Our results suggest the need for more research into how citizens process performance information in contexts with few female politicians.

Keywords: gender, voter attitudes, clientelism, survey experiment, Argentina

* This research was determined exempt by Institutional Review Boards at Brown University (Protocol ID 2010002827), Tulane University (Protocol ID 2020-1783), and the University of Illinois (Protocol Number 21332). Thanks to Jair Moreira and Anik Willig for excellent research assistance and to Juan Manuel del Mármol for his help editing the photo used in the survey experiment. Previous versions of this paper were presented at the 2022 Annual Meetings of the Midwest Political Science Association and the American Political Science Association. Thanks to Natalia Bueno and Cesi Cruz for excellent comments on previous drafts.

Female politicians are often perceived to be “better types” than their male counterparts: voters believe them to be more likely to act in the public interest and less likely to behave in a self-serving fashion (e.g., Alexander and Andersen 1993; Barnes and Beaulieu 2019). As a result, electorates often turn to women in moments of crisis (e.g., Piazza and Diaz 2020; Funk, Hinojosa, and Piscopo 2021). Positive beliefs about women’s competence and probity, however, have mixed implications for female politicians’ prospects once in office. As voters expect more from women, they are consequently more likely to punish them when they perform poorly. Recent literature provides evidence of this differential punishment of poor performance in response to corruption and other scandals (Esarey and Schwindt-Bayer 2018; Barnes, Beaulieu, and Saxton 2020).

The dominant explanation for differential performance evaluations is that citizens punish female politicians more harshly when they violate gendered expectations (Courtemanche and Connor Green 2020). Among the general public, multiple studies find that women are less tolerant of corruption than men (Swamy et al. 2001; Blake 2009; Torgler and Valev 2010), more likely to be concerned about corruption (Senters, Weitz-Shapiro, and Winters 2018), and less likely to be involved in corrupt behavior (e.g., Rivas 2013; Gingerich et al. 2016). Moreover, women appear to be more inclined to tell the truth (e.g., Dreber and Johannesson 2008) and more risk averse (e.g., Croson and Gneezy 2009). These attributes – combined with the fact that women typically face high barriers to entry into political life (Bledsoe and Herring 1990) – suggest that voters will believe female politicians to be, on average, better types than their male counterparts (Anzia and Berry 2011; Fulton 2012).¹

Existing scholarship pays less attention to the possibility of differential rewards for good behavior. To the extent that voters have higher expectations and standards for female politicians, we subsequently might expect them to provide greater rewards to male, as opposed to female, politicians for good performance (e.g., Costa 2021).

In this paper, we examine how citizens respond to information about performance for both male and female politicians. We run a survey experiment among an online sample of Argentine residents to examine the possibility of differential punishment and/or differential

¹ This belief may not be present in all contexts. In explaining equivalent punishment of corruption in a study in Brazil, Batista Pereira (2021) argues that President Dilma Rousseff’s association with the Lava Jato scandal weakened the belief that women are less corrupt than men.

rewards for male versus female politicians. We randomly assign respondents to receive information about a hypothetical mayor's performance in the distribution of a government food program, signaling good (bad) performance by describing the selection of beneficiaries as unbiased (biased). We also manipulate the gender of the mayor. Follow-up questions elicit respondents' vote intentions for the hypothetical mayor as well as evaluations of the mayor and the food program.

We find evidence of both differential punishment for poor performance (with male politicians punished more, contrary to extant scholarship) and differential rewards for good performance (with male politicians rewarded more, which follows more logically from the literature). We explore whether these differences can be attributed to variation in baseline expectations or differing interpretations of program performance across genders and find little evidence for either explanation. Our results suggest the need for more research into baseline beliefs about male versus female politicians and how citizens interpret performance information across genders in settings where female politicians are rare.

Research Design

Context

We carry out our empirical research in Argentina, which we consider a mixed case with respect to female political representation. On the one hand, it was the first country in the world to introduce a legislative gender candidate quota in national elections (1991),² female representation in the national legislature is high, and the country had a female president from 2007 to 2015. At the same time, as a federal country with over 2000 municipalities, only about 13 percent of the mayors in the country are female, suggesting that women still face significant barriers to political success.³

² See Caminotti (2014) for more details on the origins of this law.

