Estimating Causal Relationships Between Women's Representation in Government and Corruption*

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

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

January 30, 2019

Abstract

Does increasing the representation of women in government lead to less corruption, or does corruption prevent the election of women? Are these effects large enough to be substantively meaningful? Some research suggests that having women in legislatures reduces corruption levels, with a variety of theoretical rationales offered to explain the finding. Other research suggests that corruption is a deterrent to women's representation because it reinforces clientelistic networks that privilege men. Using instrumental variables, we find strong evidence that women's representation decreases corruption *and* that corruption decreases women's participation in government; both effects are substantively significant.

*Our thanks to: participants at the Gender and Corruption Workshop at the University of Gothenburg (May 23-24, 2016), especially workshop organizers Helena Stensöta and Lena Wängnerud; participants at the 2018 Comparative Politics Annual Conference (CPAC) at Washington University in St. Louis, especially our discussant Miguel Pereira; participants at the 2018 Political Economy of Corruption and Accountability Workshop at the Instituto Carlos III-Juan March in Madrid, Spain, especially our discussant Catherine de Vries; and participants at a colloquium in the Department of Politics and International Affairs at Wake Forest University, especially Sara Dahill-Brown. All provided helpful comments on previous versions of this paper.

Introduction: Identifying causality in the relationship between women's representation and corruption

At present, there is widespread scholarly agreement that higher representation of women in government is empirically connected to lower corruption in that government. However, there is much less agreement about why this relationship exists or the causal direction in which it flows. The extant research makes two arguments about the direction of causality. One argument is that reduced corruption leads to more women being represented in politics because corrupt elites keep women out of government in order to preserve the spoils of corruption for themselves and maintain the security of their corruption networks (Bjarnegård, 2013; Grimes and Wängnerud, 2012; Stockemer, 2011; Sundström and Wängnerud, 2014). The other argument is that the presence of women in government leads to less corruption; the rationales offered to explain this second argument are myriad. These rationales range from an essentialist position that women are more honest than men (Dollar, Fisman and Gatti, 2001) to the hypothesis that women are excluded from opportunities to wield power (Branisa and Ziegler, 2011; Goetz, 2007; Tripp, 2001; Heath, Schwindt-Bayer and Taylor-Robinson, 2005; Barnes, 2016; Schwindt-Bayer, 2010, 2018) to the theory that women are treated differently by voters or are more risk averse than their male counterparts and thus engage in corruption less often (Alatas et al., 2009; Esarey and Chirillo, 2013; Esarey and Schwindt-Bayer, 2018).

Unfortunately, although most of the empirical research in this field shows a strong correlation between women's representation and corruption, little of this research is designed to directly identify the direction of causality or the magnitude of the effect in each direction. Without even the basic question of directionality resolved, it is difficult to further refine our theories using evidence. Given that we do not fully understand how or why the link between gender and corruption exists, it is unsurprising that policies increasing women's representation in various government positions have had an inconsistent effect on corruption (see, for example Moore, 1999; Quinones, 1999; McDermott, 1999; Karim, 2011; Kahn, 2013). More precise knowledge of the mechanisms by which gender and corruption interact would presumably make such interventions more reliably successful.[1]

In this paper, we use instrumental variable models to determine which way causality flows in the relationship between gender and corruption. Our goal is to establish whether (a) women's representation affects corruption, (b) corruption affects women's representation, or (c) both. This goal is relatively modest, as it leaves many micro-level questions unaddressed. However, it is an important goal because it allows future work to focus on studying those theories that predict a causal relationship validated by evidence.

Instrumental variable models allow the identification of causal relationships via an *instrument* that influences the strength or presence of the independent variable being studied but does not itself directly cause the outcome (except through its effect on the independent variable); an instrument in an observational study plays a role analogous to that of random assignment in a laboratory experiment (Angrist and Pischke, 2009, Chapter 4). We instrument women's representation with two variables—female enrollment in secondary schools and female labor force participation—to determine how much an increase in women's representation in the lower house of parliament changes corruption. To examine how much an increase in corruption lowers women's representation, we instrument corruption with two measures: ethnolinguistic fractionalization and political stability. As a robustness check and a second identification strategy, we also instrument women's representation and corruption with their first and second lags (Reed, 2015). Using all of these different approaches, we gain leverage on causality in the women's representation and corruption relationship. We employ a dataset that includes 76 partial and full democracies with annual data from 1990-2010.

Our overall conclusion is that greater representation of women in the lower house of

[1] This argument is also made in Esarey and Chirillo (2013, p. 364).

parliament causes decreased corruption, *and* greater corruption in government causes lower representation of women in government. Both relationships are statistically significant and substantively strong across a variety of different model specifications, although some models find smaller and statistically less certain relationships.

We believe that this paper's main lesson is that both arguments in the literature—women's representation lowers corruption, and corruption reduces women's representation—are empirically supported. This lesson is important because it suggests that the theoretical frameworks supporting both of these arguments need to be pursued in future work, and also that any future work must take care to use a research design that clearly identifies the direction of causality in the relationship.

Theories of causality in the relationship between women's representation and corruption

A growing literature has explored the empirical relationship between women's representation and corruption and developed a correspondingly rich theoretical discussion about why this relationship may exist (see, for example, Dollar, Fisman and Gatti, 2001, Swamy et al., 2001, Sung, 2003, Alhassan-Alolo, 2007, Wängnerud, 2012, Esarey and Chirillo, 2013, Barnes and Beaulieu, 2014, Watson and Moreland, 2014). From this work, we may infer that there is a substantively meaningful empirical connection between women's representation and corruption. However, the causal direction in which the relationship operates is still unclear. There are theoretical frameworks supporting each direction of causality, and each of these frameworks hosts debates about the precise causal mechanism in operation.

We cannot hope to resolve all the debates within each framework in this paper. However, each framework makes a big prediction: either increases in women's representation in government reduce corruption, or increases in corruption reduce the representation of women in government. The aim of this paper is to validate or invalidate these big predictions using observational data. In this section, we briefly review the theoretical frameworks supporting each prediction.

The effect of women's representation on corruption

The initial studies of women's representation and corruption either avoided committing to a theory (Swamy et al., 2001) or suggested that women are more honest and trustworthy than men and therefore less likely to be corrupt (Dollar, Fisman and Gatti, 2001). However, this essentialist argument found little empirical support in follow-up work; research has repeatedly shown that the link is context-dependent (Goetz, 2007; Sung, 2003; Alatas et al., 2009; Esarey and Chirillo, 2013; Esarey and Schwindt-Bayer, 2018), a finding inconsistent with the essentialist theory.

