Do Legislative Gender Quotas Reduce Corruption?*

Justin Esarey† and Natalie Valdes

December 7, 2021

Abstract

Does increased participation by women in government cause decreased corruption in that government? We answer this question by investigating whether gender quotas for public office cause reduced corruption using time-series cross-section data from the international system. Gender quotas create an increase in women's representation that we leverage to make a difference-in-differences comparison of subsequent changes in corruption. This design also allows us to validate the use of gender quotas as an anti-corruption policy. We find that legislative gender quotas reduce corruption, but only where women have substantive influence on governance.

*Word Count: 9721. We gratefully acknowledge financial support for this project provided by the Wake Forest University Undergraduate Research and Creative Activities Center (URECA), the WFU Associate Provost for Research, and the WFU Department of Politics and International Affairs. We thank Tilda Källqvist for helpful comments and corrections on an earlier version of the manuscript.

†Corresponding author: justin@justinesarey.com. Department of Politics and International Affairs, Wake Forest University, 1834 Wake Forest Road, F.M. Kirby Hall 319, Winston-Salem, NC 27109-7568.

A consensus has emerged in the empirical literature: in at least some circumstances, greater participation by women in government can cause lower corruption in that government. However, policies that increase women's representation do not necessarily reduce corruption despite that causal link. Gender quotas may be ineffective at increasing women's representation, or might be limited to institutions without real political power. Quotas might be used to extend patronage to women loyalists in a corrupt regime. Thus, we do not know whether imposing gender quotas in a country's legislature will reduce corruption despite the fact that we know that an increased proportion of women in government can decrease corruption. This is an important gap in our knowledge for two reasons. First, gender quotas create the opportunity for a "smoking gun" test (Collier, 2011): if we find that corruption drops after quotas are implemented, we may be reasonably certain that increased women's representation causes less corruption in that government. Second, mandating women's representation provides an innovative and progressive approach to fighting corruption that would be attractive to governments *if we can show that it is effective*. Based on prior research, we expect that its effectiveness will vary depending on the political context in which it operates.

In this paper, we study whether and when gender quotas for the legislature reduce corruption. We look to past research to develop theoretical expectations. First, quotas should only reduce corruption if they are effective at raising women's representation, as there is no other plausible mechanism by which quotas would influence corruption. Second, quotas will only reduce corruption where women have meaningful access to power; if they do not have power, we would not expect them to be able to change the policies and institutions that enable corruption. Finally, we expect that only when government is accountable to voters will the women put into office by quotas have the incentive to fight corruption; however, if women's resistance to corruption is intrinsic or value-based, this will not be a moderator for the effect of quotas on corruption. We test these hypotheses with a difference-in-differences (DID) research design (with both country and year fixed effects) designed to identify causal relationships in international panel data. Difference-in-differences models have a potentially important advantage over other designs for our study; they look specifically for changes in corruption before and after quota adoption *within* a country, excluding a potential source of omitted variable bias from unit heterogeneity that may be present in designs that leverage between-country variance. We also use instrumental variable models to rule out the possibility that simultaneity (reverse causality) explains any relationship between implementing gender quotas and corruption.

Our analysis of over twenty-five years of observational data from 174 countries yields three findings. First, gender quotas do (on average) raise the representation of women in government to a substantively meaningful degree; this is a necessary precondition for quotas to reduce corruption and is consistent with results from earlier studies. Second, gender quotas for the legislature reduce corruption both in the legislature and in government as a whole, but these effects are only politically and statistically meaningful among states where women have substantive influence on governance. This makes sense if women are less likely to engage in corruption than men: their greater aversion to corruption can only change outcomes when their actions are politically meaningful. However, we do *not* find a stronger negative effect of gender quotas on corruption in countries with strong electoral democracy compared to those without. This is consistent with the idea that women's aversion to corruption is at least partially intrinsic, not fully attributable to women's greater risk aversion or differential treatment of women by voters that relate to their extrinsic incentives. Finally, we find no evidence for simultaneity in the relationship between quota adoption and corruption. We conclude that gender quotas for the legislature and similar initiatives for other branches of government can only be expected to reduce corruption when they are accompanied by efforts to provide those women with meaningful influence on policy making.

Theory development

There is little existing scholarship that directly studies whether gender quotas reduce government corruption. One paper that does so, Bjarnegård, Yoon and Zetterberg (2018), argues that:

> [I]f women elected through quotas are recruited from new networks and with no exposure to a corrupt political system, *and* they are given their own mandate to act on a range of issues once in parliament, then quotas may constitute a 'clean slate' and thus help reduce corruption. However, if the reform is designed in a manner that recruits women from already existing, corrupt networks, and the elected women are expected to protect an already corrupt party line, then quotas may just provide non-democratic regimes with yet another 'tool on the menu of manipulation' (p. 106).

They further describe (based on interviews and field work) how candidate recruitment procedures in Tanzania have made that country's gender quotas ineffective at reducing corruption by emphasizing party loyalty and discipline. These quotas simply reproduce the existing corruption networks that are already present in the political system.

Another study that examines the connection between quotas and corruption, Beaman et al. (2009), examines local governments in India that are randomly assigned to implement an electoral gender quota reserving leadership to women. They find that "on the average, individuals in currently reserved GPs [gram panchayats, or village councils] are less likely to have paid a bribe to obtain a BPL card [entitling the holder to government benefits] or drinking water connection. This is true for both GPs reserved for the first and second time" (p. 1520).

Although we hesitate to draw firm conclusions from just two studies, their very different results lead us to believe that the impact of gender quotas on corruption will depend on where and how these quotas are implemented. First (and most obviously), these quotas can only reduce corruption where they are effective at increasing the representation of women: we have no reason to expect that quotas will reduce corruption by any mechanism other than increased women's representation. Second, women must have sufficient political independence to influence policy in order to be capable of changing corruption levels, as consistent with the arguments of Bjarnegård, Yoon and Zetterberg (2018). Third, gender quotas will only reduce corruption when the women in office are directly accountable to voters and not to government or party patrons—that is, in systems with free, fair, and competitive elections. Prior empirical work has already established that higher women's representation is only associated with lower corruption among states where politicians are strongly accountable to voters (Esarey and Chirillo, 2013; Esarey and Schwindt-Bayer, 2018).[1] Finally, we anticipate that simultaneity is possible in observational data: government with endemic corruption may adopt gender quotas to impress international and domestic audiences without disrupting networks of bribery and patronage (Krook, 2006). All of these arguments have roots in prior research concerning the link between gender and corruption, and we detail these links in this section.

Why might gender quotas lower corruption?

Scholars and policy-makers became interested in increasing women's participation in government as a strategy to fight corruption almost immediately after Dollar, Fisman and Gatti (2001) and Swamy et al. (2001) demonstrated a correlation between women's representation in the legislature and lower corruption in government in country-year panel data. The logic is straightforward: if governments with greater female representation have lower corruption, then perhaps boosting the number of women in government will cause reduced corruption. Several governments have even tried feminization as a corruption-fighting measure (Moore,

[1]See also Tavits (2007) and Schwindt-Bayer and Tavits (2016).

1999; McDermott, 1999; Karim, 2011; Kahn, 2013; Wills, 2015). The potential advantages of this strategy are obvious. Corruption is by its very nature difficult to observe, may not be considered unethical by its practitioners, and is infrequently or unevenly punished when endemic to a system; these features empower the interests that benefit from corruption to successfully resist reforms. By contrast, it is easy to observe whether there are more women participating in government. Many consider gender parity to be morally important on its own terms. These factors make it harder to subvert or oppose gender quotas compared to other anti-corruption programs, and in turn make it tempting to see quotas as a way of achieving quick and politically palatable reductions in corruption.

The strategy may be even more appealing in light of recent empirical research showing a *causal* (not just correlational) impact of increased women's representation in government on corruption. For example, Jha and Sarangi (2018) study a cross-section of countries worldwide using instrumental variables for women's representation in the legislature and find that greater representation of women causes lower corruption. Esarey and Schwindt-Bayer (2019) use a different set of instrumental variables to establish that an increased proportion of women in parliament causes reduced corruption in a panel data set of 76 democratic-leaning countries. Paweenawat (2018) does the same, but for Asian countries. Correa Martínez and Jetter (2016) instrument participation of women in the labor force and find that greater participation causes lower corruption. All of these papers use different instrumental variables for women's representation, and some (e.g. Esarey and Schwindt-Bayer, 2019) use multiple combinations of instrumental variables; this suggests that the overall finding is robust to the researcher's choice of instrument. Brollo and Troiano (2016) use a completely different approach, regression discontinuity design, to establish that female mayors in Brazil are less involved in corruption (measured as a part of randomly administered government audits) and more effective at providing public goods than their male counterparts, at least when elections are competitive.

We might hesitate to trust a "black box" finding that more women causes less corruption, but experiments and survey data provide behavioral microfoundations and mechanisms for the macro-level relationship that we observe. Examining data from the World Values Survey, Torgler and Valev (2010) find that women consistently report greater aversion to corruption and tax evasion compared to men worldwide. Why? Three explanations are frequently discussed in extant scholarship. First, a long and cross-disciplinary literature has consistently found that women are more averse to risk than men (Sundén and Surette, 1998; Byrnes, Miller and Schafer, 1999; Bernasek and Shwiff, 2001; Watson and McNaughton, 2007; Eckel and Grossman, 2008; Croson and Gneezy, 2009); relatedly, women may be more motivated by guilt, shame, and regret than men (Ward and King, 2018). Thus, where corruption is risky (i.e. subject to discovery and punishment) or stigmatized, women may be more reticent to take that risk in order to gain the reward. This explanation is supported by observational evidence that the gender-corruption linkage only exists in places where accountability for corruption is high (Esarey and Chirillo, 2013; Esarey and Schwindt-Bayer, 2018) and by experiments demonstrating that women are only less willing to engage in bribery than men when detection and punishment are possible (Schulze and Frank, 2003; Armantier and Boly, 2013). Risk aversion also appears to explain why voters expect women politicians to be less corrupt (Barnes and Beaulieu, 2014): once this factor is accounted for, gender differences disappear (Barnes and Beaulieu, 2019).

Second, women may be held to a higher standard when it comes to corruption compared to men. For example, Wagner et al. (2017) found that male police officers in Uganda were more lenient than women in evaluating and punishing fellow male police officers for corrupt activities, but equally strict when evaluating female police officers. Eggers, Vivyan and Wagner (2018) find that women voters in Britain more harshly punish female members of parliament involved in misconduct compared to male MPs involved in the same conduct; male voters treated MPs of both genders equally. However, Schwindt-Bayer, Esarey and

Schumacher (2018) found no tendency to more harshly judge women suspected of corruption among voters in Brazil or the United States in their survey data.

Finally, it is possible that women are less likely to engage in corruption (at least in part) because they are intrinsically less willing to participate in it. Dollar, Fisman and Gatti (2001) speculated that such a difference could be attributed to gender differences in socialization, biology, and/or cultural norms that translate into a greater innate resistance to corruption among women. Wängnerud (2020, p. 1) makes the similar argument that "an influx of women in elected assemblies is accompanied by an influx of empathic and other-regarding values" that leads to reduced corruption because "self-regarding values, rather than individual men, are replaced." If this is true, we would expect women to reduce corruption wherever they have power even if they are not accountable to voters.

Our overall interpretation of this evidence is that there are good reasons to suspect that an exogenous increase in women's representation in government, such as the increase caused by a gender quota, might reduce corruption in that government. Specifically, we expect that (for several possible reasons) women will be less willing to participate in corruption than equivalent men in the same position. However, we should expect this relationship to be highly variable across contexts. For example, an experiment by Alatas et al. (2009) finds that women are less willing to pay or accept bribes and more willing to punish them, but in Australia (and not in India, Indonesia, or Singapore). As another example, Le Foulon and Reyes-Housholder (2021) found that Uruguayan voters actually *preferred* women political candidates accused of corruption compared to equivalent men. In short, we expect that gender quotas might fail to lower corruption in many circumstances. The literature provides us guidance on where and when we would expect gender quotas to succeed at reducing corruption.

What conditions are required for gender quotas to successfully lower corruption?

Because we are studying gender quotas for the legislature, our argument focuses on what makes these sorts of quotas effective at reducing corruption. We argue that three conditions are required for legislative gender quotas to cause decreased corruption:

1. The quotas must increase the representation of women in the legislature;
2. the women who gain office through quotas must have meaningful and independent influence over policy; and
3. the women who gain office through quotas must be held accountable for corruption, e.g., by free and fair elections.

Our first criterion implies that the only effect that legislative gender quotas have on corruption is via increasing the representation of women; we know of no other way that the two could be causally connected. Fortunately, there is already evidence that gender quotas are effective at increasing women's representation. For example, in their empirical analysis of Sweden's gender quota, O'Brien and Rickne (2016) find that the quota caused an increase in women's political leadership. Schwindt-Bayer (2009, p. 21) finds that "quota size affects women's representation regardless of whether or not the quota includes placement mandates and enforcement mechanisms." We will verify that quotas increase women's representation in our study as well.