³ <https://www.infobae.com/politica/2019/12/01/el-mapa-de-las-intendencias-son-minoria-gobiernan-al-9-de-la-poblacion-y-la-mayoria-gestiona-zonas-rurales/> (accessed 22 March 2022). In large part due to the gender quotas, the representation of women in national legislative positions is substantial: 40 percent of senators and 41 percent of

Vignette Experiment

Our experiment focuses on a hypothetical incumbent mayor's role in implementing a food distribution program, a common form of welfare. We describe that program as either distributed fairly, to those who really need it, or else as distributed in a biased fashion, wherein individuals with connections to the municipality are favored. We also include a control condition in which we provide no information about program implementation. We consider performance on this dimension to be a valence issue: evidence shows that most Argentines prefer unbiased distribution of social welfare benefits (Weitz-Shapiro 2014).

For each respondent, we randomized the characteristics of program implementation (biased, unbiased, or not mentioned) and the gender of the mayor (male or female).⁴ We used simple randomization, such that each respondent had an equal probability of being assigned to any of the treatment conditions.⁵ The vignette was available to respondents on a series of screens with follow-up questions. Respondents were thus repeatedly exposed to the treatment to which they were assigned.

We recruited N=1,802 respondents from Netquest's online panel in Argentina. The sample was designed to closely mirror the composition of the national population in terms of gender, age, region, and socioeconomic status.⁶ We implemented the survey in Qualtrics; respondents could complete it on a computer, tablet, or smartphone. Table 1 summarizes assignment to the relevant manipulations.

deputies in 2019-21. Source: Ministerio del Interior, Argentina (<https://www.argentina.gob.ar/interior/observatorioelectoral/analisis/mujeres> (accessed 26 March 2022)).

⁴ The online appendix discusses other manipulations and shows the full vignette text.

⁵ To examine the possibility of imbalance in the randomization, we use a multinomial logit model with the six possible treatment combinations as the outcome variable and age, education, gender, and social class as the explanatory variables. We find no evidence that the included covariates predict treatment assignment.

⁶ Online Appendix Table A1 reports sample characteristics.

Table 1. Vignette Experiment Research Design

Mayor Gender	Program Implementation		
	Biased	No Info	Unbiased
Male	342	352	355
Female	321	328	342

Note: Each cell shows the number of respondents assigned to that combination of treatments.

Respondents Take Note of Female Mayors

Before turning to the results, we provide evidence of respondent attentiveness to the mayor's gender. All respondents learn about the mayor's gender in the text of the vignette. The first sentence of the vignette asks respondents to imagine a mayor: "*Imagine **un** intendente*" if they are assigned to a vignette about a male mayor or "*Imagine **una** intendenta*" if they are assigned to a vignette about a female mayor.⁷

As a manipulation check immediately after measuring the outcome variables, we asked respondents whether they remembered the gender of the mayor who was mentioned, with possible responses of male, female, "that information was not provided," or "I don't recall." Information about a female mayor was substantitally more noteworthy to respondents.⁸ Among respondents assigned to a vignette with a male mayor, 51 percent replied that they read about a male mayor, 39 percent say that they were not given this information, and 9 percent of respondents said that they could not recall. Only 1 percent of respondents incorrectly recalled a female mayor. For respondents assigned to a vignette with a female mayor, 91 percent correctly reported learning about a female mayor. Four percent incorrectly recalled a male mayor, 2.4 percent reported not receiving this information, and only 3 percent say that they did not know.⁹

These results establish that respondents were attentive to the gender of the mayor in the vignette. This gives us confidence that respondents assigned to the female-mayor condition were

⁷ The male and gender-neutral articles are the same in Spanish; thus some respondents may not infer gender based on the use of the masculine form. The gender of the mayor was further reinforced for some respondents by an additional treatment that manipulated the inclusion of a picture with the mayor's full name.

⁸ See Online Appendix Table C4.

⁹ Online Appendix Table C5 shows that these results do not depend on the inclusion of a picture with the mayor's name.

thinking about a female mayor when they answered the outcome questions. It is also a striking descriptive finding about how notable Argentines find female mayors.

Results

We asked respondents two questions to assess the electoral impact of the information in the vignette. The first asked the respondent their own likelihood of voting for the hypothetical mayor in the next election. The second asked whether the respondent believed that the food program would help the mayor secure reelection.¹⁰ Figure 1 shows the results for these two outcomes. The left panels show the mean response for each combination of mayor gender and program implementation; the right panels show the differences in means across conditions.¹¹

We first compare vote intention among respondents who learned about biased implementation to vote intention among respondents in the control group who did not receive any information about implementation. Among respondents who learned about a male mayor, information about biased implementation reduces vote intention by 0.12 points ($p = 0.08$), compared to a reduction of only 0.04 points for female mayors ($p = 0.58$). The punishment for male mayors is three times as great, although the difference between these point estimates is not statistically significant ($p = 0.43$). This is contrary to the literature's dominant finding that voters punish women more harshly for poor performance.