Some papers have argued that women's representation reduces corruption because women have less opportunity to engage in corruption as a result of being excluded from power and patronage (Branisa and Ziegler, 2011; Goetz, 2007; Tripp, 2001). Research on women in legislatures has shown that women have been marginalized in politics, and even where they get elected to office, they do not have access to the same leadership posts, committees, bills, and other legislative spoils as men (Heath, Schwindt-Bayer and Taylor-Robinson, 2005; Barnes, 2016; Schwindt-Bayer, 2010, 2018). Other research has shown that male-dominated patronage networks exclude women (Beck, 2003; Bjarnegård, 2013; Arriola and Johnson, 2014). If women do not have access to corrupt networks, then this would imply that the presence of more women means fewer corrupt politicians in government and thus less corruption. This argument assumes that the increased presence of women will crowd out corrupt networks and reduce corruption. It is very possible, however, that those corrupt networks will grow more skilled at operating at high levels with a reduced number of politicians leave corruption levels as they are. Alternatively, corrupt networks could bring more women into the fold over time as women's increased number and seniority garners them more access to powerful leadership posts that they were once kept out of. Indeed, one study exploring women's representation and corruption found that political opportunity had no effect on the relationship (Torgler and Valev, 2010).

Yet another set of studies focuses on the risk associated with engaging in corruption. Esarey and Schwindt-Bayer (2018) argue that women's representation is likely to lead to reduced corruption, but only in contexts where corruption is particularly risky. They offer two reasons for this. First, women may be disproportionately punished by voters at the polls for engaging in corruption (at least where corruption is stigmatized by law and custom) because doing so violates stereotypes about women in office. Anticipating greater punishment, women may be more likely to refrain from entering into corrupt transactions while in office. Unfortunately, the evidence for this proposition is mixed (Schwindt-Bayer, Esarey and Schumacher, 2018; Barnes and Beaulieu, 2018; Benstead and Lust, 2018; Eggers, Vivyan and Wagner, 2018; Benstead, Jamal and Lust, 2015; Żemojtel-Piotrowska et al., 2016). Second, a great deal of research shows that women are more risk averse than men (Byrnes, Miller and Schafer, 1999; Bernasek and Shwiff, 2001; Sundén and Surette, 1998; Watson and McNaughton, 2007; Croson and Gneezy, 2009; Eckel and Grossman, 2008). If this is true, then women should be less likely to engage in corruption when it is risky even if the objective probability of detection and/or the severity of punishment does not differ by gender.

A growing body of empirical research supports the contention that women's representation leads to reduced corruption but only in environments where corruption is risky. Esarey and Chirillo (2013) show that greater female participation in legislatures is only associated with less corruption in democracies, but not in autocracies. Esarey and Schwindt-Bayer (2018) showed that the relationship between women's representation and corruption is conditional upon electoral accountability, measured as contexts where corruption norms are absent, in parliamentary systems, where the press operates freely, and in more personalistic electoral systems. Another study conducted an experiment designed to measure willingness to engage in corruption and found that women are less susceptible to corruption in some countries but not in others (Alatas et al., 2009). Similar findings exist at the micro level as well. Examining the World Values Survey data, Esarey and Chirillo (2013, p. 372) find that "there is little difference in corruption tolerance between men and women for countries that rank lowest on the Polity scale [viz., autocracies]. In more democratic countries, however, men are considerably more tolerant of corruption than women." Similarly, Schwindt-Bayer (2010) finds no relationship between women's representation in legislatures and citizens' perceptions of corruption in government in Latin American countries using mass survey data from the Americas Barometer (LAPOP).

Although most of the studies above were not explicitly designed to eliminate simultaneity and endogeneity, a few studies have employed research designs that aim to identify the causal effect of women's representation on corruption. One study found a causal relationship between women's representation and corruption in experimental evidence that female candidates cause reduced perception of election fraud compared to men (Barnes and Beaulieu, 2014). In addition, two papers have used instrumental variables to try to establish a causal relationship between women's representation and corruption; however, the instruments they employ are not beyond question. Correa Martínez and Jetter (2016) employ plow usage two centuries ago to instrument female labor force participation (which is often strongly related to women's representation in politics), but their instrument requires us to believe that countries' differences on gender inequality two hundred years ago mirror those of today. Jha and Sarangi (2018) use an instrumental variable model to study the effect of women in parliament on corruption by instrumenting women in parliament with the year that women attained suffrage in that country and the years since transition to agricultural society. However, the recent adoption of gender quotas in countries over the past twenty years has largely broken any correlation that existed between suffrage and the representation of women in parliament; the agricultural instrument is predicated on a speculative theoretical argument about how prehistoric transition to farming led to division of labor by sex. Thus, it is prudent to study the causal relationship between women's representation and corruption using a new (and hopefully more robust) set of instruments, with particular attention to whether these instruments are valid.

The effect of corruption on women's representation

At the same time that scholars were exploring the reasons why women's representation might lead to reduced corruption, another set of studies was exploring the opposite phenomenon: corruption may be a deterrent to women's representation in politics. The primary argument made in this literature is that networks of corrupt officials suppress women's representation in government as a means of ensuring that outsiders do not penetrate these networks and disrupt the stream of benefits from corruption (Bjarnegård, 2013; Grimes and Wängnerud, 2012; Stockemer, 2011; Sundström and Wängnerud, 2014).

Stockemer (2011, p. 697) argues that corruption hurts women's election chances because it "perpetuates gender inequalities, reinforces traditional networks and prevents women from gaining human and financial resources." With a statistical analysis of African countries, he shows a significant correlation between women's representation and corruption. In slightly stronger terms, Sundström and Wängnerud (2014, p. 355) argue that the political recruitment of women is more difficult in clientelistic or corrupt societies because women are more likely to be excluded from the male-dominated networks from which candidates are selected. Specifically, they highlight "the presence of shadowy arrangements that benefit the already privileged, which in most countries tend to be men" and suggest that this is the reason why they find fewer women represented in European local councils with higher corruption. Bjarnegård (2013) substantiates her argument about corruption reinforcing clientelistic relationships among men that exclude women with compelling qualitative evidence from Thailand.

The quantitative evidence presented in these studies is (for the most part) focused on the empirical association between corruption and female representation. Generally speaking, their research designs have a limited ability to explicitly identify how the proportion of women in government will change when corruption changes. Thus, as with the literature studying the effect of women's representation on corruption, more analysis is needed to establish causality in the opposite direction too.

Hypotheses

We have presented arguments that women's representation may influence corruption and that corruption may influence women's representation and highlighted the most common explanations for why the relationship may go in each direction. This leads to two clear causal hypotheses to test in this paper. First, we hypothesize that greater women's representation in legislatures should cause reduced corruption. Second, we hypothesize that higher corruption levels should cause lower women's representation in legislatures.[2]

Importantly, these hypotheses are not mutually exclusive. The theoretical logic for why women's representation may lead to less corruption is distinct from the logic for why corruption may affect women's representation. It is entirely possible that, as Grimes and Wängnerud (2012, p. 26) put it, "causation runs in both directions." We allow for finding support for both hypotheses in our empirical analyses.