Second, gender quotas must also give women the agency to effect change in government. In Morocco, for example, seats in parliament are reserved for women. But parliament has no power: Moroccan parliamentary seats are more akin to the plums of patronage than to levers of influence. Loyalty to the monarchy (the source of real political power in Morocco) prevents members of parliament from acting independently (Sater, 2012). Similarly, although

Rwanda's parliament included 48.8 percent women in 2003, parliament is powerless to criticize the regime. High women's representation in such parliaments may legitimize the ruling party but is unable to create change (Longman, 2006; Burnet, 2012).

Even representation in a parliament with power may not result in policy influence for women. In Pakistan, for example, male legislators vote to select female legislators for the reserved seats. Because of this, "whenever female legislators took positions on issues of concern to women, their male colleagues reminded them that they had been elected by men and not women" (Krook, 2009, p.66). The example illustrates that women who are not directly responsible to a voting consistuency have a weaker base of legitimacy from which to make policy (Matland, 2006; Hassin, 2010). South Africa provides another example:

> ANC [African National Congress] governing elites used their electoral dominance, the PR [proportional representation] system, and the quota to undermine women's counterpublics and discipline female MPs [members of parliament]. ...An expansion of the ANC's voluntary quota to 50 percent has not resolved these problems. Instead, elites continue to herald the quota, claiming a commitment to participatory poltics and women's rights that no longer exists (Walsh, 2012, p. 130).

A similar situation exists in Tanzania (Bjarnegård, Yoon and Zetterberg, 2018). Under these circumstances, we would not expect a gender quota to produce changes in policy (including reductions in corruption) even though the legislature has real power.

Finally, gender quotas could simply enable corrupt officials to install female allies in government positions who are willing to participate in corrupt activity. The recent examples of Cristina Fernàndez de Kirchner in Argentina and Dilma Rousseff in Brazil show that women can be active participants in government corruption. Although prior scholarship has shown that networks of people involved in corrupt activities fear exposure by women outsiders (Bjarnegård, 2013; Grimes and Wängnerud, 2012; Stockemer, 2011; Sundström and Wängnerud, 2014), corruption networks may be able to help women who are insiders gain power. If so, a gender quota would not reduce corruption. This possibility leads to our third criterion: the women put into office by gender quotas must face accountability for corruption from the public. Following Tavits (2007), Esarey and Schwindt-Bayer (2018), and Schwindt-Bayer and Tavits (2016), we stipulate that this sort of accountability requires that voters must be able to identify and punish corrupt politicians at the ballot box. Voters' ability to link individual politicians to corrupt acts may vary according to the structure of the political system, but a foundation of electoral democracy is necessary for this link to matter. On the other hand, if women are more *intrinsically* resistant to corruption we will not observe a stronger link between quotas and reduced corruption in electoral democracies compared to other systems.

Why might corruption cause gender quotas to be implemented?

Studying the effect of gender quotas on corruption is difficult because governments, and especially corrupt governments, may adopt quotas specifically to answer criticisms emanating from foreign governments or international and domestic NGOs (Krook, 2006; Bush, 2011; Hughes, Krook and Paxton, 2015). Governments receiving foreign assistance are especially susceptible to such pressure. That pressure is a potential source of simultaneity in the causal process: our dependent variable (corruption) can cause our independent variable (gender quotas) because corruption generates demands for reform. Even worse, and as described in the previous section, quotas may be implemented with no intent of genuinely changing the structure of power and patronage relationships. Of four theoretical explanations for gender quota adoption that Krook (2006) discusses, at least two[2] are suggestive of simultaneity between corruption and gender quotas: "political elites recognize strategic advantages for

[2]The other two explanations for quotas from Krook (2006), "women mobilize for the adoption of quotas to increase women's representation" and "quotas are consistent with existing or emerging notions of equality and representation" (p. 307), are neutral with respect to corruption.

supporting quotas" and "quotas are supported by international norms and spread through transnational sharing" (p. 307; see also Krook, 2009).

There are many case studies of governments and parties that implement gender quotas as a mechanism of patronage or to crowd out political opposition without conferring real power to women. For example, gender quotas imposed by the ruling party in Senegal in the 1980s were "motivated primarily by competition between men... for control of the ruling party" where a new leader "sought to create new clients who would be dependent upon his political largesse in order to detract from the power of the party 'barons'" (Beck, 2003, p. 156). Thus, gender quotas were implemented specifically as a means of creating patronage, a kind of corruption that facilitates corruption in other forms. In Rwanda, a case mentioned in the previous section, women's relatively subordinate position in Rwandan society may have made them more susceptible to pressure from the regime and therefore a favorable target for patronage. The government includes women to present a false front of legitimacy:

> One person told me: "The RPF [Rwandan Patriotic Front] focuses on diversity so that they can appear democratic even though they control all power. They put women in the National Assembly because they know they [the women] will not challenge them" (Longman, 2006, p. 148).

Countries like Senegal and Rwanda might be particularly prone to implement quotas precisely because their endemic corruption creates incentives to broaden and reinforce existing clientelistic politics.

Pressure from the international community can motivate the imposition of gender quotas, but these quotas could be superficial if women are socially, politically, and economically unable to take advantage of their positions (Liu and Dionne, 2019). Bush (2011) finds "strong evidence that international incentives are positively and significantly related to a country's likelihood of adopting a gender quota" (p. 104); these incentives are foreign aid, support from the United Nations for post-conflict operations that supported political liberalization, and/or election observers. But in Afghanistan, where gender quotas were imposed as a part of the post-war reconstruction process (prior to the Taliban's 2021 second conquest of the country), "women's considerable presence in the parliament has not led to the substantive representation (or definition) of the interests of 'women in general' (Larson, 2012, p. 136)." Similarly, in Latin America, transnational organization and activism motivated the proposal and (in some cases) passage of gender quotas in some countries without necessarily increasing the political power of women (Htun, 2016, pp. 52-54).

Most concerningly, increased global pressure for quotas is *less* effective in countries with strong domestic ties to women's transnational organizations (Hughes, Krook and Paxton, 2015). One possible explanation for this paradoxical moderation effect is that male elites "see quotas as a challenge to their power and position" (p. 359) and are more threatened by quotas when domestic women's interests groups are stronger and more organized. That is, quotas may be less likely to be implemented precisely where they are more likely to be effective at reducing corruption because extant (male) elites are most harmed by them in those circumstances.

Theoretical expectations

To summarize our expectations based on the theory presented above, we anticipate gender quotas are most effective at causing lower corruption when:

1. they successfully raise women's representation in government;
2. women have independent influence in policy-making; and
3. female politicians are accountable to voters (unless resistance to corruption is primarily value-based or otherwise intrinsic).

However, we believe it is possible that corruption *also* causes a greater likelihood of adopting a gender quota intended to placate internal and external pressure groups, especially in authoritarian or clientelistic governments and/or in countries most susceptible to pressure from foreign governments and NGOs. The remainder of our paper examines the empirical support for these expectations in observational panel data from the international system over the last quarter-century.

Data

Our empirical analysis relies on data from version 9 of the Varieties of Democracy (V-Dem) country-year data set (Coppedge et al., 2019; Pemstein et al., 2019) with some additional data from the 2019 edition of the Quality of Government time-series cross-sectional data set (Teorell et al., 2019) and the 2019 release of the International Country Risk Guide (Political Risk Services Group, 2018). Summary statistics for our data are presented in Table 1.[3]

The two most important measures for our study are (a) our measure of gender quotas, and (b) our measures of corruption. The existence of a gender quota in the lower (or only) chamber of the legislature (including reserved seats or statutory quotas, but excluding voluntary party quotas) is available in the V-Dem data set and sourced from the QAROT data (Coppedge et al. 2019, p. 144, Hughes et al. 2019). Our data set also includes the proportion of women in this chamber of the legislature as compiled by the V-Dem authors using multiple sources.

Our primary measure of corruption is the V-Dem legislative corruption index. This index uses evaluations from multiple country experts to determine whether "members of the legislature abuse their position for financial gain," including bribery, nepotism, and forms of graft (Coppedge et al., 2019, pp. 134-135). These ratings are then converted to a continuous measure (with a range of about -3.3 to 3.3) using an item response model (Pemstein et al., 2019). We recoded the measure so that larger values indicate more corruption. This measure

[3]All tables in this paper and the online appendix were produced using `estout` (Jann, 2005, 2007).

Table 1: Summary statistics

	N	mean	sd	min	max
V-Dem Legislative Corruption	4458	0.225	1.340	-3.322	3.265
V-Dem Public Sector Corruption	4609	0.509	0.302	0.004	0.979
V-Dem Executive Corruption	4609	0.507	0.298	0.011	0.978
V-Dem Judicial Corruption	4605	0.042	1.533	-3.454	3.400
V-Dem Overall Corruption	4596	0.533	0.300	0.006	0.976
Bayesian Corruption Index	4237	48.462	15.778	6.450	74.963
ICRG Corruption Risk	3581	3.184	1.276	0.000	6.000
% Women in Parliament	4411	15.695	11.022	0.000	63.800
Gender Quota in Legislature	4618	0.210	0.408	0.000	1.000
V-Dem Gender Power Index	4618	0.872	1.108	-2.854	3.876
log Mean GDP PC, 2010 USD	4529	8.293	1.520	5.098	11.453
V-Dem Electoral Democracy Index (Polyarchy)	4616	0.521	0.267	0.014	0.948

Data are present for 174 countries between 1992 and 2018. Panels are unbalanced due to missing data.

is particularly suitable because it measures corruption specifically *among legislators*, and is therefore a strong match for a study of how legislative gender quotas influence corruption in government. The measure is also available for a very large number of countries and time periods.

Prior empirical studies have examined the connection between women's representation in the legislature and *overall* government corruption, not just corruption in the legislature. Accordingly, we also consider the effect of gender quotas on other measures of corruption from the V-Dem project. V-Dem measures corruption in the legislature, the public sector generally, among high-level members of the executive branch, in the judiciary, and in government overall. We use each of these measures as dependent variables in our analysis. While all are correlated with one another, as shown in Appendix Figure 5, they are distinct. All measures are (re)coded so that larger values indicate more corruption.

To ensure that our results are not overly sensitive to the V-Dem measurement methodology, we also examine the Bayesian Corruption Index (BCI) and the corruption measure from the International Country Risk Guide (ICRG). The BCI uses a latent variable measurement methodology (Standaert, 2015) based on that of the World Bank's World Governance Indicators (World Bank, 2016) to combine multiple measures of corruption into a single country-year measure on a 0-100 scale. The ICRG is produced by the Political Risk Services group and consists of a 0-6 expert rating of "actual or potential corruption in the form of excessive patronage, nepotism, job reservations, 'favor-for-favors,' secret party funding, and suspiciously close ties between politics and business" in a country during a particular year (Political Risk Services Group, 2018, pp. 4-5). As before, measures are (re)coded so that higher values indicate more corruption. As shown in Appendix Figure 6, these measures are correlated with V-Dem legislative corruption but are distinct from it.

The other variables used in our analysis are contextual factors that might influence the causal relationship between gender quotas and corruption. The V-Dem gender power index measures how "political power [is] distributed according to gender" (Coppedge et al., 2019, p. 191) using country expert ratings converted to a continuous scale by an item response model; larger values indicate a more equal distribution of power between men and women. The V-Dem electoral democracy score indicates the extent to which "the ideal of electoral democracy in its fullest sense" is achieved via free, fair, and competitive elections without restrictions on speech (Coppedge et al., 2019, p. 39); this score varies between 0 (least democratic) and 1 (most democratic). Finally, average per capita GDP in 2010 prices by country comes from the Quality of Government data set (Teorell et al., 2019) and the World Bank's World Development Indicators (World Bank, 2016).

Difference-in-difference analysis

We present difference-in-difference (DID) plots showing the effect of legislative quota imposition on (a) women's representation in government and (b) corruption. These plots are created using a panel linear model:

$$y_{it} = \beta_p \times q_i + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1}$$

where α and γ are vectors of fixed effects for country (indexed by i) and year (indexed by t) respectively.[4] q_i is a binary variable indicating whether a legislative gender quota was ever adopted by the country (= 1) or not (= 0). β is a vector of treatment effects measuring the difference-in-differences in the dependent variable y_{it} between quota adopters (treated countries, $q_i = 1$) and non-adopters (untreated countries, $q_i = 0$); β is indexed by the number of years p before or after the treated country adopts the quota. Each difference between treated and untreated countries at p years before quota adoption is compared to

[4]These models were estimated in R using the `plm` library (Croissant and Millo, 2008) and using `clubSandwich` to estimate cluster-robust standard errors (Pustejovsky, 2019).

the same difference at $p = -1$ (just before the quota goes into effect).[5] Thus, $\beta_{-1} = 0$, as there is no difference in differences in the dependent variable when comparing the baseline category to itself. Other values of β_p are defined relative to the difference at $p = -1$. ε_{it} is a country-year error term; standard error estimates are clustered on country.