We next examine the effects of information about unbiased implementation. When compared to the control condition, male mayors described as implementing the program in an unbiased fashion receive a 0.21-point increase in vote intention on the four-point scale ($p < 0.01$). The difference for female mayors is a 0.06-point increase ($p = 0.45$). The difference between these two estimates is substantively large (men receive rewards almost four times as large as women), although again not statistically significant at conventional levels ($p = 0.13$).

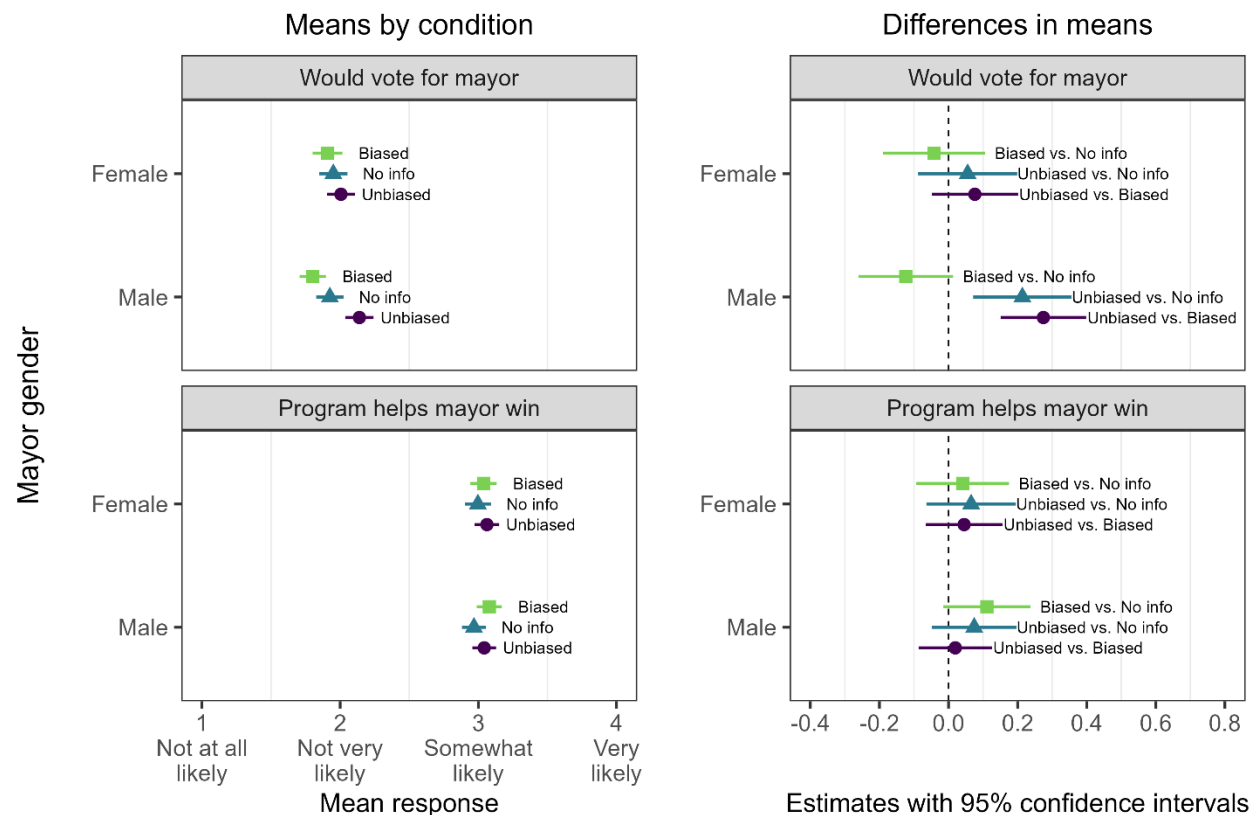
We also examine the difference between the two treatment conditions – unbiased implementation versus biased implementation. Relative to the biased implementation condition, unbiased implementation is associated with a 0.34-point difference in the likelihood that the

¹⁰ Unless otherwise noted, all outcomes of interest are coded on a scale ranging from 1 (not at all likely) to 4 (very likely).