[2]Esarey and Chirillo (2013) and Esarey and Schwindt-Bayer (2018) argue that electoral accountability to voters is a key reason why female representatives are disproportionately deterred from engaging in corruption. If this is true, we may expect to find that increased representation of women in government only lowers corruption in the most consolidated democracies, where electoral accountability is strongest. We test this hypothesis in an appendix; the results are inconclusive, as our findings depend on modeling choices where the best choice is uncertain.

An instrumental variable model of women's representation and corruption

An instrumental variable model offers a useful method for exploring the direction of causality in the women's representation and corruption relationship. Specifically, it allows us to estimate the Local Average Treatment Effect (LATE) of women's representation on corruption (and the LATE of corruption on women's representation).[3] The LATE is the extent to which a change in the independent variable causes a change in the dependent variable for the subset of cases whose value of the independent variable is influenced by the instrument. For example, if we use female enrollment in secondary education as an instrument for women's participation in government, the corresponding IV model will estimate the change in corruption caused by a change in women's participation in government only among those states whose female representation in government is actually influenced by female enrollment in secondary education.[4]

Our instrumental variable model requires that we find instruments for both women's representation in legislatures and corruption. To explore the effect of women's representation on corruption, we need an instrument for the percentage of women in the legislature that is associated with women's representation but not with corruption (except through its effect on women in government). To determine the causal relationship in the opposite direction, the effect of corruption on women's representation, we need an instrument for corruption that is associated with corruption but not with the percentage of women in government (except

[3]The assumptions necessary to sustain this interpretation of an instrumental variable model with a continuous treatment condition are described in Angrist and Imbens (1995); see also Angrist, Imbens and Rubin (1996).

[4]The cases whose value of the independent variable is changed by manipulating the instrument are sometimes called *compliers* (Angrist, Imbens and Rubin, 1996). This terminology is linked to LATE's usefulness as an estimate of the degree of change in the dependent variable that can be prompted by a policy action: policy changes are an instrument that causes a change in some independent variable, and the "compliers" are those units who will actually respond to (or comply with) the policy change via a corresponding change in the independent variable (Esarey, 2017).

through its effect on corruption).

Instruments

We propose two different sets of instruments to separately identify our causal effects of interest. The first set of instruments are observable variables that we believe are likely to be closely associated with the target independent variable, but to have no alternate causal pathways to the dependent variable that cannot be blocked by control variables. For determining the effect of women's representation on corruption, we use two instruments:

1. female enrollment in secondary education; and

2. the proportion of women in the labor force.

Theoretically, these instruments are linked to women's representation but not directly to corruption. Women's representation has been influenced by "incremental mechanisms," such as cultural and socioeconomic changes focused on getting women into the pool of potential candidates for office. Female enrollment in secondary education and the involvement of women in the labor force have been identified in many global studies as two of the most important indicators of societal changes that have helped make women viable candidates (Rule, 1981; Norris, 1985; Oakes and Almquist, 1993; Kenworthy and Malami, 1999; Paxton and Kunovich, 2003; Schwindt-Bayer, 2005; Paxton and Hughes, 2007).[5] Our research design

[5] Of course, any choice of instruments is going to be debatable and it is not surprising to find other scholars making different choices in the literature. As one example, Uslaner and Rothstein (2016) argue that the mean years of schooling in a country in 1870 is a predictor of corruption in 2010. Furthermore, they argue that education and corruption are endogenously related; they instrument for education using (a) the share of the population that was Protestant in 1980, (b) the share that was European in 1900, and (c) whether the country is a former colony. However, Uslaner and Rothstein (2016) do not explicitly state precisely *why* education levels in 1870 are endogenous with present-day corruption. As corruption in 2010 cannot directly cause education levels in 1870, simultaneity is presumably not a concern. Confounding is the primary other reason why one might instrument; that is, there may be a third variable that explains both education in 1870 and corruption in 2010. For our purposes, the potential concern is that schooling in 1870 might be a confounding variable that causes both contemporary female school enrollment and corruption. However, this potentially confounding pathway is blocked in our models via the inclusion of region and country fixed effects, which absorbs the effect of any variable that is constant within countries (such as schooling in 1870).

requires that exogenously imposed changes in either of these variables will cause changes in the representation of women in government, which would in turn cause changes in government corruption, and that no other unblocked pathway from the instruments to corruption exists.[6] In certain subsets of countries (like Latin America, Sub-Saharan Africa, and developing countries), research has not always found strong relationships between women's education and workforce levels and women's representation (Yoon, 2004; Lindberg, 2004; Schwindt-Bayer, 2010; Fallon, Swiss and Viterna, 2012). Yet these studies focus on regions or groupings of countries where socioeconomic differences are minimal across countries and women have often taken different paths to power. Globally, where wider variation in these factors exists, a relationship is evident (particularly when these regions and groupings of countries are included).

One potential concern with these instruments is that systematic differences between regions or countries might change both corruption and our instruments. For example, increased economic development (itself endogenously related to corruption) may increase women's education, creating a potential spurious relationship. As a consequence, we add region and country fixed effects to block confounding influences that are constant within units.[7] We also assess the robustness and validity of our results using diagnostic tests for instrument

[6]Sara Dahill-Brown noted that it is possible that both of these instruments may be correlated with other forms of women's participation that may in turn affect corruption. For example, greater participation by women in the labor force is probably also associated with a greater proportion of women in corporate executive positions, in the bureaucracy, and so on. These provide a potential pathway to corruption that does not lead through women's participation in the legislature specifically. However, all these pathways *do* lead to women's overall representation in leadership positions. Women's representation in parliament may be considered a measure of female leadership participation overall, and from this perspective the instruments are valid. However, it is important to note that our design's ability to separately identify the contributions of women's representation in parliament, in corporate leadership, in the bureaucracy, and so on to corruption is less than ideal.

[7]Several commenters at presentations of this work suggested that we control for the presence of gender quotas in the legislature to ensure that our instruments are not associated with corruption except through the presence of women in office. Consequently, we add a control for electoral quotas or reserved seats in the lower house of the legislature to the models of Table 1; the results are shown in Appendix Table 14. The magnitude and uncertainty of the effect of women's representation on corruption estimated in this model are similar to those found in Table 1.

validity.[8]

For determining the effect of corruption on the proportion of women in government, we instrument corruption with two variables:

1. ethnolinguistic fractionalization; and
2. political stability.