Fixed effects are especially important for a DID analysis because we aim to study changes in the dependent variable (or DV) *within* units that adopt quotas, net of common time trends or long-term differences between countries. This represents a different methodology compared to many prior observational studies, some of which partially attributed between-country variance in corruption to differences in women's representation and therefore may have allowed a source of bias from unmodeled unit heterogeneity. However, using country fixed effects comes at the cost of more imprecision in the estimates (Clark and Linzer, 2015), so that both choices have potential disadvantages that are hard to weigh without knowing the data generating process. Thus, our design should be considered complementary to prior work rather than critical of it.

Valid identification of causal effects with a DID analysis requires that the trends in women's representation among countries that adopt a quota and those that do not would remain the same over time if quota adoption were held constant (Angrist and Krueger, 1999, pp. 1296-1299); this is sometimes called the "common trends" assumption (Angrist and Pischke, 2009, p. 230). For example, countries that eventually adopt a gender quota must not have women's representation that would grow faster than non-adopting countries if the adopters never actually imposed a quota. To assess the validity of this assumption, we examine whether trends in the dependent variable are different among countries that

[5]Because not all countries that adopted a gender quota did so in the same year, we cannot compare treated and untreated countries in each year, one by one. The treated countries have a different history of experience with the quota in any given year, and thus year-by-year comparisons cannot accurately capture growth in the effectiveness of the quota over time. Consequently, we instead compare treated countries p years before or after they adopted the quota to the overall average of untreated countries. Year fixed effects in the model of equation 1 remove any common time trends in the data, making this comparison more appropriate. This is also the approach taken by Bertrand, Duflo and Mullainathan (2004, p. 250).

eventually adopt a quota, but *before the quota is actually adopted*, compared to trends in the dependent variable among those countries that never adopt a quota. In terms of the model of equation 1, we expect $\beta_p \approx 0$ for $p < -1$. If there is no consistent difference between the two groups before quota adoption, we gain confidence that quota adopters and non-adopters are not on different paths that would diverge regardless of the quota.

Identification in a DID design also requires that quota adoption q_i not be caused by any omitted variable that also causes y_{it}, including reverse causality (simultaneity); this is sometimes called conditional ignorability. The country and year fixed effects control for any influences that are constant within countries over time, or constant within a year across all countries. But it is possible that there are influences that vary within countries and times and could confound the relationship between gender quotas in the legislature and the dependent variable. We therefore verify that an instrumental variables model robust to simultaneity and omitted variable bias produces results that are substantively similar to our DID analysis (Angrist and Krueger, 1999, pp. 1300-1305). A test for the exogeneity of quotas can also help us establish whether we need to employ instrumental variables (Baum, Schaffer and Stillman, 2003), as these models are less efficient and therefore a basic DID model is preferred if instruments are unnecessary. These results are presented in Appendix A; we find little evidence that corruption causes quota adoption or that the two variables are endogenous conditional on our controls.

We also estimate a simpler and more efficient binary treatment dynamic fixed effects model:

$$y_{it} = \zeta \times y_{i(t-1)} + \beta \times q_{it} + \alpha_i + \gamma_t + \varepsilon_{it} \tag{2}$$

This model, which allows the data to estimate a single coefficient β as an average effect of the quota on y_{it} over all countries and times, should improve the efficiency of our estimates because the same amount of data is being used to estimate only one β instead of many.[6]

[6]Note that the quota indicator is now indexed by country i and time t, reflecting the fact that the quota

We also include a lag of the dependent variable as a predictor in this simplified model. Although models with both fixed effects and lagged dependent variables as predictors suffer from Nickell bias (Nickell, 1981; Judson and Owen, 1999), this bias gets smaller as the time dimension of the data set increases. In most of our models, on average a country is observed for more than 25 years. It is also plausible to expect persistence in corruption levels over time, even if institutional factors change, and omitting this persistence could cause its own form of misspecification bias. We also report models without a lagged dependent variable and without fixed effects; these models may provide a "bracket" around the correct results if either the parallel trends or conditional ignorability assumptions are false (Ding and Li, 2019); we refer to these additional results in the text below.

Quotas and women's representation

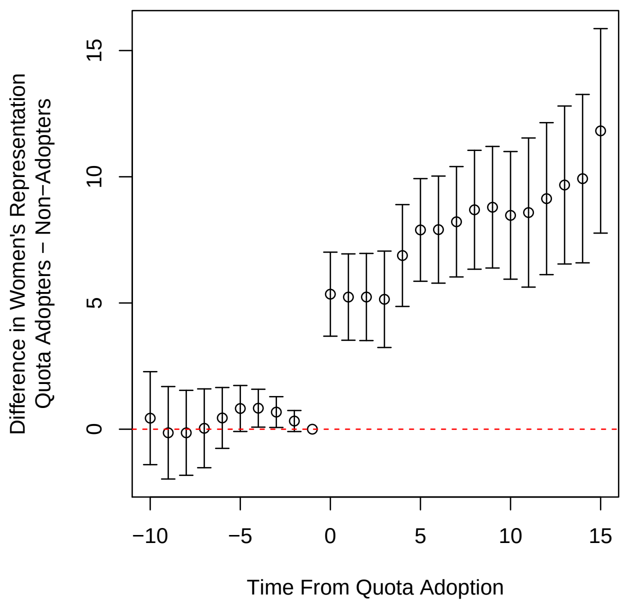
As noted above, we argue that quotas must successfully increase women's representation if they are going to cause reduced corruption. Figure 1 shows the DID estimate of the effect of legislative gender quotas on the proportion of women in parliament. Each point is an element from the vector of β estimates from the model of equation 1; the bars represent 95% confidence intervals for β_p where $-10 \leq p \leq 10$ based on standard errors clustered by country.

The figure makes clear that, on average, the gap in women's representation between countries that adopt a gender quota for their legislature and those that do not is stable for ten years before quota adoption. However, countries that adopt a quota immediately gain an average of 5 percentage points of women in parliament compared to countries that never adopt; this effect of quotas continues to rise over time. Our finding is consistent with that of Schwindt-Bayer (2009), who also finds a positive effect of quotas on women's representation.

Appendix Table 14 reports estimates for the model in equation 2 with country and time

indicator $q_{it} \in \{0, 1\}$ indicates whether country i has adopted a quota at time t or not.

Figure 1: Estimated effect of legislative gender quota on women's representation



Each point represents an estimate of the difference-in-differences for the percentage of women in parliament between countries that adopt a gender quota for their legislature and those that do not; this is β_p in the model described by 1 for the value of p on the x-axis. The baseline category of comparison is one year before quota adoption; the estimate is definitionally zero for this time. 95% confidence intervals for each estimate are indicated by bars and based on standard errors clustered by country.

FEs only, lagged DV only, or both FEs and lagged DVs.[7] All models report a statistically significant and positive effect of quotas on women's representation, with at least a 6 percentage point increase in women's representation caused by the quota. The models with lagged dependent variables imply an even larger effect of quotas on representation that cumulates over time, which can be calculated using the long-run multiplier described in Keele and Kelly (2006); for the model with FEs and lagged DV, quotas implementation causes just under a 10 percentage point cumulative increase in women in the legislature.

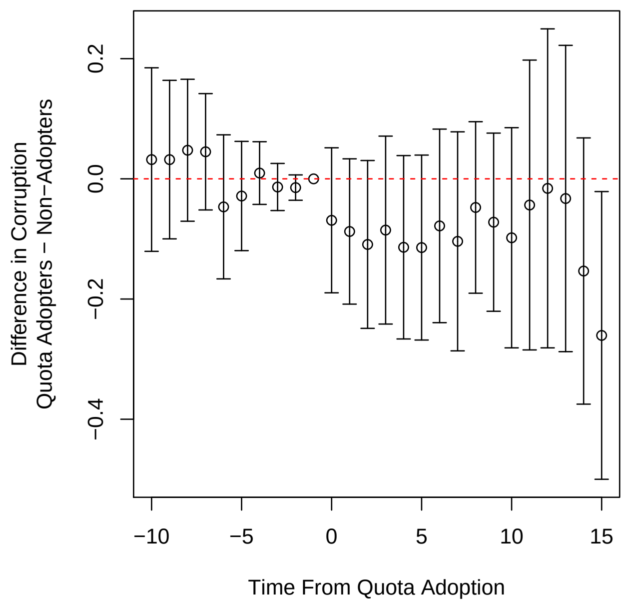
Our evidence is consistent with the conclusion that gender quotas raise women's representation in the legislature by between 5-10 percentage points on average in the long run, as is necessary for the causal mechanism we believe might connect gender quotas to reduced corruption. Based on the estimates of Esarey and Schwindt-Bayer (2019), this magnitude of change in women's representation in parliament should be sufficient to cause a substantively meaningful change in corruption; they report that a 10 percentage point increase in women in parliament typically causes a change in corruption on the order of 10-30% of the possible range of the measure.

Quotas and corruption in the legislature

Figure 2 shows a difference-in-differences plot of the effect of gender quotas on the V-Dem legislative corruption index using the model of equation 1. Although corruption seems to decline by about -0.1 (on the roughly seven-point scale of the measure) after the quota is adopted, this change is not statistically significant. The magnitude of the change is roughly consistent for ten years after adoption, suggesting that we should use the more efficient model of equation 2 that allows the quota effect to be constant over time; this model might be able to statistically resolve that effect.

[7] These models are estimated using `xtreg` in Stata 15.1.

Figure 2: Estimated effect of legislative gender quotas on V-Dem legislative corruption



Each point represents an estimate of the difference-in-differences for the V-Dem legislative corruption variable between countries that adopt a gender quota for their legislature and those that do not. The baseline category of comparison is one year before quota adoption; the estimate is definitionally zero for this time. 95% confidence intervals for each estimate are indicated by bars and based on standard errors clustered by country.

Table 2 reports the results of a fixed effects linear model in the form of equation 2 assuming a constant effect of quota adoption.[8] The first column indicates the results when using the V-Dem measure of legislative corruption as the dependent variable. According to this model, the implementation of a gender quota causes an immediate drop of nearly 0.04 points on the roughly seven point scale of the measure. This reduction is substantively minuscule, though statistically significant ($\alpha = 0.1$, two-tailed). However, the presence of a lagged dependent variable in the model implies a larger cumulative effect on corruption (Keele and Kelly, 2006). We calculate that implementation of a gender quota results in a 0.222 point drop in the V-Dem legislative corruption variable, about one-sixth of a standard deviation on the scale or about 3.4% of the maximum change possible. This change is statistically significant at conventional levels ($p = 0.054$, two-tailed). A model with no lagged dependent variable (shown in Appendix Table 15) or no fixed effects (Appendix Table 16) produces similar but less-certain results: a gender quota is estimated to reduce corruption by just under 0.1 points, but the effect is statistically insignificant at conventional levels.[9]

Quotas and corruption outside of the legislature

The four rightmost columns of Table 2 display results using the other V-Dem measures of corruption as the dependent variable in the model of equation 2. In all of these models, and also all of the versions of the model with no lagged dependent variable in Appendix Table 15, there is no statistically detectable effect of a legislative gender quota on corruption. Nor do difference-in-difference plots (based on the model of equation 1) for these alternative measures, shown in Appendix Figure 7, indicate any consistent effect of quotas on corruption in non-legislative branches or in government corruption overall (with the possible exception of

[8]The models in Table 2, 15, and FE models in Appendix Table 17 are estimated using `xtreg` in Stata 15.1. Models in Table 16 and models without FEs in Table 17 are estimated using `regress` in Stata 15.1.

[9]In the dynamic model, the long-run effect is larger but statistically insignificant.

Table 2: Fixed effects linear model estimates for the effect of legislated gender quotas on corruption

	Leg	Pub	Exec	Jud	Overall
lag corruption	0.822***	0.865***	0.862***	0.855***	0.884***
	(29.00)	(45.14)	(52.64)	(46.80)	(52.52)
presence of legislated quota	-0.0396*	-0.000157	-0.0000919	-0.0161	-0.00191
	(-1.83)	(-0.04)	(-0.02)	(-1.24)	(-0.60)
Observations	4251	4435	4435	4431	4422
Countries	173	173	173	173	173
Years	26	26	26	26	26
Country FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: V-Dem measure of corruption indicated in the column heading. Independent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. Standard errors are clustered by country.

a substantively small and statistically uncertain negative impact of legislative gender quotas on the measure of judicial corruption). Dynamic models without fixed effects (in Appendix Table 16) do show a statistically significant but small short-run negative effect of gender quotas on both overall and judicial corruption. Adoption of a quota is predicted to cause a long-run 1.8 point decline in the judicial corruption index ($p = 0.116$, two-tailed) and a 0.96 point decline on the overall corruption scale ($p = 0.103$, two-tailed) but both effects are statistically insignificant at conventional levels.