¹¹ Section B of the online appendix shows full tables for the findings described in this paper.

respondent will vote for the male mayor ($p < 0.01$) but with only a 0.10-point difference that a respondent will vote for the female mayor ($p = 0.19$). This 0.24-point difference-in-differences is statistically significant at the 95-percent confidence level, indicating that survey respondents were more responsive to performance information provided about male mayors.

Figure 1. Means by treatment condition and differences in means for electoral performance outcomes



Turning to the second question, which asks respondents whether the food program will help the mayor win reelection, the different treatment conditions elicit more limited differences in patterns of responses. For both male and female mayors, *either* biased or unbiased implementation increases respondents' perceptions that the program will be electorally valuable compared to the control. For male mayors, the 0.11-point increase in the biased-implementation condition is marginally significant ($p = 0.08$), whereas the 0.07-point increase in the unbiased-implementation condition is not ($p = 0.23$). For female mayors, the perceived benefit of the

program is small in both conditions; the differences are not statistically significant, nor are they distinguishable from the differences observed among male mayors. The pattern of responses suggests that voters think biased implementation may be electorally valuable, even if they themselves react negatively to it.

It is possible that the existence of differential punishment and/or rewards will vary by the gender of the citizen assessing performance (see, for instance, Costa and Schaffner 2018; Schwarz and Coppock 2022). If women have even higher expectations for female politicians than men do, women could be harsher when punishing female politicians for not meeting expectations and less likely to reward them for good performance. Thus, one could expect both differential punishment and rewards to be stronger among female respondents. We find no evidence for this in our data.¹²

Do Differential Reactions Originate in Differential Baseline Expectations?

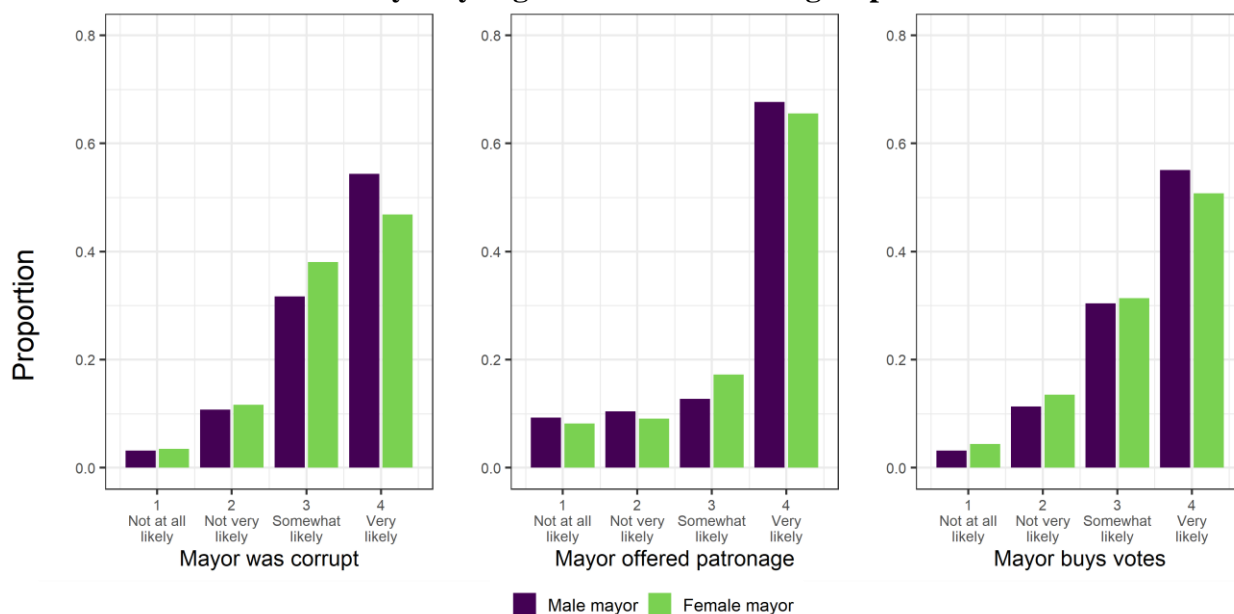
The section above establishes that respondents react more strongly to performance information for male mayors than for female mayors. We find evidence both of differential punishment—male mayors are punished more for biased implementation—and of differential rewards—with male mayors rewarded more for unbiased implementation. Is this a result of different baseline preferences for mayors of different genders? Our evidence suggests not. In the control condition, respondents are equally likely to say that they would vote to reelect the male or female mayor (1.93 versus 1.95; $p = 0.73$) and equally likely to say that the program will help the mayor win reelection (2.97 vs. 3.00; $p = 0.66$). Furthermore, if there were different baseline expectations across genders, this should lead to either greater rewards or punishment for men, but not both.