Ethnolinguistic fractionalization has a long history of being used as an instrument for corruption, dating back at least to Mauro (1995). As Mauro (1995) explains (on p. 693):

> A number of mechanisms may explain this relationship. Ethnic conflict may lead to political instability and, in extreme cases, to civil war. The presence of many different ethnolinguistic groups is also significantly associated with worse corruption, as bureaucrats may favor members of the same group. Shleifer and Vishny (1993) suggest that more homogenous societies are likely to come closer to joint bribe maximization, which is a less deleterious type of corruption than noncollusive bribe setting.

However, to our knowledge ethnolinguistic fractionalization has not been linked to women's representation. Following a similar logic, we also use political stability as an instrument for corruption. As Mauro (1995) says, decreased political stability directly lowers the efficiency of institutions and creates room for corruption. It also decreases the "shadow of

[8] In order to sustain the LATE interpretation, it is also necessary that our instruments be monotonically related to the treatment of interest. That is, for units i, a treatment variable T, and binary instrument Z, $T_i(Z=1) \geq T_i(Z=0)$ with strict equality in at least one i (Angrist, Imbens and Rubin, 1996, 434-435); this interpretation can be carried through to multivalued instruments by interpreting these instruments as multiple orthogonal binary variables (Angrist and Imbens, 1995). Although this assumption cannot be directly tested, as it involves counterfactual quantities, Angrist and Imbens (1995) proposes an indirect test by examining the cumulative distribution function for the treatment variable for $Z=0$ and $Z=1$; if the monotonicity assumption holds, then these CDFs should not cross. In Appendix Figures 4 and 5, we examine the CDFs of ICRG corruption score and the percentage of women in parliament separately for those observations below and above the median value of each of our proposed (non-lag) instruments. Two of our instruments' CDFs do not cross at all, while the other two have minimal overlap near one edge of the distribution of the treatment variable.

the future" needed to sustain cooperative agreements and may lead to less efficient forms of corruption.[9] Some research (e.g., Hughes, 2009; Hughes and Tripp, 2015) has argued that, in the aftermath of political crises and conflict, women's representation increases as a result of international pressure to conform to gender equality norms or voters and elites seeking political newcomers to re-legitimize government. Yet, these studies focus on the impact of post-conflict or post-crisis periods rather than political instability during such periods (where we would not expect increased opportunities for women), and they apply these arguments to subsets of countries around the world where such crises have been common (African countries and low-income countries). We expect little relationship between political stability and women's representation, more generally.

As a robustness check on the instruments we propose above, we repeat our analysis with an entirely different set of instruments based on an identification strategy proposed by Reed (2015). This strategy assumes that, conditional on any controls (especially the lag of the dependent variable), lags in the independent variable only influence the current value of the dependent variable through their effect on the current value of the independent variable. This imposes restrictions on allowable dynamics within the model, but conditional on those restrictions the causal effect of an independent variable on a dependent variable can be identified despite simultaneity between the two. We therefore propose to use the first and second lags of the target independent variable as an instrument for the current value of that variable.

We estimate statistical models for these two sets of instruments separately to assess the robustness of our results to the validity of the assumptions that support our identification

[9] Indeed, Mauro notes that ethnolinguistic fractionalization and stability are intertwined: "Strictly speaking, the ELF index is a valid instrument only for the institutional efficiency index [in a regression examining the effect of institutions on economic growth], as fractionalization affects both corruption and political instability" (pp. 693-694). Ethnolinguistic fractionalization and political stability are related to one another, but we argue that their effect on women's representation in government comes solely through their effect on corruption.

strategy. Each set of instruments relies on different assumptions for correct identification of causal effects. Consequently, comparing our estimates across the two models will ensure that our results are not overly sensitive to these assumptions.

Data and Variables

Our data set contains 76 democratic-leaning countries observed over 21 years (from 1990 to 2010), though our models vary in spatio-temporal coverage according to the availability of the variables they include (and the panels are unbalanced due to some data missingness). Except where noted, the variables come from the time-series cross-national dataset compiled by Schwindt-Bayer and Tavits (2016). Democratic-leaning countries are those with a Freedom House average Civil Liberties and Political Rights score of 5 or lower (www.freedomhouse.org) and a Polity IV polity2 score of zero or more for twelve years or more (Marshall, Gurr and Jaggers, 2014).

Our key variables are measures of corruption perceptions and of the proportion of women in parliament.[10] We measure corruption perceptions with two common indices. The first is the Political Risk Services International Country Risk Guide's (ICRG) corruption risk measure, which runs annually from 1990-2010 and varies between 0 (least corruption) and 6 (most corruption). The second is the Transparency International Corruption Perceptions Index (TI CPI), which measures "the abuse of public office for private gain" (World Bank Group, 1997), annually from 1995 to 2010 and varies between 0 (least corruption) and 10 (most corruption).[11] We report details from our ICRG results in the main paper and summarize the TI CPI results in the text; detailed TI CPI results are presented in an appendix

[10]Measuring corruption as corruption perceptions is not perfect. Yet, it is the most common way in that the literature has measured a concept that involves informal and hidden behavior. Corruption perceptions emerge from corrupt interactions serving as a useful proxy for corruption.

[11]Note that the ICRG and TI CPI were originally coded so that higher values indicated less corruption; we have reversed the coding in our models.

to save space. The percentage of women in the lower house of the legislature comes from the Inter-Parliamentary Union (2012).

Our four non-lag instrumental variables come from the Quality of Government (QoG) dataset (Teorell et al., 2015). The gross enrollment ratio of females in secondary school is measured by UNESCO (UNESCO Institute for Statistics, 2014). The proportion of the total labor force that is female is measured by the World Bank's World Development Indicators (World Bank, 2014). An ethnolinguistic fractionalization index measures the probability that two randomly selected people in a state will not belong to the same group in the year 1985 (Roeder, 2014). Finally, a political stability estimate comes from the World Bank's Governance Indicators (Kaufmann, Kraay and Mastruzzi, 2010), a model-derived aggregate index that measures "perceptions of the likelihood that the government in power will be destabilized or overthrown by possibly unconstitutional and/or violent means, including domestic violence and terrorism" (Teorell et al., 2015, p. 532); the index varies between -2.39 and 1.67.

Statistical Modeling

Our time-series cross-sectional dataset presents special challenges to inference owing to its panel nature. To ensure that our results are not an artifact of unit or temporal heterogeneity in the data, we take four different approaches to the problem. First, we estimate cross-sectional models for each year's data separately and compare our results across years. Second, we estimate pooled models with no region, country, or year fixed effects. Third, we estimate models with fixed effects for region or country that control for spatial heterogeneity. Finally, we estimate a model with multiple lags of the dependent variable to account for dynamics within each panel.