Models using our alternative measures of overall government corruption, the ICRG Corruption Risk measure and Bayesian Corruption Index, support the conclusion that legislative gender quotas do not influence corruption on average in the full sample of countries. Appendix Table 17 reports results for models based on equation 2 using these alternative dependent variables, with and without fixed effects or lagged dependent variables.[10] All the

[10]Exploratory analysis indicated the possibility for a deeper lag structure for the ICRG and BCI measures

variants of these models reported in the table indicate no statistically meaningful effect of gender quotas on overall government corruption. Difference-in-difference plots (based on the model of equation 1) shown in Appendix Figure 7 indicate no relationship between these overall measures of corruption and gender quotas.

Contextual sensitivity in the effect of gender quotas on legislative corruption

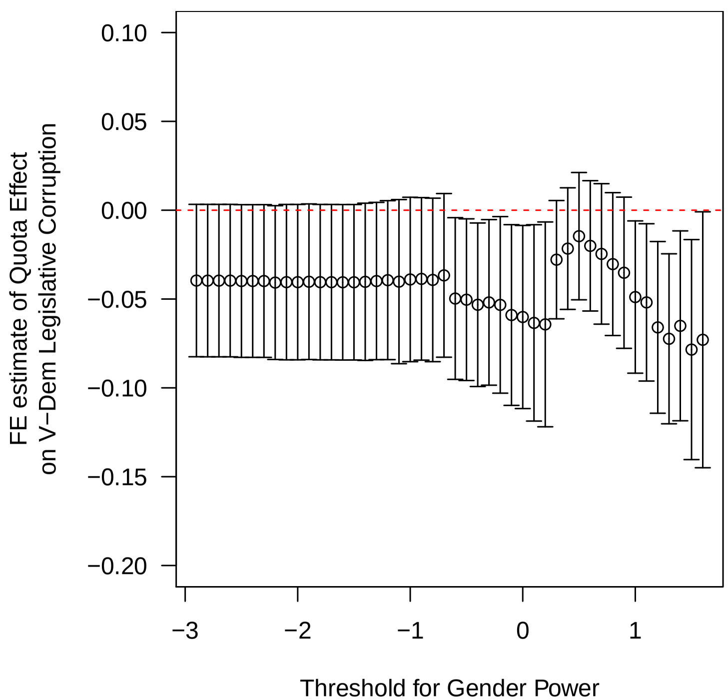
As we stipulated in our theory, we expect the effect of quotas on corruption to be strongest in places (1) where women are empowered to influence policy and (2) are accountable to voters. We therefore repeated the analysis of Table 2 in two subsets of our data: (1) among countries with high ratings on the V-Dem gender power index, and (2) among countries with high ratings on the V-Dem electoral democracy index. The results of these models are graphically depicted in Figures 3a and 3b which show the estimated marginal effect of gender quotas on legislative corruption in different subsamples.

There is little difference between the effect of quotas in states with high electoral democracy compared to states without.[11] However, quotas are more negatively associated with corruption in the legislature among countries with higher ratings on the V-Dem gender power index. The effect is statistically significant ($\alpha = 0.05$, two-tailed) for country-years with high gender power on the V-Dem index. For example, when the V-Dem gender power index is above 1.4, the substantive size of the instantaneous effect of quotas on corruption in the legislature is nearly double the effect estimated in the full sample in Table 2. The long-run impact of the quota is a reduction in corruption of just under half a point on the seven-point legislative corruption scale. This is roughly equivalent to the difference in legisla-

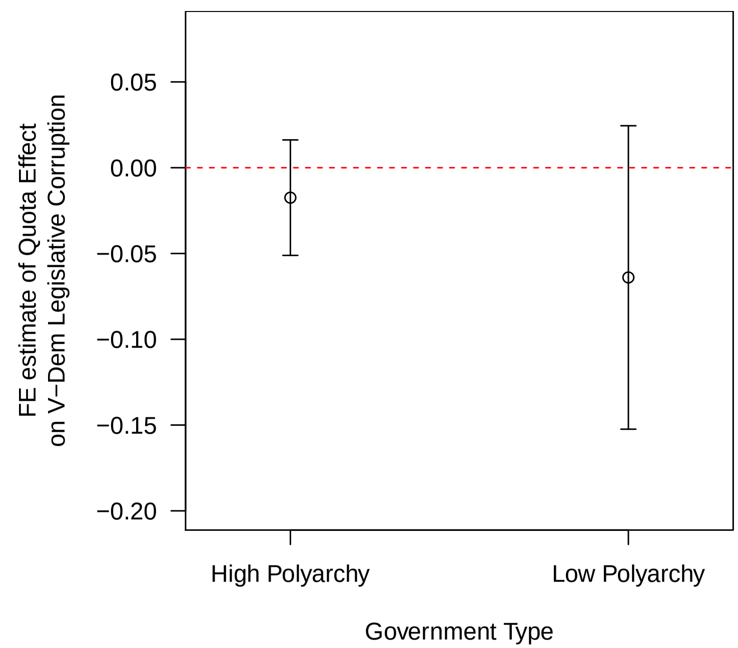
of corruption compared to the V-Dem measures. We therefore estimated models with no lag, one lag, and four lags of the dependent variable.

[11] This is also true for all our other measures of corruption; see Appendix Figure 8.

Figure 3: Estimated effect of legislative gender quotas on corruption in the legislature, by subgroups



(a) by gender power-sharing



(b) by electoral democracy score

Each point represents the estimate of the effect of legislative gender quotas on the V-Dem legislative corruption measure for a subset of the data with V-Dem gender power index greater than the value on the x-axis (left panel) or defined by a range of V-Dem electoral democracy scores (right panel), where "High" indicates an electoral democracy (polyarchy) score above the sample median and "Low" indicates a score equal to or below the median. The estimate comes from a fixed effects model with a single lag of the dependent variable plus country and year FEs (as in equation 2). 95% confidence intervals for each estimate are indicated by bars and based on standard errors clustered by country.

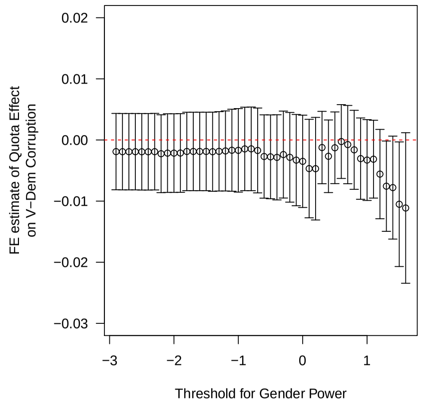
tive corruption in the United States (average value ≈ -1.07) compared to Senegal (average value ≈ -0.487) or Lithuania (average value ≈ -0.643). Appendix Table 23 reports results from alternative models with no lagged DV or no FEs. A smaller but statistically significant effect is estimated in a model with country and year FEs but no lagged DV; a model with only a lagged DV reports a larger but statistically insignificant long-run effect.

Figure 4 repeats the analysis of Figure 3a for our other measures of corruption. Interestingly, there is a statistically significant and negative effect on overall government corruption at high levels of gender power for two out of three of our measures of overall corruption throughout government: the V-Dem measure of overall corruption (panel 4a) and ICRG corruption risk (panel 4b). For example, among country-years with a V-Dem gender power rating above 1.4, adopting a gender quota for the legislature is associated with an instantaneous decrease in corruption of about 0.008 points ($p = 0.060$, two-tailed) and a long-run decrease of about 0.05 points ($p = 0.037$, two-tailed) on the V-Dem overall corruption scale. The long-run decrease constitutes about 5% of the largest possible change on the scale of the dependent variable.[12] For country-years with V-Dem gender power score over 0.5, a gender quota is associated with a long-run decrease in ICRG corruption risk of about 0.37 points, also about 5% of the largest possible change on this scale.[13] None of the individual branch measures of V-Dem corruption nor the Bayesian Corruption Index show any statistically significant or substantively meaningful effect of gender quotas on corruption (panels 4c through 4f).

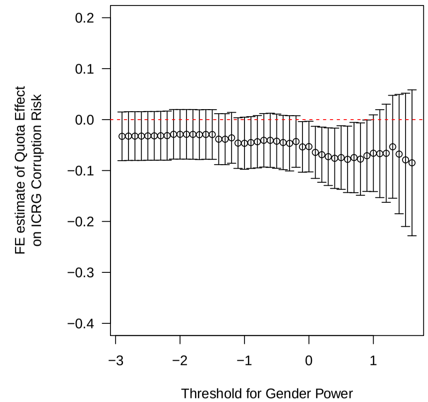
[12] Alternative models without fixed effects or lagged DVs are reported in Appendix Table 24. All show a negative and statistically significant relationship between quotas and V-Dem overall corruption.

[13] Alternative models without fixed effects or lagged DVs are reported in Appendix Table 25. All show a negative relationship between quotas and ICRG corruption risk, but only models with both country and year FEs as well as lagged DVs show a statistically significant relationship.

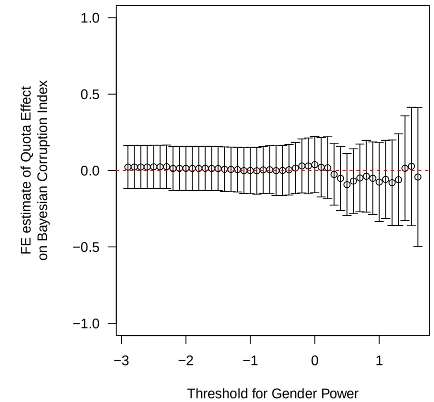
Figure 4: Estimated effect of legislative gender quotas on other measures of corruption, by gender power



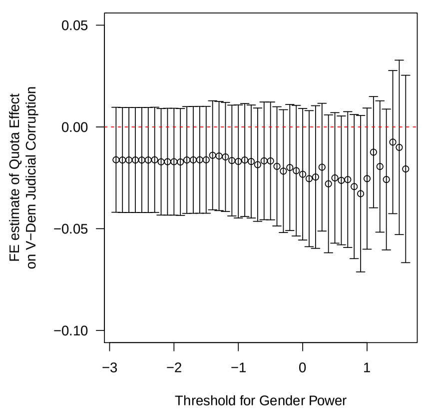
(a) V-Dem Overall Corruption



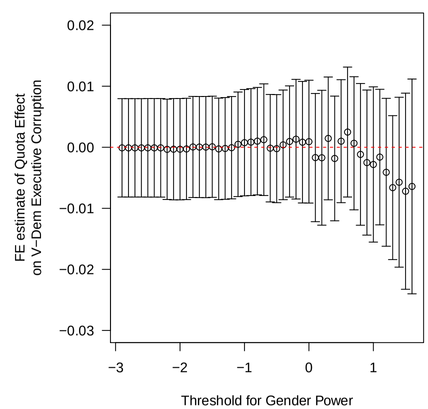
(b) ICRG Corruption Risk



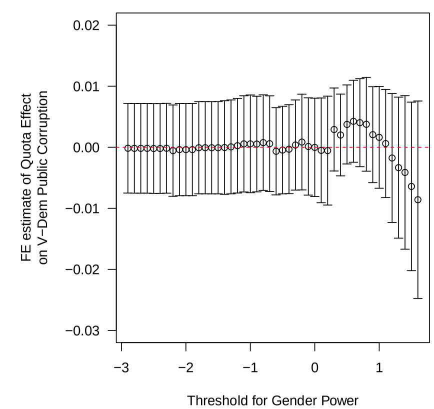
(c) Bayesian Corruption Index



(d) V-Dem Judicial Corruption



(e) V-Dem Executive Corruption



(f) V-Dem Public Sector Corruption

Each point represents the estimate of the effect of legislative gender quotas on the corruption measure listed in a subset of the data with V-Dem gender power index greater than the value on the x-axis. The estimate comes from a model country and year FEs and a single lag of the dependent variable (as in equation 2). 95% confidence intervals for each estimate are indicated by bars and based on standard errors clustered by country.

Conclusion

In this paper, we studied the causal relationship between legislative gender quotas and corruption. There are many plausible theoretical explanations of how this relationship works and what contextual factors affect its magnitude. Based on prior scholarship, we expected that these quotas could only reduce corruption if they increased women's representation in government. We also theorized that the best chance for gender quotas to lower corruption existed in country-years where (a) women have agency to participate in policy making and (b) they are accountable to voters. Finally, we acknowledged the possibility that electoral accountability may not matter if women's resistance to corruption is intrinsic and not driven by incentives.

Among countries where women have a meaningful share of power, we find a substantively important causal impact of legislative gender quotas on corruption both in the legislature and in government overall. These effects are relatively small at first but cumulate over time and are large enough to be politically relevant. The effect of gender quotas on corruption does not appear to be substantially different in countries with a system of electoral accountability of politicians to voters compared to those without. Nor do we detect simultaneity in the link between gender quotas and corruption within countries.