Other questions from the survey also suggest that Argentine respondents view male and female mayors similarly at baseline. After reading the vignette, respondents assessed whether the hypothetical mayor was likely to have engaged in corruption, patronage, or vote buying. Figure 2 shows the distribution of responses across gender treatments for respondents in the control group. Although respondents seem somewhat more likely to say that it is “very likely” that male

¹² See Table C2 in the online appendix.

mayors are corrupt, use patronage, or engage in vote buying, both difference-in-means tests and chi-squared tests of differences across the whole distributions return insignificant results: in the absence of a treatment, most respondents believe that both male and female mayors are similarly likely to engage in illicit behaviors.

Figure 2. Perceptions of corruption, patronage, and vote buying by mayor gender in the control group



Note: p-values from χ^2 tests: $p = 0.27$ for the corruption outcome; $p = 0.40$ for the patronage outcome; $p = 0.60$ for the vote buying outcome

Does Performance Information Differentially Change Program Perceptions?

Another possible explanation for our findings is that the information about program implementation might lead respondents to update their perceptions of the described social welfare program in different ways for male and female mayors. To examine this possibility, we explore questions that ask respondents if they would be satisfied with the program if it were run in their municipality and whether they believe the program was distributed to those most in need.

Figure 3 shows results for these outcomes. The top panels provide some evidence of differences across genders.¹³ As compared to the control condition, biased implementation

¹³ For this variable only, there is a more notable – but still not statistically significant – difference between male and female mayors in the control condition. When asked how satisfied they would be with the program if it were

decreases respondent satisfaction by about 0.13 points for both male and female mayors ($p = 0.09$ for male; $p = 0.07$ for female). In contrast, information about unbiased implementation is somewhat more meaningful for male mayors. Compared to the control condition, unbiased implementation by male mayors leads to a 0.27-point increase, on average, in program satisfaction ($p < 0.01$). This difference is more than twice the size of the difference observed for female mayors (0.13, $p = 0.11$), although the two differences are again not statistically distinguishable from one another ($p = 0.21$). We also see that the comparisons across the unbiased and biased implementation conditions are significant for both male and female mayors; the difference is larger for male mayors but not statistically distinguishable from the difference estimated for female mayors ($p = 0.25$).