For our cross-sectional analyses, we use two-stage least squares (2SLS) analysis. For the panel analyses, we a two-stage generalized method of moments (GMM2S) model. Both the 2SLS and GMM2S models are implemented in the ivreg2 package for Stata (Baum, Schaffer and Stillman, 2003, 2007). Two-stage GMM allows us to account for the potentially heteroskedastic nature of the panel data, including clustering on countries, in a way that ordinary 2SLS does not. For the panel analyses, we cluster our standard error estimates according to country (Cameron, Gelbach and Miller, 2008). Clustering on year is contraindicated due to the small number of years available in the data (Cameron, Gelbach and Miller, 2008; Angrist and Pischke, 2009, ch. 8; Esarey and Menger, 2017); however, we include year fixed effects as a part of the model to control for trends or other temporal shocks in the dataset.[12]

The dynamic panel model is a good fit for our lagged independent variable instruments, because (1) it is inadvisable to estimate a model including both country fixed effects and lags of the dependent variable in a dataset with a relatively short temporal window due to Nickell bias (Nickell, 1981; Judson and Owen, 1999), and (2) the presence of lags of the dependent variable as control variables in the model makes its assumptions more plausible. Conversely, our other instruments are a good fit for the fixed effects models, as the fixed effects terms in the model make it more plausible that alternative pathways between these instruments and the dependent variable have been blocked.

Our empirical results include several important diagnostic tests. The first is the Sargan/Hansen's J test for instrument validity (Baum, Schaffer and Stillman, 2007, pp. 481-483). This test establishes whether the "orthogonality conditions" needed for valid instruments (i.e., that the instruments are independent of the dependent variable, net of their impact on the instrumented independent variables) are met in the data. The null hypoth-

[12]At the request of a reviewer, we also estimated full-information maximium likelihood (FIML) models on the full system of equations using `sem` in Stata (for the ICRG corruption measure). These results are reported in Appendix Tables 4 and 5. These models support the conclusion that increases in women's representation lower corruption, but present conflicting results (depending on modeling choices) about the effect of increased corruption on women in government.

esis of this test is that the orthogonality conditions are valid; thus, a rejection of the null hypothesis indicates that at least one of the instruments is invalid. This test can only be performed when multiple instruments are available.

An F-statistic for the joint significance of excluded instruments is estimated for the first stage of each model (Baum, Schaffer and Stillman, 2007, pp. 489-491). This test establishes whether the variables being used to instrument for endogenous independent variable(s) in the second stage (and therefore "excluded" from that second stage) are jointly capable of predicting the endogenous variable. The rule of thumb proposed by Staiger and Stock (1997) is that this F-statistic should be 10 or more to ensure consistent estimates. We report two versions of the test:

1. the Cragg-Donald (1993) statistic proposed by Stock and Yogo (2005) that assumes identically and independently distributed error terms, and

2. the Kleibergen-Paap (2006) statistic proposed by Baum, Schaffer and Stillman (2007) that allows for non-IID error terms.

Finally, we conduct a test for endogeneity (Baum, Schaffer and Stillman, 2007, pp. 481-483). The null hypothesis of the test is that the variable is exogenous; thus, a rejection of the test indicates that the independent variable must be treated as endogenous. We report the results of these tests for every model we estimate in the accompanying table.

Empirical Results

We have two sets of results to present. Our results are IV/2SLS and IV/GMM2S estimate of the LATE for:

1. women's representation on corruption, and

2. for corruption on women's representation in government.

We present each set of results separately.

LATE of women's representation on corruption

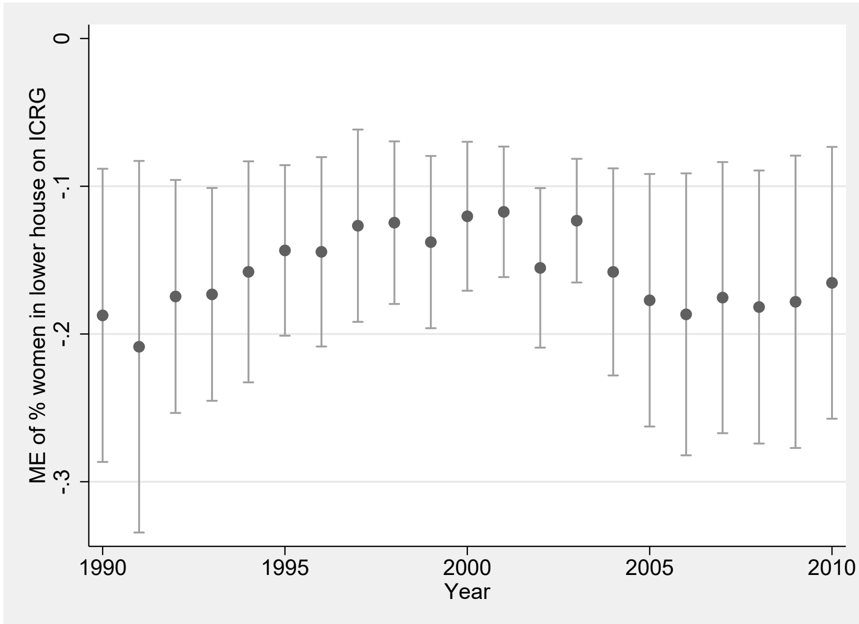
We begin by estimating the LATE of increased women's representation in government on corruption levels. Figure 1 shows the results of two-stage least squares models run separately on each year of data; for these cross-sectional models we use the gross enrollment ratio of females in secondary school and the proportion of females in the labor force as instruments. Panel 1a shows estimates of the marginal effect of women's representation on ICRG corruption score with 95% confidence intervals; panel 1b shows the results of the Sargan test of instrument validity and first-stage F-test for significance of the instruments.

The cross-sectional models show a relatively stable, negative marginal effect of increased women's representation on ICRG corruption risk score. The relationship varies between -0.117 and -0.209, with a mean of -0.158. The mean effect corresponds to a substantively important effect of women's representation on corruption: a 10 percentage point increase in women's representation causes a 1.58 point decline in ICRG score, corresponding to 22.6% of the maximum possible change in this corruption measure. This is roughly the equivalent of the difference in corruption between Guatelmala (average ICRG score in our data set = 3.54) and Hungary (average ICRG score in our data set = 2.01).[13] Although some of the Cragg-Donald F-statistics fall below the guideline value of 10 put forward by Staiger and Stock (1997), all are above 5. Eighteen of the twenty-one Sargan tests support the validity of the instruments using an $\alpha = 0.05$ test.