From a theoretical perspective, the fact that electoral democracy does not make quotas more effective at reducing corruption suggests that neither the documented greater risk aversion of women nor a tendency for voters to hold women to higher ethical standards can fully explain the link between women's representation and corruption. That does not mean that we have falsified these explanations: given the wealth of prior evidence, we expect they are impactful and even dominant in some contexts. Similarly, while corruption can cause gender quotas to be adopted (as shown in prior work), we were able to exclude this reverse causality as an explanation for our present findings. It appears that some form of *intrinsic* resistance to corruption among women is the best explanation for those findings. The intrinsic resistance of women to corruption explains why women's representation only reduces corruption when women have access to power: their preferences must have an influence on political decision-making in order for their presence to lower corruption.

As a practical policy matter, our research supports the conclusion that imposing gender quotas for the legislature as an anti-corruption measure can work. However, it will only work where women have equal access to genuine political power. Consequently, any plan to reduce corruption by boosting women's representation should simultaneously seek to ensure women's political equality in government. This additional work is necessary to ensure that their representation does merely reproduce existing corruption (Bjarnegård, Yoon and Zetterberg, 2018) but leverages women's intrinsic preferences to reduce it.

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.

Angrist, Joshua D. and Alan B. Krueger. 1999. Empirical strategies in labor economics. In *Handbook of Labor Economics*. Vol. 3 Elsevier pp. 1277–1366.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. 1 edition ed. Princeton: Princeton University Press.

Arellano, Manuel and Stephen Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The review of economic studies* 58(2):277–297.

Armantier, Olivier and Amadou Boly. 2013. "Comparing Corruption in the Laboratory and in the Field in Burkina Faso and in Canada." *The Economic Journal* 123(573):1168–1187. **URL:** *http://onlinelibrary.wiley.com/doi/10.1111/ecoj.12019/abstract*

Barnes, Tiffany D. and Emily Beaulieu. 2014. "Gender Stereotypes and Corruption: How Candidates Affect Perceptions of Election Fraud." *Politics & Gender* 10(3):365–391.

Barnes, Tiffany D. and Emily Beaulieu. 2019. "Women Politicians, Institutions, and Perceptions of Corruption." *Comparative Political Studies* 52(1):134–167.

Baum, Christopher F., Mark E. Schaffer and Steven Stillman. 2003. "Instrumental variables and GMM: Estimation and testing." *Stata Journal* 3(1):1–31.

Baum, Christopher F., Mark E. Schaffer and Steven Stillman. 2007. "Enhanced routines for instrumental variables/GMM estimation and testing." *Stata Journal* 7(4):465–506.

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.

Beck, Linda J. 2003. "Democratization and the Hidden Public: The Impact of Patronage Networks on Senegalese Women." *Comparative Politics* 35(2):147–169.

Bernasek, Alexandra and Stephanie Shwiff. 2001. "Gender, Risk, and Retirement." *Journal of Economic Issues* 35(2):345–356.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119(1):249–275. URL: `https://dx.doi.org/10.1162/003355304772839588`.

Bjarnegård, Elin. 2013. *Gender, Informal Institutions and Political Recruitment: Explaining Male Dominance in Parliamentary Representation.* Basingstoke, UK: Palgrave Macmillan.

Bjarnegård, Elin, Mi Yung Yoon and Pär Zetterberg. 2018. Gender Quotas and the Re (pro) duction of Corruption. In *Gender and Corruption.* Springer pp. 105–124.

Blundell, Richard and Stephen Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics* 87(1):115–143.

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.

Burnet, Jennie. 2012. Women's Empowerment and Cultural Change in Rwanda. In *The Impact of Gender Quotas*, ed. Susan Franceschet, Mona Lena Krook and Jennifer M. Piscopo. Oxford University Press chapter 12, pp. 190–206.

Bush, Sarah Sunn. 2011. "International Politics and the Spread of Quotas for Women in Legislatures." *International Organization* 65(1):103–137. URL: https://www.jstor.org/stable/23016105.

Byrnes, James P., David C. Miller and William D. Schafer. 1999. "Gender Differences in Risk-Taking: A Meta-analysis." *Psychological Bulletin* 125(3):367–383.

Clark, Tom and Drew Linzer. 2015. "Should I Use Fixed or Random Effects?" *Political Science Research and Methods* 3(2):399–408. URL: https://doi.org/10.1017/psrm.2014.32.

Collier, David. 2011. "Understanding Process Tracing." *PS: Political Science & Politics* 44(4):823–830.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Valeriya Mechkova, Johannes von Römer, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi ting Wang, Tore Wig, and Daniel Ziblatt. 2019. "V-Dem Codebook v9." Online. URL: https://www.v-dem.net/en/data/archive/previous-reference-materials/reference-materials-v9/ accessed 7/15/2021.

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.

Croissant, Yves and Giovanni Millo. 2008. "Panel Data Econometrics in R: The plm Package." *Journal of Statistical Software* 27(2):1–43.

Croson, Rachel and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47(2):448–474.

Ding, Peng and Fan Li. 2019. "A Bracketing Relationship between Difference-in-Differences and Lagged-Dependent-Variable Adjustment." *Political Analysis* 27(4):605–615.

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.

Eckel, Catherine C. and Philip J. Grossman. 2008. Men, Women, and Risk Aversion: Experimental Evidence. In *Handbook of Experimental Economic Results*, ed. Charles Plott and Vernon Smith. Vol. 1 New York: Elsevier pp. 1061–1073.

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 Andrew Menger. 2017. "Practical and Effective Approaches to Dealing with Clustered Data." *Political Science Research and Methods* forthcoming:1–35. URL: `http://jee3.web.rice.edu/cluster-paper.pdf`.

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 Schwindt-Bayer. 2018. "Women's Representation, Accountability, and Corruption in Democracies." *British Journal of Political Science* 48(3):659–690. URL: `https://doi.org/10.1017/S0007123416000478`.

Esarey, Justin and Leslie Schwindt-Bayer. 2019. "Estimating Causal Relationships Between Women's Representation in Government and Corruption." *Comparative Political Studies* 52(11):1713–1741. URL: `https://doi.org/10.1177/0010414019830744`.

Grimes, Marcia and Lena Wängnerud. 2012. "Good Government in Mexico: The Relevance of the Gender Perspective." *QoG Working Paper Series* (11):1–31. URL: `https://goo.gl/WsKzyY`.

Hassin, Shireen. 2010. *Political Representation.* Cambridge University Press chapter 8: Perverse Consequences? The impact of quotas for women on democratization in Africa, pp. 211–235.

Holtz-Eakin, Douglas, Whitney Newey and Harvey S Rosen. 1988. "Estimating vector autoregressions with panel data." *Econometrica: Journal of the Econometric Society* pp. 1371–1395.

Htun, Mala. 2016. *Inclusion without Representation in Latin America*. Cambridge University Press.

Hughes, Melanie M., Mona Lena Krook and Pamela Paxton. 2015. "Transnational Women's Activism and the Global Diffusion of Gender Quotas." *International Studies Quarterly* pp. 357–372.

Hughes, Melanie M., Pamela Paxton, Amanda Clayton and Pär Zetterberg. 2019. "Global Gender Quota Adoption, Implementation, and Reform." *Comparative Politics* 51(2):219–238.

Jann, Ben. 2005. "Making Regression Tables from Stored Estimates." *Stata Journal* 5(3):288–308.

Jann, Ben. 2007. "Making Regression Tables Simplified." *Stata Journal* 7(2):227–244.

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.

Judson, Ruth A. and Ann L. Owen. 1999. "Estimating dynamic panel data models: a guide for macroeconomists." *Economics Letters* 65(1):9–15.

Kahn, Carrie. 2013. "Mexican State's Anti-Corruption Plan: Hire Female Traffic Cops." Online. URL: `http://www.npr.org/2013/09/28/226903227/mexican-state-s-anti-corruption-plan-hire-women-traffic-cops` accessed 7/11/2019.

Karim, Sabrina. 2011. "Madame Officer." *Americas Quarterly* 5(3). URL: `http://www.americasquarterly.org/node/2802/` accessed July 20, 2012.

Keele, Luke and Nathan J. Kelly. 2006. "Dynamic models for dynamic theories: The ins and outs of lagged dependent variables." *Political Analysis* 14(2):186–205.

Krook, Mona Lena. 2006. "Reforming Representation: The Diffusion of Candidate Gender Quotas Worldwide." *Politics & Gender* 2(3):303–327.

Krook, Mona Lena. 2009. *Quotas for Women in Politcs*. Oxford University Press.

Larson, Anna. 2012. Collective Identities, Institutions, Security, and State Building in Afghanistan. In *The Impact of Gender Quotas*, ed. Susan Franceschet, Mona Lena Krook and Jennifer M. Piscopo. Oxford University Press chapter 9, pp. 136–153.

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`.

Liu, Shan-Jan Sarah and Kim Yi Dionne. 2019. "Constraints on Women Politicans' Role Model Effects: Evidence from Africa." Working Paper.

Longman, Timothy. 2006. Rwanda: Achieving Equality or Serving an Authoritarian State? In *Women in African Parliaments*, ed. Gretchen Bauer and Hannah E. Britton. Lynne Rienner Publishers chapter 6, pp. 133–150.

Matland, Richard E. 2006. Electoral quotas: Frequency and effectiveness. In *Women, Quotas and Politics*, ed. Drude Dahlerup. Routledge chapter 13, pp. 275–292.

McDermott, Jeremy. 1999. "International: Women Police Ride In on a Ticket of Honesty." *The Daily Telegraph* . July 31, 1999.

Moore, Molly. 1999. "Mexico City's Stop Sign to Bribery; To Halt Corruption, Women Traffic Cops Replace Men." *The Washington Post* . July 31, 1999. URL: http://www.highbeam.com/doc/1P2-605613.html.

Nickell, Stephen. 1981. "Biases in dynamic models with fixed effects." *Econometrica: Journal of the Econometric Society* 49(6):1417–1426.

O'Brien, Diana Z. and Johanna Rickne. 2016. "Gender Quotas and Women's Political Leadership." *American Political Science Review* 110(1):112–126.

Paweenawat, Sasiwimon W. 2018. "The gender-corruption nexus in Asia." *Asian-Pacific Economic Literature* 32(1):18–28. URL: https://doi.org/10.1111/apel.12214.

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri and Johannes von Römer. 2019. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." Online. V-Dem Working Paper No. 21. 4th edition. University of Gothenburg: Varieties of Democracy Institute.

Political Risk Services Group. 2018. "ICRG Methodology." Online. URL: https://www.prsgroup.com/wp-content/uploads/2018/01/icrgmethodology.pdf accessed November 11, 2019.

Pustejovsky, James. 2019. *clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections.* R package version 0.3.3. URL: https://CRAN.R-project.org/package=clubSandwich.

Reed, William Robert. 2015. "On the Practice of Lagging Variables to Avoid Simultaneity." *Oxford Bulletin of Economics and Statistics* 77(6):897–905.

Roodman, David. 2009. "How to do xtabond2: An introduction to difference and system GMM in Stata." *The Stata Journal* 9(1):86–136.

Sater, James N. 2012. *The Impact of Gender Quotas.* Oxford University Press chapter 5: Reserved Seats, Patriarchy, and Patronage in Morocco, pp. 72–86.

Schulze, Günther G. and Björn Frank. 2003. "Deterrence versus intrinsic motivation: Experimental evidence on the determinants of corruptibility." *Economics of Governance* 4:143–160.

Schwindt-Bayer, Leslie A. 2009. "Making Quotas Work: The Effect of Gender Quota Laws On the Election of Women." *Legislative Studies Quarterly* 34(1):5–28.

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.

Standaert, Samuel. 2015. "Divining the level of corruption: A Bayesian state-space approach." *Journal of Comparative Economics* 43(3):782–803. URL: `https://doi.org/10.1016/j.jce.2014.05.007`.

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

Sundén, Annika E. and Brian J. Surette. 1998. "Gender Differences in the Allocation of Assets in Retirement Savings Plans." *The American Economic Review* 88(2):207–211.

Sundström, Aksel and Lena Wängnerud. 2014. "Corruption as an obstacle to women's political representation: Evidence from local councils in 18 European countries." *Party Politics* 22(3):364–369.

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

Tavits, Margit. 2007. "Clarity of responsibility and corruption." *American Journal of Political Science* 51(1):218–229.

Teorell, Jan, Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Natalia Alvarado Pachon and Richard Svensson. 2019. "The Quality of Government Dataset, version Jan19." URL: http://www.qog.pol.gu.se.

Torgler, Benno and Neven T. Valev. 2010. "Gender and public attitudes toward corruption and tax evasion." *Contemporary Economic Policy* 28(4):554–568.

Wagner, Natascha, Matthias Rieger, Arjun Bedi and Wil Hout. 2017. "Gender and policing norms: Evidence from survey experiments among police officers in Uganda." *Journal of African Economies* 26(4):492–515.

Walsh, Denise. 2012. Party Centralization and Debate Conditions in South Africa. In *The Impact of Gender Quotas*, ed. Susan Franceschet, Mona Lena Krook and Jennifer M. Piscopo. Oxford University Press chapter 8, pp. 119–135.