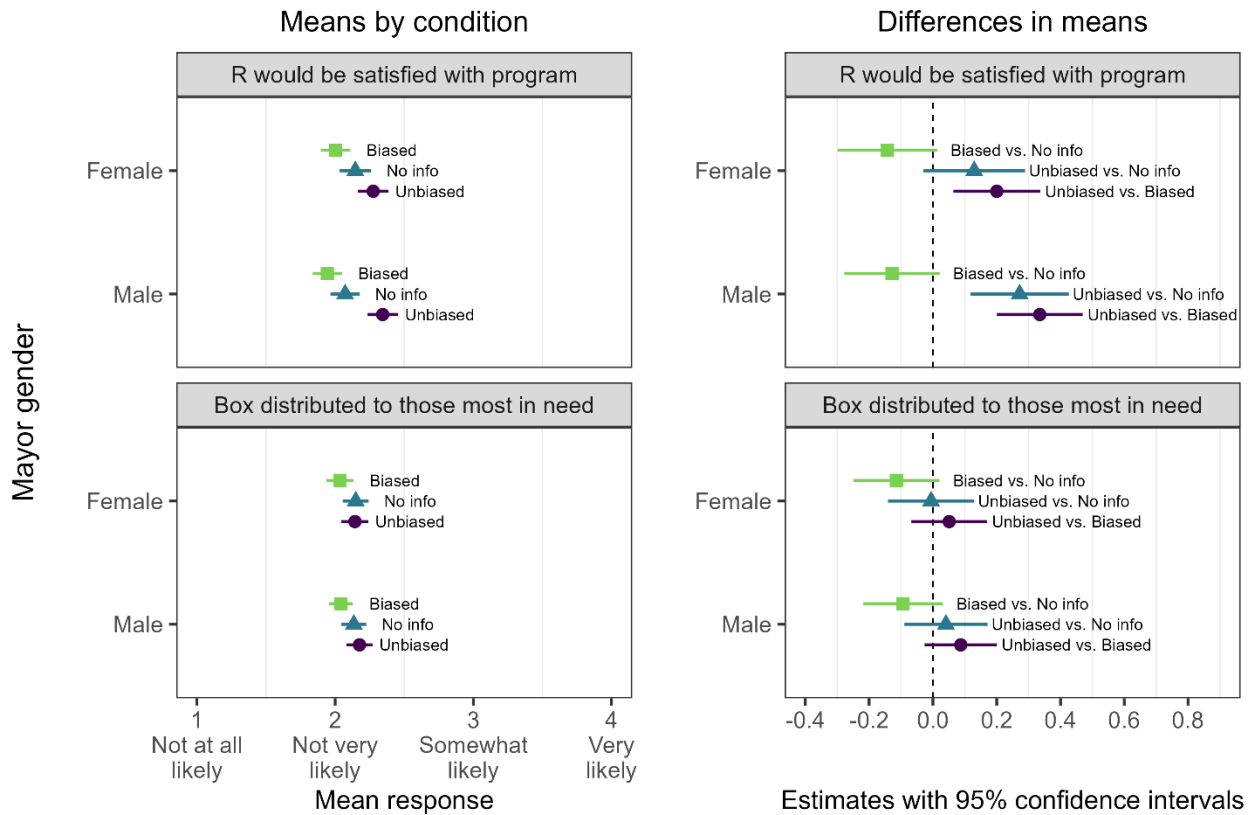
While this pattern might provide some insight into why male mayors benefit more from information about unbiased implementation, it does not help us understand why male mayors also appear to be punished more for biased implementation.

In the bottom panels, we show that information about biased implementation reduces perceptions that the program will benefit those most in need for both male and female mayors. The effect for female mayors is slightly larger and is significant at the 90-percent confidence level. The effects of this treatment across genders, however, are not distinguishable from one another. Information about unbiased implementation, on the other hand, appears to have no effect on beliefs about whether the program benefits those in need for either gender mayor.

Finally, we explore whether implementation information changes perceptions of other mayor characteristics. Online Appendix Table C3 shows no evidence of an effect of performance information on whether the respondent thinks the mayor is likely to have engaged in corruption, patronage, or vote buying.

implemented in their municipality, respondents confer a slight advantage on female mayors (2.15 versus 2.07, $p = 0.35$).

Figure 3. Means by treatment condition and differences in means for program satisfaction outcomes



Discussion

Existing literature establishes that the public tends to hold female politicians in higher regard and that this can imply more punishment if a female politician fails to deliver in office or becomes embroiled in scandal. Women seem to pay a higher cost for inappropriate behavior than men. Less evidence has been collected on whether male politicians receive greater rewards for good performance in office when compared to their female counterparts.

We study how Argentines react to information about social welfare program implementation when the mayor in charge of the program is either female or male. In contrast to the existing literature, we find that male mayors are punished more strongly for poor implementation. We also find that they are rewarded more strongly for good implementation. Although the cross-gender differences are not significant for either individual treatment, the total

difference between the two treatments is significantly greater for male mayors than for female mayors.

We show that these differences do not originate in different baseline expectations in the control condition – either on the outcome variables of interest or on other variables. In other words, our findings of differential rewards and punishments do not result from a gap between performance and expectations. While positive implementation information may affect perceptions of the program for male mayors, this provides only a partial explanation for the overall findings.

One possibility is that, given the relative rarity of female mayors, respondents focus on their gender identity and are thus less responsive to performance information compared to when the mayor is male. Our findings that respondents have very strong recall of female mayors, that there are no differences across male and female mayors in the control condition, and that there is positive updating about the program in the unbiased implementation condition for males only is consistent with that scenario. Future research designs might make use of pre- and post-treatment question batteries to explore this possibility and might also compare responsiveness to performance information across genders in contexts with varied shares of female elected politicians.