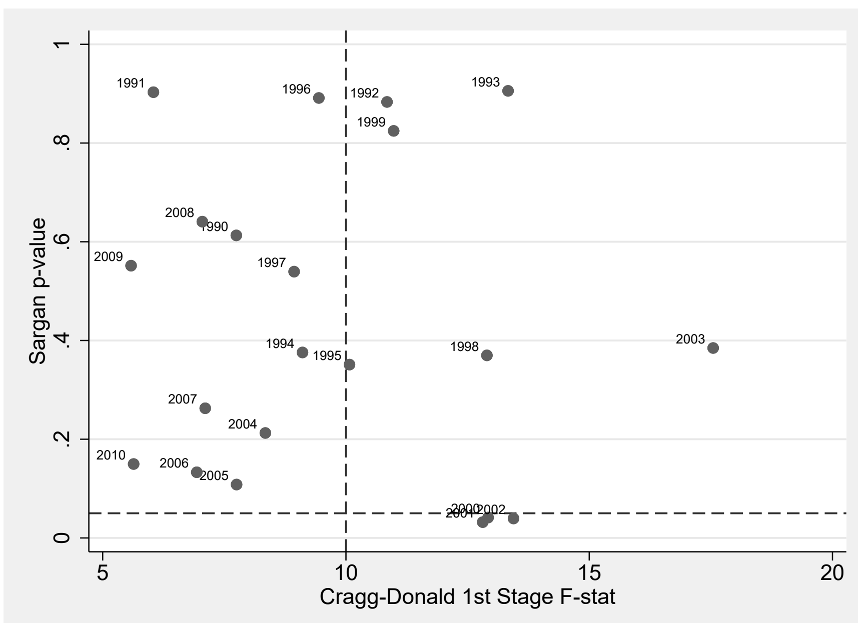
Results for similar models using the TI CPI dependent variable are shown in Appendix Figure 6 and yield similar results. The coefficient in a cross-sectional 2SLS model by year varies between -0.267 and -0.403 with a mean value of -0.329. This coefficient indicates

[13] All average figures for the ICRG score in this section are taken from the sample used to estimate Model 2 in Table 1.

Figure 1: IV/2SLS Estimates of Marginal Effect of Women's Representation on ICRG Corruption, with 95% Confidence Intervals



(a) Marginal Effects



(b) Sargan / F-Statistics

Marginal effect estimates in panel 1a and Sargan / F-statistics in panel 1b are from cross-sectional two-stage least squares models predicting ICRG corruption score using % women in the lower house of the legislature for each year of data between 1990 and 2010. Instrumental variables are gross enrollment ratio of females in secondary school and proportion females in the labor force.

that a 10 percentage point increase in women's representation in the lower house lowers the corruption score by an average of 3.3 points, 33% of the maximum possible change in the measure. This change is similar to the difference in corruption between Guatemala (average TI CPI score = 7.14) and Estonia (average TI CPI score = 3.88).[14] All the models support the validity of the instruments using the Sargan test, and five of sixteen first stage F-statistics are larger than ten and all but one are larger than five.

Our panel model results for the ICRG Corruption Risk variable are reported in Table 1, including dynamic models using the lag-based instruments.[15] All our models indicate that increased women's representation will decrease corruption, although the relationship is statistically insignificant at conventional levels when country-level fixed effects are included. For our model with region-level and year fixed effects (Model 2 in Table 1), an increase of 10 percentage points in women's representation in government causes an 0.911 point decline in ICRG score, about 13% of the maximum change possible. This is roughly the equivalent of the difference in corruption between Guatemala (average ICRG score = 3.54) and Botswana (average ICRG score = 2.64).

The models with lag-based instruments for women's representation report a statistically significant, but smaller, effect of women's participation in government on corruption. In model 4, a 10 percentage point increase in women's representation in the lower house of parliament causes an instantaneous 0.0657 point decline in the ICRG index, just under 1% of its maximum span. However, this instantaneous effect is multiplied over the long run through its effect on the lag values of the dependent variable in future periods (Keele and Kelly, 2006). The long run effect of a variable x is measured by:

$$LRM = \frac{\beta_x}{(1 - \sum_{j=1}^{T} \beta_{y_{(t-j)}})} \tag{1}$$

[14]Average figures for the TI CPI score in this section are taken from the sample used to estimate Model 2 in Appendix Table 6.

[15]The estimates corresponding to the first stage of the model are in Appendix Table 10.

In the long run, a 10% increase in women's representation causes a 0.751 point decline in the ICRG index ($p < 0.001$), about 11% of the maximum change possible. This is substantively close to the difference measured in Model 2 of Table 1.

Results for panel models using the TI Corruption Perception Index (reported in Appendix Table 6) are similar (though not identical) to the results for the ICRG variable in Table 1. The effect of women's representation in the lower house on the TI CPI is positive, but substantively small and statistically insignificant when country fixed effects are used. However, the instruments for the country fixed effects model are also weak according to the first stage F-statistics. In the model with lagged instruments, the instantaneous coefficient is statistically insignificant but the LRM indicates that a 10% increase in representation causes a 1.02 point decline in the TI CPI ($p = 0.029$).

LATE of corruption on women's representation in government

We now move to presenting the local average treatment effect of women's representation in the lower house of parliament on corruption. Figure 2 shows the results of two-stage least squares models run separately on each year of data. Note that, because the World Bank Governance Indicator's estimate of political stability is only available starting in 1996 and only measured biannually before 2002, estimates of the marginal effect do not exist before 1996 or for the years 1997, 1999, and 2001.[16] Panel 2a shows the estimates of marginal effects

[16]The limited availability of data for certain years leads us to explore alternative instruments available for the full set of years between 1990 and 2010. We substitute dummy variables for Spanish, British, and French colonial origin of Hadenius and Teorell (2007) as catalogued in the Quality of Government dataset (Teorell et al., 2015) for the original political stability instrument. The idea behind this instrument is that these countries established different institutions in their colonies which eventually impacted their corruption level (Acemoglu, Johnson and Robinson, 2001); this idea is suggested by Mauro (1995, p. 694). The results are shown in Appendix Figure 9 and Table 9 and are largely consistent with the results reported here, with four differences. First, these alternative instruments are weaker than our original choices. Second, the marginal effect of corruption on women's representation is statistically insignificant ($\alpha = 0.05$, two-tailed) after 2006, though its magnitude is similar to the statistically significant estimates in and before 2005. Third, and perhaps relatedly, corruption is a negative but statistically insignificant ($\alpha = 0.05$, two-tailed) predictor of women's representation in a model that includes fixed effects for region and time. Fourth, the magnitudes of estimated causal effects are smaller, though still substantively important.