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.

Ward, Sarah J. and Laura A. King. 2018. "Gender Differences in Emotion Explain Women's Lower Immoral Intentions and Harsher Moral Condemnation." *Personality and Social Psychology Bulletin* 44(5):653–669. PMID: 29291658.

Watson, John and Mark McNaughton. 2007. "Gender differences in risk aversion and expected retirement benefits." *Financial Analysts Journal* 63(4):52–62.

Wills, Kate. 2015. "Kruger National Park: Meet the women hunting South Africa's poachers." *The Independent* September 28. URL: `https://bit.ly/31HKrSR` accessed 6/20/2019.

World Bank. 2016. *World Development Indicators*. The World Bank.

A Online Appendix: Instrumental variables analysis

Difference-in-difference analysis is only valid if there is no confounding from omitted variables and no simultaneity between corruption and quota imposition. Therefore we use instrumental variables models to assess (a) whether corruption causes quota adoption, and (b) whether there is a detectable effect of quota adoption on corruption once quotas are instrumented. Toward these ends, we employ two models:

1. a fixed effects (FE) model with instrumental variables (IVs) using two-step feasible GMM (Baum, Schaffer and Stillman, 2003, 2007); and

2. a dynamic panel data (DPD) model with year fixed effects (Roodman, 2009), in both system (Blundell and Bond, 1998) and difference (Holtz-Eakin, Newey and Rosen, 1988; Arellano and Bond, 1991) one-step GMM variants with robust standard errors.

Two- and three-year lags of corruption serve as our instruments for the one-year lag of corruption in the fixed effects IV model. This instrumentation strategy, which is similar to the strategy employed in the DPD model and one recommended by Reed (2015), makes the exclusion restriction that an instrumented independent variable $x_{(t-1)}$ observed at time $t-1$ is independent of y_t conditional on x_t and y_{t-1}. Including the lagged dependent variable in the model is important because it serves to block a plausible back-door pathway of influence that could contaminate the instruments and make the results invalid.

DPD models overcome the possibility of Nickell bias, but replace it with the possibility of sensitivity to specification (e.g., in the number of lags used as instruments or in whether both difference and level moment conditions are used in estimation). Given the possible deficiencies in each approach, we believe that our conclusions are most robust when most or all models indicate a similar answer.

Effect of corruption on adoption of a gender quota for the legislature

We predict the presence of a gender quota for the legislature with a one-period lag of corruption to reflect the fact that policies are determined on information available at the time—that is to say, in the recent past—and not contemporaneous information that may not yet be available.[14] Standard errors are clustered on country.

Our instrumental variable analysis for the effect of corruption in the legislature on enactment of parliamentary gender quotas in the full data set is presented in Table 3. In this analysis, we find that legislative corruption has no substantive effect on the implementation of gender quotas. Moreover, a chi-squared test for endogeneity indicates that (one-year lagged) corruption can safely be treated as exogenous in all our models.

Tables 4, 5, and 6 repeat the analysis of Table 3 with the V-Dem overall corruption measure, ICRG Corruption Risk, and Bayesian Corruption Index as the dependent variable. In eight of the nine models, a chi-square test fails to reject the null that corruption is exogenous

[14]Models are estimated in Stata 15.1 using the `xtivreg2` and `xtabond2` routines.

Table 3: Dynamic instrumental variable model estimates for the effect of legislative corruption on enactment of legislative gender quotas

	FE IV	System DPD	Diff. DPD
lag corruption	-0.00718	0.000596	-0.00734
	(-0.92)	(0.12)	(-0.38)
lag presence of legislated quota	0.856***	0.967***	0.802***
	(85.66)	(169.70)	(21.62)
Observations	3877	4284	4078
Countries	173	173	173
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	0.969	173.3	172.1
Hansen's J p-value	0.325	1	1
1st stage F-stat (Kleibergen-Paap)	747.0		
Endogeneity test stat.	0.168	0.802	38.52
Endogeneity test, p-value	0.682	1.000	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of legislative corruption score as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

to quota adoption. Six models find no statistically significant effect of corruption on quota adoption, two find a statistically significant positive effect, and one finds a statistically significant *negative* effect. Our overall conclusion is that there is little support for any consistent effect of corruption on the probability that a gender quota will be adopted for the legislature.

Theory led us to expect that corruption might be most likely to prompt the adoption of legislative gender quotas in authoritarian governments and/or in countries most susceptible to pressure from foreign governments and NGOs. We examine this possibility in Tables 7 through 10, where we repeat the analysis of Table 3 but among countries with low polyarchy scores (Table 7) or among the countries with the lowest mean GDP per capita over the time span of the data set (Tables 8 through 10).

There is little evidence that corruption in the legislature prompts quota adoption in autocracies, as the estimated causal relationship is statistically significant and in the wrong direction for two of our three models. However, in the very poorest countries (with mean log per capita GDP below the 10th or 15th percentile), we estimate an immediate 2-3 percentage point increase in the probability of adopting a legislative gender quota for every one point increase in the V-Dem legislative corruption measure in most of our models. The long-run effect is even greater; for example, the fixed effects instrumental variable model of Table 9 predicts that a one unit increase in legislative corruption raises the long-run probability of adopting a gender quota by over 16 percentage points ($p = 0.056$, two-tailed). However, this analysis is limited to a small subset of countries in the international system (27 countries and 500-600 observations in the models of Table 9).[15]

We note, however, that we find little evidence for impact of overall government corruption on quota adoption among the poorest countries. Tables 11 through 13 use lagged values of the V-Dem overall corruption measure, ICRG Corruption Risk, and Bayesian Corruption Index as predictors of the presence of a gender quota for the legislature among countries at or below the 15th percentile of mean log GDP per capita. Only one out of nine models shows any statistically significant relationship between overall corruption and quota adoption, and only one of the nine tests for endogeneity rejects the null hypothesis that corruption is exogenous to legislative gender quotas.

[15]The small number of countries in this subgroup analysis leads us to question the use of clustered standard errors because data sets with few clusters these standard errors are typically overconfident (Esarey and Menger, 2017). We therefore re-analyzed the models with ordinary robust standard errors (for FE IV models) or vanilla standard errors (for DPD models) and report these results in Appendix Tables 18 and 19. All instantaneous effects are statistically insignificant in these alternative models, but the long-run impacts of corruption on the probability of quota adoption in the fixed effects instrumental variable model of Appendix Table 19 are still statistically significant for countries at or below the 15th percentile of mean log GDP per capita ($p = 0.089$, two-tailed).

Table 4: Dynamic instrumental variable model estimates for the effect of V-Dem overall corruption on enactment of legislative gender quotas

	FE IV	System DPD	Diff. DPD
lag corruption	0.00363	0.00812	0.0295
	(0.08)	(0.51)	(0.28)
lag presence of legislated quota	0.848***	0.964***	0.851***
	(98.74)	(176.52)	(34.42)
Observations	4076	4422	4249
Countries	173	173	173
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	1.853	2254.7	173.0
Hansen's J p-value	0.173	2.87e-169	1
1st stage F-stat (Kleibergen-Paap)	1336.8		
Endogeneity test stat.	0.157	2083.1	36.02
Endogeneity test, p-value	0.692	1.000	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of V-Dem overall corruption risk as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 5: Dynamic instrumental variable model estimates for the effect of ICRG corruption risk on enactment of legislative gender quotas

	FE IV	System DPD	Diff. DPD
lag corruption	0.00147	-0.00386	-0.0248**
	(0.20)	(-1.18)	(-2.48)
lag presence of legislated quota	0.837***	0.961***	0.809***
	(76.52)	(147.68)	(29.49)
Observations	3169	3443	3306
Countries	137	137	137
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	2.855	119.9	125.9
Hansen's J p-value	0.0911	1	1
1st stage F-stat (Kleibergen-Paap)	1388.4		
Endogeneity test stat.	1.202	-13.36	63.64
Endogeneity test, p-value	0.273	1	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of ICRG corruption risk as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 6: Dynamic instrumental variable model estimates for the effect of the Bayesian Corruption Index on enactment of legislative gender quotas

	FE IV	System DPD	Diff. DPD
lag corruption	0.000520	0.000379***	0.00344*
	(0.40)	(2.89)	(1.84)
lag presence of legislated quota	0.836***	0.961***	0.818***
	(89.84)	(167.64)	(26.65)
Observations	3889	4235	4062
Countries	173	173	173
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	0.0342	173.0	172.7
Hansen's J p-value	0.853	1	1
1st stage F-stat (Kleibergen-Paap)	14260.2		
Endogeneity test stat.	1.096	-1.38892e+14	102.8
Endogeneity test, p-value	0.295	0	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of the Bayesian Corruption Index as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 7: Dynamic instrumental variable model estimates for the effect of corruption in the legislature on enactment of legislative gender quotas, countries with V-Dem Electoral Democracy Scores less than or equal to the sample median (0.524)

	FE IV	System DPD	Diff. DPD
lag corruption	-0.000414	0.0188**	-0.0159
	(-0.04)	(2.37)	(-0.60)
lag presence of legislated quota	0.886***	0.950***	0.710***
	(62.09)	(94.81)	(10.91)
Observations	1805	2050	1922
Countries	110	115	112
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	0.141	90.90	89.40
Hansen's J p-value	0.707	1	1
1st stage F-stat (Kleibergen-Paap)	282.1		
Endogeneity test stat.	0.652	46.45	61.38
Endogeneity test, p-value	0.419	1	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of legislative corruption score as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 8: Dynamic instrumental variable model estimates for the effect of corruption in the legislature on enactment of legislative gender quotas, countries at or below the 10th percentile of mean log GDP per capita in the data set

	FE IV	System DPD	Diff. DPD
lag corruption	0.0299^{*}	0.00290	0.0188^{*}
	(1.87)	(0.49)	(1.66)
lag presence of legislated quota	0.843^{***}	0.969^{***}	0.821^{***}
	(75.38)	(70.76)	(43.38)
Observations	367	422	393
Countries	19	19	19
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	0.0532	4.442	6.474
Hansen's J p-value	0.818	1	1
1st stage F-stat (Kleibergen-Paap)	35.09		
Endogeneity test stat.	0.868	3.607	3.231
Endogeneity test, p-value	0.352	1	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of legislative corruption score as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 9: Dynamic instrumental variable model estimates for the effect of corruption in the legislature on enactment of legislative gender quotas, countries at or below the 15th percentile of mean log GDP per capita in the data set

	FE IV	System DPD	Diff. DPD
lag corruption	0.0269*	-0.00571	0.0241*
	(1.91)	(-0.68)	(1.87)
lag presence of legislated quota	0.837***	0.965***	0.757***
	(35.71)	(62.23)	(21.14)
Observations	538	617	576
Countries	27	27	27
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	0.0512	5.666	8.762
Hansen's J p-value	0.821	1	1
1st stage F-stat (Kleibergen-Paap)	44.82		
Endogeneity test stat.	1.031	5.658	4.871
Endogeneity test, p-value	0.310	1	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of legislative corruption score as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 10: Dynamic instrumental variable model estimates for the effect of corruption in the legislature on enactment of legislative gender quotas, countries at or below the 20th percentile of mean log GDP per capita in the data set

	FE IV	System DPD	Diff. DPD
lag corruption	0.0124	0.00255	0.00584
	(0.77)	(0.35)	(0.30)
lag presence of legislated quota	0.854***	0.953***	0.780***
	(42.54)	(67.12)	(26.50)
Observations	762	867	813
Countries	37	37	37
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	0.102	8.351	12.44
Hansen's J p-value	0.749	1	1
1st stage F-stat (Kleibergen-Paap)	62.95		
Endogeneity test stat.	0.576	2.117	3.968
Endogeneity test, p-value	0.448	1	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of legislative corruption score as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 11: Dynamic instrumental variable model estimates for the effect of V-Dem overall corruption on enactment of legislative gender quotas, countries at or below the 15th percentile of mean log GDP per capita in the data set

	FE IV	System DPD	Diff. DPD
lag corruption	-0.00767	-0.0188	0.00968
	(-0.08)	(-0.73)	(0.09)
lag presence of legislated quota	0.876***	0.971***	0.772***
	(45.96)	(102.29)	(21.46)
Observations	647	701	674
Countries	27	27	27
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	1.474	1.624	9.242
Hansen's J p-value	0.225	1	1
1st stage F-stat (Kleibergen-Paap)	164.1		
Endogeneity test stat.	0.458	0.549	6.318
Endogeneity test, p-value	0.498	1	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of V-Dem overall corruption score as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 12: Dynamic instrumental variable model estimates for the effect of ICRG Corruption Risk on enactment of legislative gender quotas, countries at or below the 15th percentile of mean log GDP per capita in the data set