References

- Alexander, Deborah, and Kristi Andersen. 1993. "Gender as a Factor in the Attribution of Leadership Traits." *Political Research Quarterly* 46(3): 527–45.
- Anzia, Sarah F., and Christopher R. Berry. 2011. "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?" *American Journal of Political Science* 55(3): 478–93.
- Barnes, Tiffany D., and Emily Beaulieu. 2019. "Women Politicians, Institutions, and Perceptions of Corruption." *Comparative Political Studies* 52(1): 134–67.
- Barnes, Tiffany D., Emily Beaulieu, and Gregory W. Saxton. 2020. "Sex and Corruption: How Sexism Shapes Voters' Responses to Scandal." *Politics, Groups, and Identities* 8(1): 103–21.
- Batista Pereira, Frederico. 2021. "Do Female Politicians Face Stronger Backlash for Corruption Allegations? Evidence from Survey-Experiments in Brazil and Mexico." *Political Behavior* 43(4): 1561–80.
- Blake, Charles H. 2009. "Public Attitudes toward Corruption." In *Corruption and Democracy in Latin America*, eds. Charles H. Blake and Stephen D. Morris. Pittsburgh: University of Pittsburgh Press, 103–16.
- Bledsoe, Timothy, and Mary Herring. 1990. "Victims of Circumstances: Women in Pursuit of Political Office." *American Political Science Review* 84(1): 213–23.
- Caminotti, Mariana. 2014. "Ideas, Legados y Estrategias Políticas En La Reforma de Las de Selección de Candidatos: La Ley de Cuotas Pionera de Argentina." *Revista Uruguaya de Ciencia Política* 23(spe): 65–85.
- Costa, Mia. 2021. "He Said, She Said: The Gender Double Bind in Legislator–Constituent Communication." *Politics & Gender* 17(4): 528–51.
- Costa, Mia, and Brian F. Schaffner. 2018. "How Gender Conditions the Way Citizens Evaluate and Engage with Their Representatives." *Political Research Quarterly* 71(1): 46–58.
- Courtemanche, Marie, and Joanne Connor Green. 2020. "A Fall from Grace: Women, Scandals, and Perceptions of Politicians." *Journal of Women, Politics & Policy* 41(2): 219–40.
- Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic literature* 47(2): 448–74.
- Dreber, Anna, and Magnus Johannesson. 2008. "Gender Differences in Deception." *Economics Letters* 99(1): 197–99.
- Esarey, Justin, and Leslie A. Schwindt-Bayer. 2018. "Women's Representation, Accountability and Corruption in Democracies." *British Journal of Political Science* 48(3): 659–90.

- Fulton, Sarah A. 2012. "Running Backwards and in High Heels: The Gendered Quality Gap and Incumbent Electoral Success." *Political Research Quarterly* 65(2): 303–14.
- Funk, Kendall D., Magda Hinojosa, and Jennifer M. Piscopo. 2021. "Women to the Rescue: The Gendered Effects of Public Discontent on Legislative Nominations in Latin America." *Party Politics* 27(3): 465–77.
- Gingerich, Daniel W., Virginia Oliveros, Ana Corbacho, and Mauricio Ruiz-Vega. 2016. "When to Protect? Using the Crosswise Model to Integrate Protected and Direct Responses in Surveys of Sensitive Behavior." *Political Analysis* 24(2): 132–56.
- Piazza, Kelly Senters, and Gustavo Diaz. 2020. "Light in the Midst of Chaos: COVID-19 and Female Political Representation." *World Development* 136: 105125.
- Rivas, M. Fernanda. 2013. "An Experiment on Corruption and Gender." *Bulletin of Economic Research* 65(1): 10–42.
- Schwarz, Susanne, and Alexander Coppock. 2022. "What Have We Learned about Gender from Candidate Choice Experiments? A Meta-Analysis of 67 Factorial Survey Experiments." *The Journal of Politics* 84(2): 000–000.
- Senters, Kelly, Rebecca Weitz-Shapiro, and Matthew S. Winters. 2018. "Continuity and Change in Public Attitudes toward Corruption." In *Routledge Handbook of Brazilian Politics*, ed. Barry Ames. New York: Routledge.
- Swamy, Anand, Stephen Knack, Young Lee, and Omar Azfar. 2001. "Gender and Corruption." *Journal of development economics* 64(1): 25–55.
- Torgler, Benno, and Neven T. Valev. 2010. "Gender and Public Attitudes toward Corruption and Tax Evasion." *Contemporary Economic Policy* 28(4): 554–68.
- Weitz-Shapiro, Rebecca. 2014. *Curbing Clientelism in Argentina: Politics, Poverty, and Social Policy*. New York: Cambridge University Press.