Table 1: IV/GMM2S Estimate, Effect of Women's Representation in Government on ICRG Corruption Score

	(1)	(2)	(3)	(4)
% women in lower house	-0.118***	-0.0911***	-0.0368	-0.00657***
	(-7.71)	(-4.63)	(-0.69)	(-4.91)
lag ICRG				1.052***
				(33.76)
lag (2) ICRG				-0.139***
				(-4.79)
Observations	1177	1177	1177	1341
Countries	74	74	74	76
Years	21	21	21	19
Region FE	No	Yes	No	No
Country FE	No	No	Yes	No
Time FE	No	Yes	Yes	Yes
Hansen's J	2.013	0.133	0.299	0.245
Hansen's J p-value	0.156	0.715	0.585	0.621
1st stage F-stat (Cragg-Donald)	219.3	125.0	29.55	6038.4
1st stage F-stat (Kleibergen-Paap)	18.75	11.02	2.618	6352.7
endog. test	16.79	8.152	0.761	0.293
endog. p-value	0.0000417	0.00430	0.383	0.589

t statistics in parentheses

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

Instrumental variables for models (1), (2), and (3): gross enrollment ratio of females in secondary school and proportion females in the labor force. Instrumental variables for model (4): lag and second lag of % women in the lower house. Estimates and standard errors are clustered on country. Constants and fixed effects omitted from table. First stage results shown in Appendix Table 10.

with 95% confidence intervals; panel 2b shows the results of the Sargan test of instrument validity and first-stage F-test for significance of the instruments.

Our essential finding from these cross-sectional models is that corruption exerts a statistically significant and substantively meaningful negative effect on women's representation in the lower house of parliament. The magnitude of the estimated effect varies between -5.16 and -6.76, with a mean effect of -5.93. This means that a one unit increase in the ICRG corruption score causes nearly a 6 percentage point drop in women's representation in government. This is roughly the difference in women's representation between Canada (with an average of 20.25% women's representation in the House of Commons in our data set) and the United States (with an average of 14.64% women in the House of Representatives in our data set).[17] Sargan tests accept the validity of the instruments for every cross-sectional model, and Cragg-Donald F-statistics are well above the threshold of 10 suggested by Staiger and Stock (1997). Results for similar models using the Transparency International CPI score (reported in Appendix Figure 8) yield similar inferences.

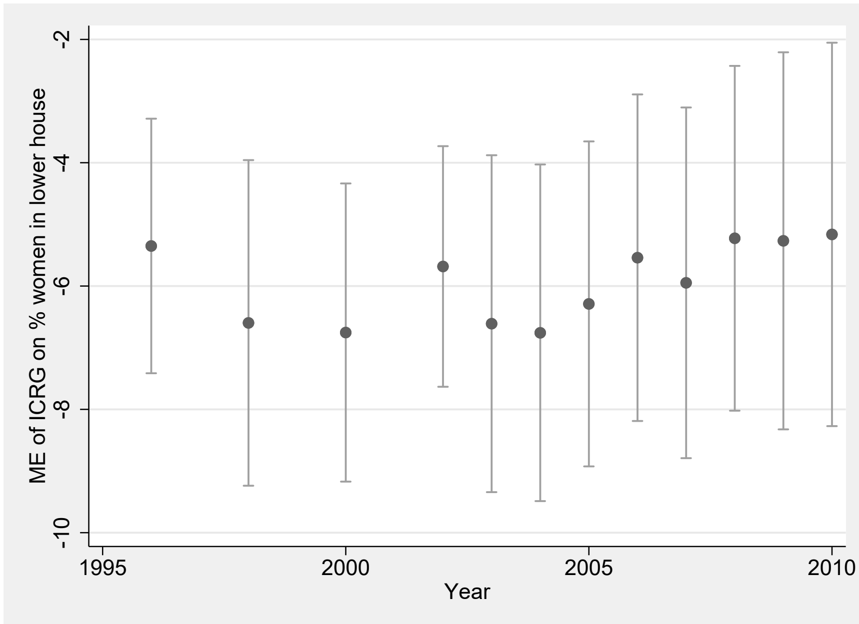
Table 2 shows our panel model estimates of the LATE of increased ICRG corruption score on the percentage of women in the lower house of parliament for both our non-lag-based and lag-based instruments.[18] Although we show results for a model with regional and year fixed effects, we do not present a model with country fixed effects because these effects would be perfectly collinear with our ethnolinguistic fractionalization instrument.

Both sets of instruments show a negative and statistically significant effect of increased corruption on the proportion of women in government. In the model with region and year fixed effects (model 2), a one point increase in ICRG corruption causes a 4.7 percentage point decrease in women's representation in the lower house. This is roughly the same as the difference between women's representation in the U.S. House of Representatives (with average

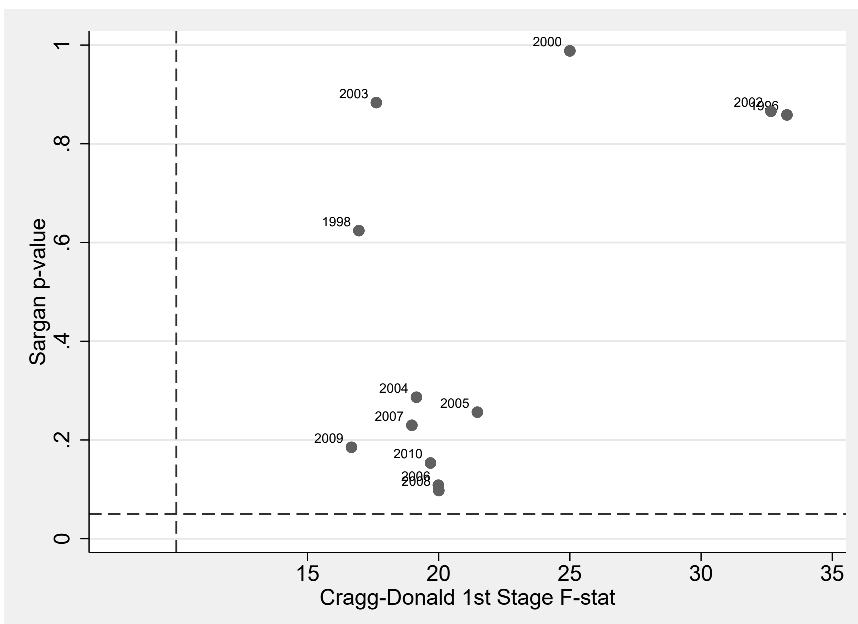
[17]All average figures for women's representation in this section are taken from the sample used to estimate Model 2 in Table 2.

[18]The estimates for the first stage of the model are in Appendix Table 11.

Figure 2: IV/2SLS Estimates of Marginal Effect of ICRG Corruption on Women's Representation, with 95% Confidence Intervals



(a) Marginal Effects



(b) Sargan / F-Statistics

Marginal effect estimates in panel 2a and Sargan / F-statistics in panel 2b are from cross-sectional two-stage least squares models predicting % women in the lower house of the legislature using ICRG corruption score using for each year of available data between 1996 and 2010. Instrumental variables are ethnolinguistic fractionalization and political stability.

women's representation of 14.64%) and that of the National Assembly of Mali (average women's representation = 10.04%). For the lagged instrument model, a one point increase in ICRG corruption score is associated with an instantaneous 0.23 percentage point decrease in the representation of women in government and with a long-run decrease of 6.64 percentage points ($p < 0.001$). This is slightly larger than the difference between Canada and the United States in our data set, as previously noted.