	FE IV	System DPD	Diff. DPD
lag corruption	0.00331	0.00886**	-0.00605
	(0.25)	(2.36)	(-0.52)
lag presence of legislated quota	0.881***	0.966***	0.828***
	(31.86)	(71.12)	(28.27)
Observations	431	467	449
Countries	18	18	18
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	2.689	4.872	5.193
Hansen's J p-value	0.101	1	1
1st stage F-stat (Kleibergen-Paap)	271.1		
Endogeneity test stat.	1.369	4.872	3.882
Endogeneity test, p-value	0.242	1	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of ICRG Corruption Risk as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

Table 13: Dynamic instrumental variable model estimates for the effect of the Bayesian Corruption Index on enactment of legislative gender quotas, countries at or below the 15th percentile of mean log GDP per capita in the data set

	FE IV	System DPD	Diff. DPD
lag corruption	0.00172	-0.000356	-0.000879
	(0.94)	(-0.61)	(-0.33)
lag presence of legislated quota	0.847***	0.965***	0.769***
	(38.73)	(89.90)	(38.51)
Observations	611	665	638
Countries	27	27	27
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Hansen's J	0.148	5525.8	8.197
Hansen's J p-value	0.701	0	1
1st stage F-stat (Kleibergen-Paap)	2276.4		
Endogeneity test stat.	1.763	-2670875.8	3.962
Endogeneity test, p-value	0.184	0	1

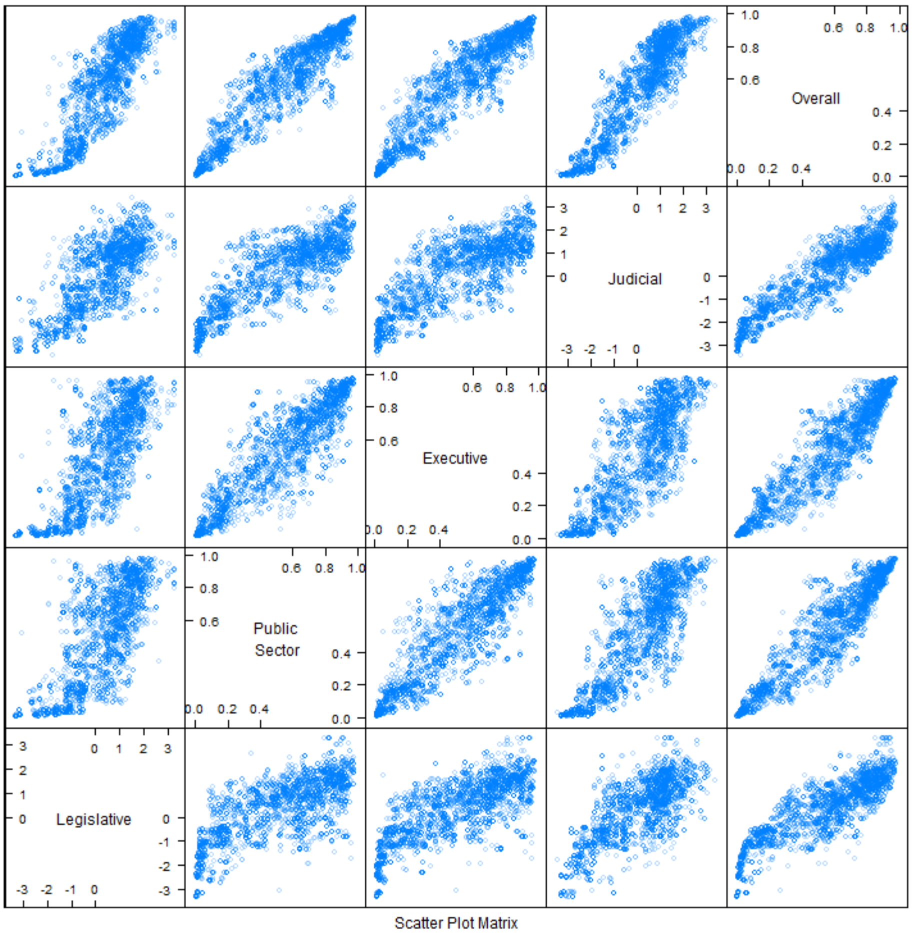
t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of the Bayesian Corruption Index as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors are clustered by country.

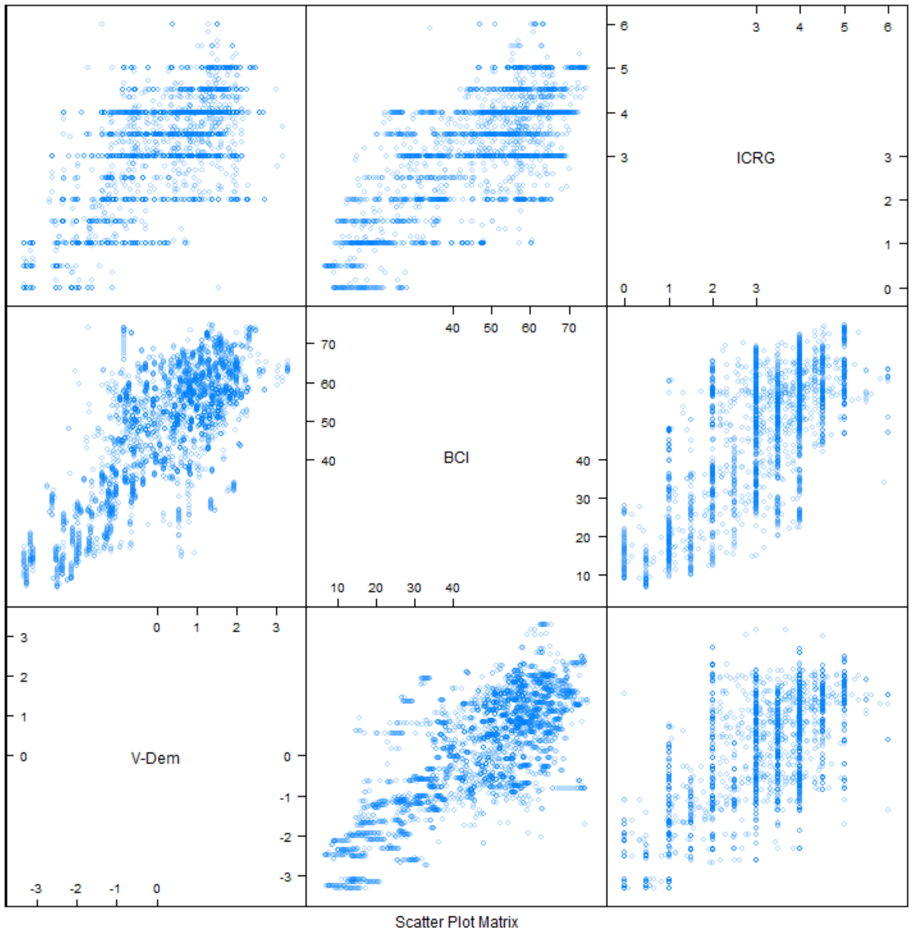
B Online Appendix: Additional Figures and Tables

Figure 5: Comparison of V-Dem corruption measures



This scatterplot shows the relationship between five different measures of corruption from the V-Dem data set: Overall, Judicial, Executive, Public Sector, and Legislative. Each point represents measurement of a country-year.

Figure 6: Comparison of corruption measures



This scatterplot shows the relationship between our three measures of corruption: the V-Dem legislative corruption measure (first column and third row), the Bayesian Corruption Index (second column and second row), and the International Country Risk Guide measure of corruption (third column and first row). Each point represents measurement of a country-year.

Table 14: Alternative model estimates for the effect of legislated gender quotas on women's representation

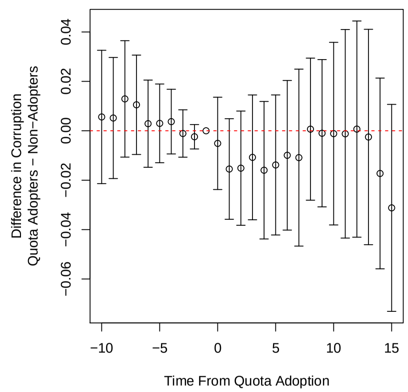
	FE only	FE w/ lag	lag only
presence of legislated quota	6.114***	1.890***	0.810***
	(6.57)	(6.69)	(7.02)
lag % women in legislature		0.809***	0.976***
		(58.03)	(232.00)
Observations	4411	4204	4204
Countries	174	173	173
Years	27	26	26
Country FE	Yes	Yes	No
Time FE	Yes	Yes	No

t statistics in parentheses

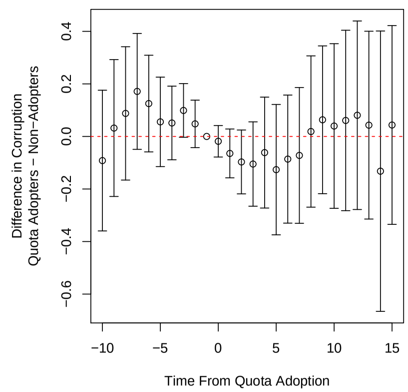
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: percentage of women in parliament. Standard errors are clustered by country.

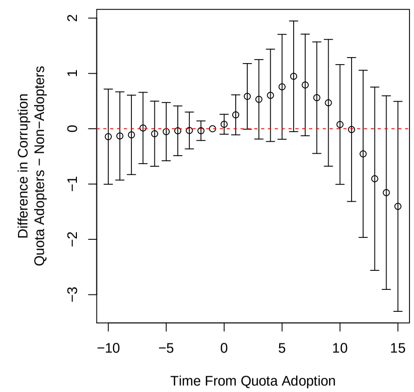
Figure 7: Estimated effect of legislative gender quotas on other measures of corruption



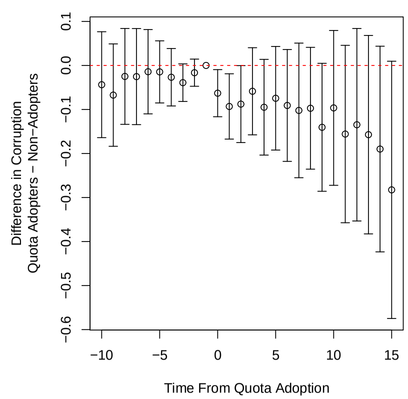
(a) V-Dem Overall Corruption



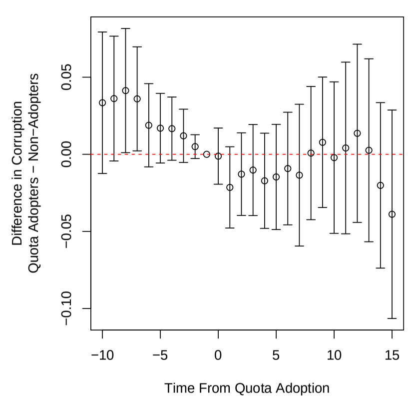
(b) ICRG Corruption Risk



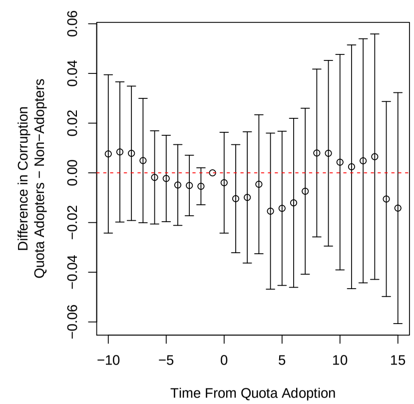
(c) Bayesian Corruption Index



(d) V-Dem Judicial Corruption



(e) V-Dem Executive Corruption



(f) V-Dem Public Sector Corruption

Table 15: Fixed effects linear model estimates for the effect of legislated gender quotas on corruption, no lagged dependent variable

	Leg	Pub	Exec	Jud	Overall
presence of legislated quota	-0.0970	-0.00236	-0.0233	-0.0465	-0.00964
	(-1.24)	(-0.15)	(-1.40)	(-0.78)	(-0.68)
Observations	4458	4609	4609	4605	4596
Countries	174	174	174	174	174
Years	27	27	27	27	27
Country FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: V-Dem measure of corruption indicated in the column heading. Independent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. Standard errors are clustered by country.

Table 16: Dynamic linear model estimates for the effect of legislated gender quotas on corruption, no fixed effects

	Leg	Pub	Exec	Jud	Overall
presence of legislated quota	-0.0114	-0.00186	-0.00238	-0.0115*	-0.00295**
	(-1.48)	(-1.21)	(-1.30)	(-1.85)	(-2.28)
lag corruption	0.990***	0.993***	0.988***	0.994***	0.997***
	(337.12)	(610.12)	(526.32)	(648.64)	(832.45)
Constant	0.000716	0.00387***	0.00489***	-0.0000188	0.00162**
	(0.20)	(4.00)	(4.61)	(-0.01)	(2.32)
Observations	4251	4435	4435	4431	4422
Countries	173	173	173	173	173
Years	26	26	26	26	26
Country FE	No	No	No	No	No
Time FE	No	No	No	No	No

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: V-Dem measure of corruption indicated in the column heading. Independent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. Standard errors are clustered by country.

Table 17: Model estimates for the effect of legislated gender quotas on corruption, alternative overall corruption measures

	ICRG	ICRG	ICRG	ICRG	BCI	BCI	BCI	BCI
presence of legislated quota	-0.0709	-0.0328	-0.0300	-0.00788	0.446	0.0228	0.0449	-0.0141
	(-0.90)	(-1.36)	(-1.17)	(-0.57)	(1.17)	(0.32)	(1.01)	(-0.64)
lag corruption		0.828***	0.986***	1.132***		1.011***	1.808***	1.887***
		(61.91)	(34.25)	(39.46)		(141.23)	(62.11)	(68.44)
lag (2) corruption			-0.240***	-0.246***			-1.197***	-1.282***
			(-6.79)	(-6.05)			(-20.05)	(-21.83)
lag (3) corruption			0.0558*	0.0510			0.449***	0.473***
			(1.87)	(1.66)			(7.74)	(8.32)
lag (4) corruption			-0.0351*	0.0204			-0.119***	-0.0772***
			(-1.92)	(1.09)			(-3.79)	(-2.89)
Observations	3581	3443	3032	3032	4237	4063	3544	3544
Countries	138	137	137	137	174	173	173	173
Years	27	26	23	23	26	25	22	22
Country FE	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Time FE	Yes	Yes	Yes	No	Yes	Yes	Yes	No

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: Measure of corruption indicated in the column heading. Independent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. Standard errors are clustered by country.

Table 18: Dynamic instrumental variable model estimates for the effect of corruption in the legislature on enactment of legislative gender quotas, countries at or below the 10th percentile of mean log GDP per capita in the data set, no clustering on country

	FE IV	System DPD	Diff. DPD
lag corruption	0.0271	0.00290	0.0188
	(1.41)	(0.44)	(1.01)
lag presence of legislated quota	0.847***	0.969***	0.821***
	(15.33)	(59.88)	(29.33)
Observations	367	422	393
Countries	19	19	19
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Sargan/Hansen test stat.	0.0456	381.3	387.4
Sargan/Hansen test p-value	0.831	0.0171	0.000422
1st stage F-stat (Kleibergen-Paap)	91.16		
Endogeneity test stat.	0.797	353.9	357.3
Endogeneity test, p-value	0.372	1.000	0.996

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of legislative corruption score as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors for FE IV model are robust to arbitrary heteroskedasticity.

Table 19: Dynamic instrumental variable model estimates for the effect of corruption in the legislature on enactment of legislative gender quotas, countries at or below the 15th percentile of mean log GDP per capita in the data set, no clustering on country

	FE IV	System DPD	Diff. DPD
lag corruption	0.0243	-0.00571	0.0241
	(1.48)	(-0.82)	(1.38)
lag presence of legislated quota	0.839***	0.965***	0.757***
	(16.15)	(66.86)	(28.38)
Observations	538	617	576
Countries	27	27	27
Years	24	26	25
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Sargan/Hansen test stat.	0.0453	473.8	471.4
Sargan/Hansen test p-value	0.831	0.00118	0.0000330
1st stage F-stat (Kleibergen-Paap)	113.3		
Endogeneity test stat.	0.834	374.7	373.3
Endogeneity test, p-value	0.361	0.0227	0.0123

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: presence of a legislated gender quota (legal candidate mandates or reserved seats) for the lower or sole house of parliament. FE IV model uses second and third lag of legislative corruption score as excluded instruments. System and difference DPD models use all available exogenous lags as instruments. Standard errors for FE IV model are robust to arbitrary heteroskedasticity.

Table 20: Fixed effects instrumental variable model estimates for the effect of legislative gender quotas on V-Dem measures of corruption

	Leg	Pub	Exec	Jud	Overall
presence of legislated quota	-0.0207	0.000614	0.00157	-0.00163	-0.000166
	(-1.18)	(0.19)	(0.40)	(-0.12)	(-0.07)
lag corruption	0.836***	0.868***	0.864***	0.852***	0.887***
	(33.59)	(46.71)	(48.97)	(47.27)	(52.54)
Observations	4096	4263	4263	4259	4251
Countries	173	173	173	173	173
Years	25	25	25	25	25
Country FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Hansen's J	0.730	1.640	6.290	1.817	4.535
Hansen's J p-value	0.393	0.200	0.0121	0.178	0.0332
1st stage F-stat (Kleibergen-Paap)	9599.3	13404.0	13200.8	13768.7	13547.4
Endogeneity test stat.	1.019	0.296	0.222	3.302	0.746
Endogeneity test p-value	0.313	0.587	0.637	0.0692	0.388

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: V-Dem measure of corruption indicated in the column heading. FE IV model uses first and second lag of quota as excluded instruments. Standard errors are clustered by country.

Table 21: Dynamic panel data (level) model estimates for the effect of of legislative gender quotas on V-Dem measures of corruption

	Leg	Pub	Exec	Jud	Overall
presence of legislated quota	-0.00268	-0.000121	-0.0000880	-0.00608	-0.000901
	(-0.20)	(-0.07)	(-0.04)	(-0.74)	(-0.59)
lag corruption	0.963***	0.993***	0.989***	0.990***	0.999***
	(73.32)	(245.99)	(158.23)	(236.39)	(249.05)
Observations	4251	4435	4435	4431	4422
Countries	173	173	173	173	173
Years	26	26	26	26	26
Country FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Hansen's J	150.8	176.0	164.1	145.3	168.1
Hansen's J p-value	1	1	1	1	1
Endogeneity test stat.	74.61	100.1	86.56	70.54	92.98
Endogeneity test p-value	1	1	1	1	1

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: V-Dem measure of corruption indicated in the column heading. All available exogenous lags are used as instruments. Standard errors are clustered by country.

Table 22: Dynamic panel data (difference) model estimates for the effect of of legislative gender quotas on V-Dem measures of corruption

	Leg	Pub	Exec	Jud	Overall
presence of legislated quota	-0.0931	-0.0110	-0.0170*	-0.0594*	-0.0129*
	(-1.35)	(-1.51)	(-1.80)	(-1.93)	(-1.73)
lag corruption	0.578***	0.850***	0.805***	0.815***	0.759***
	(6.95)	(18.24)	(22.91)	(25.20)	(18.43)
Observations	4050	4262	4262	4258	4249
Countries	173	173	173	173	173
Years	25	25	25	25	25
Country FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Hansen's J	141.4	161.9	164.0	147.6	165.2
Hansen's J p-value	1	1	1	1	1
Endogeneity test stat.	-12.08	18.35	7.890	2.133	7.552
Endogeneity test p-value	1	1	1.000	1	1.000

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: V-Dem measure of corruption indicated in the column heading. All available exogenous lags are used as instruments. Standard errors are clustered by country.

Table 23: Model estimates for the effect of legislative gender quotas on corruption in the legislature, countries above 1.4 on the V-Dem gender power index

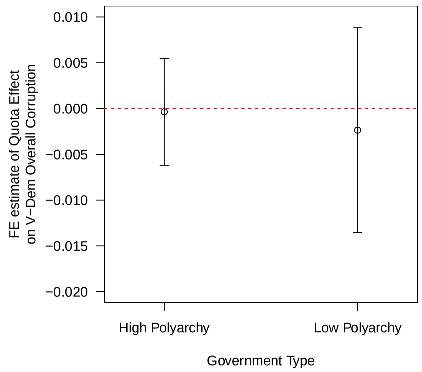
	FE w/ lag	FE only	lag only
lag corruption	0.864***		0.992***
	(29.71)		(323.80)
presence of legislated quota	-0.0651**	-0.232**	-0.0187
	(-2.52)	(-2.37)	(-1.57)
Observations	1373	1477	1373
Countries	81	86	81
Years	26	27	26
Country FE	Yes	Yes	No
Time FE	Yes	Yes	No

t statistics in parentheses

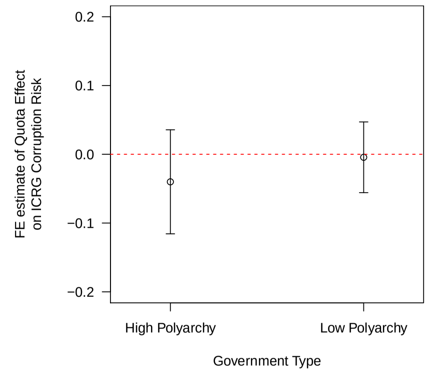
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: V-Dem measure of corruption in the legislature. Standard errors are clustered by country.

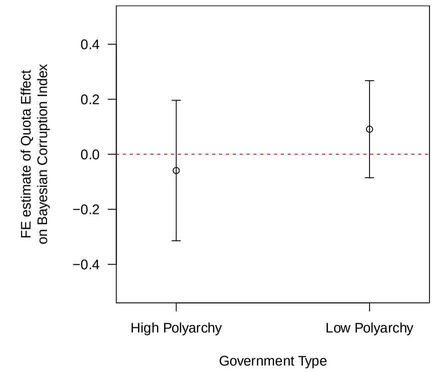
Figure 8: Estimated effect of legislative gender quotas on other measures of corruption, by democracy score



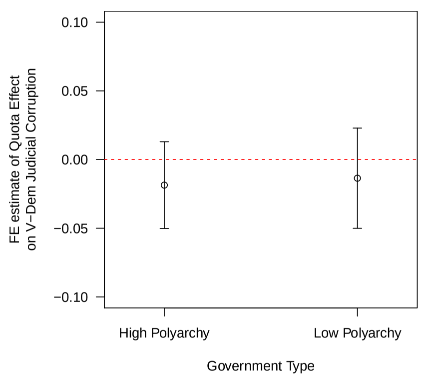
(a) V-Dem Overall Corruption



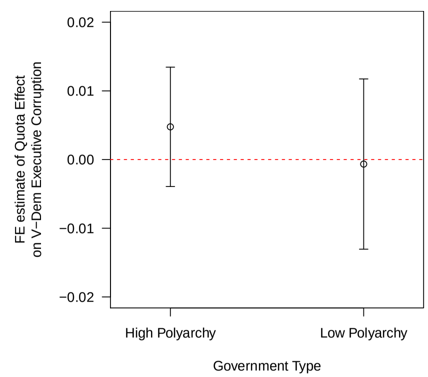
(b) ICRG Corruption Risk



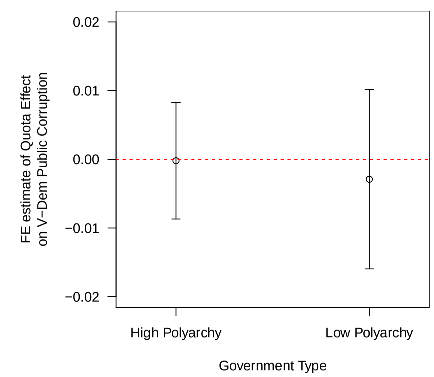
(c) Bayesian Corruption Index



(d) V-Dem Judicial Corruption



(e) V-Dem Executive Corruption



(f) V-Dem Public Sector Corruption

Each point represents the estimate of the effect of legislative gender quotas on the corruption measure listed in a subset of the data with V-Dem electoral democracy score as noted on the x-axis (“high” indicates above the sample median while “low” indicates less than or equal to the sample median). The estimate comes from a fixed effects model with a single lag of the dependent variable country and year FEs (as in equation 2). 95% confidence intervals for each estimate are indicated by bars and based on standard errors clustered by country.

Table 24: Model estimates for the effect of legislative gender quotas on V-Dem overall corruption in government, countries above 1.4 on the V-Dem gender power index

	FE w/ lag	FE only	lag only
lag corruption	0.835***		0.993***
	(29.35)		(382.14)
presence of legislated quota	-0.00779*	-0.0284**	-0.00335*
	(-1.91)	(-2.06)	(-1.98)
Observations	1382	1484	1382
Countries	81	86	81
Years	26	27	26
Country FE	Yes	Yes	No
Time FE	Yes	Yes	No

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: V-Dem measure of overall corruption. Standard errors are clustered by country.

Table 25: Model estimates for the effect of legislative gender quotas on ICRG Corruption Risk, countries above 0.5 on the V-Dem gender power index

	1	2	3	4	5
lag corruption	0.801***	0.941***		0.961***	1.107***
	(41.83)	(27.83)		(161.87)	(37.34)
lag (2) corruption		-0.193***			-0.193***
		(-5.08)			(-4.54)
lag (3) corruption		0.0140			0.00718
		(0.40)			(0.20)
lag (4) corruption		-0.0225			0.0373*
		(-0.97)			(1.79)
presence of legislated quota	-0.0745**	-0.0636**	-0.103	-0.0186	-0.0142
	(-2.40)	(-2.05)	(-1.03)	(-1.01)	(-0.77)
Observations	2366	2022	2491	2366	2022
Countries	111	106	113	111	106
Years	26	23	27	26	23
Country FE	Yes	Yes	Yes	No	No
Time FE	Yes	Yes	Yes	No	No

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: ICRG Corruption Risk. Standard errors are clustered by country.