When using the TI corruption perception index as the dependent variable, we get similar though weaker results. The results are shown in Appendix Table 8. For a model including regional and time fixed effects and using ethnolinguistic fractionalization and political stability instruments, a one point increase in the TI CPI causes a 2.1 percent decline in the proportion of women in the lower house of the legislature; this is a substantively small difference, roughly equivalent to the difference between the United States (average women's representation = 14.64%) and Zambia (average women's representation = 12.24%). However, when using the first and second lag of the TI CPI as instruments, we find no appreciable causal relationship from corruption to women's representation.

Conclusion

Does greater representation of women in the lower house of parliament cause decreased corruption, or does greater corruption in government causes lower representation of women in government? In this study, our overall impression is that the evidence supports *both* propositions. The exact magnitude and statistical certainty of the relationship that we find depends on our particular choice of instruments and model specification. However, the majority of our models show a substantively and statistically significant causal relationship in both directions.

The major substantive upshot of our finding is that we should not regard theoretical ar-

Table 2: IV/GMM2S Estimate, Effect of ICRG Corruption on Representation of Women in Government

	(1)	(2)	(3)
ICRG Corruption Score	-5.871***	-4.718**	-0.232**
	(-6.08)	(-2.73)	(-3.25)
lag % women in lower house			0.895***
			(23.20)
lag (2) % women in lower house			0.0703
			(1.72)
Observations	853	853	1341
Countries	73	73	76
Years	12	12	19
Region FE	No	Yes	No
Country FE	No	No	No
Time FE	No	Yes	Yes
Hansen's J	1.082	0.0391	1.835
Hansen's J p-value	0.298	0.843	0.176
1st stage F-stat (Cragg-Donald)	237.8	129.8	4565.7
1st stage F-stat (Kleibergen-Paap)	22.46	10.91	3475.1
endog. test	3.903	0.488	0.0378
endog. p-value	0.0482	0.485	0.846

t statistics in parentheses

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

Instrumental variables for models (1) and (2): ethnolinguistic fractionalization and political stability. Instrumental variables for model (3): lag and second lag of the ICRG score. Estimates and standard errors are clustered on country. Constants and fixed effects omitted from table. First stage results shown in Appendix Table 11.

guments in the literature in favor of corruption decreasing women's representation as being in conflict with theoretical arguments in favor of women's representation decreasing corruption. These two streams of argument are not in mutually exclusive competition with one another: in our study, there is evidence for both.

We think that future research should concentrate on examination of the specific causal mechanisms and implications of these theories—for example, trying to determine *why*, *when* and *how much* women's representation changes corruption instead of *whether* a relationship exists. Research should further explore whether the link between women's representation and corruption results from women themselves being less corrupt than men (as a result of differential risk aversion or differential treatment by voters (Esarey and Schwindt-Bayer, 2018)), hiring staff who are less corrupt, or by breaking down gendered cultures and "old boys' networks" in political networks. Research could also explore ways in which corrupt networks operate to keep women out of politics: do these networks not select women as candidates, or do women choose not to enter races for offices where corrupt networks dominate the process?

Finally, we believe that future studies must explicitly account for the possibility of reverse causality as a part of their modeling strategy. Although no single piece of evidence can be considered conclusive, we think the evidence of simultaneous causality we found in this study is sufficiently compelling to consider this the default expectation in future work.

References

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *The American Economic Review* 91(5):1369–1401.

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.

Alhassan-Alolo, N. 2007. "Gender and corruption: Testing the new consensus." *Public Administration and Development* 237:227–237.

Angrist, Joshua D. and Guido W. Imbens. 1995. "Two-stage least squares estimation of average causal effects in models with variable treatment intensity." *Journal of the American statistical Association* 90(430):431–442.

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–455.

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

Arriola, Leonardo R. and Martha C. Johnson. 2014. "Ethnic Politics and Women's Empowerment in Africa: Ministerial Appointments to Executive Cabinets." *American Journal of Political Science* 58(2):495–510.

Barnes, Tiffany D. 2016. *Gendering Legislative Behavior: Institutional Constraints and Collaboration*. New York: Cambridge University Press.

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

Barnes, Tiffany D. and Emily Beaulieu. 2018. "Women Politicians, Institutions, and Perceptions of Corruption." *Comparative Political Studies* (OnlineFirst):1–34.

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.

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

Benstead, Lindsay J., Amaney A. Jamal and Ellen Lust. 2015. "Is It Gender, Religiosity or Both? A Role Congruity Theory of Candidate Electability in Transitional Tunisia." *Perspectives on Politics* 13(1):74–94.

Benstead, Lindsay J. and Ellen Lust. 2018. Why Do Some Voters Prefer Female Candidates? The Role of Perceived Incorruptibility in Arab Elections. In *Gender and Corruption*, ed. Helena Stensöta and Lena Wängnerud. Political Corruption and Governance Cham, Switzerland: Palgrave Macmillan chapter 5, pp. 83–104.

Bernasek, Alexandra and Stephanie Shwiff. 2001. "Gender, Risk, and Retirement." *Journal of Economic Issues* 35(2):345–356.
URL: *http://www.jstor.org/stable/4227666*

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

Branisa, Boris and Maria Ziegler. 2011. Reexamining the link between gender and corruption: The role of social institutions. In *Proceedings of the German Development Economics Conference, Berlin 2011*. Verein für Socialpolitik, Research Committee Development Economics. URL: `http://ideas.repec.org/p/zbw/gdec11/15.html`.

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.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90(3):414–427.

Chiebub, Jose Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and dictatorship revisited." *Public Choice* 143:67–101.

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`.

Cragg, John G. and Stephen G. Donald. 1993. "Testing identifiability and specification in instrumental variable models." *Econometric Theory* 9(02):222–240.

Croson, Rachel and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47(2):448–474.
URL: *http://www.jstor.org/stable/27739928*

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. 2017. *Analytics, Policy, and Governance*. Yale University Press chapter 2: Causal Inference with Observational Data, pp. 40–66.

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`.

Fallon, Kathleen M., Liam Swiss and Jocelyn Viterna. 2012. "Resolving the democracy paradox: Democratization and women's legislative representation in developing nations, 1975 to 2009." *American Sociological Review* 77(3):380–408.

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

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`.

Hadenius, Axel and Jan Teorell. 2007. "Pathways from authoritarianism." *Journal of democracy* 18(1):143–157.

Heath, Roseanna M., Leslie A. Schwindt-Bayer and Michelle M. Taylor-Robinson. 2005. "Women on the Sidelines: Women's Representation on Committees in Latin American Leglislatures." *American Journal of Political Science* 49(2):420–436.

Hughes, Melanie M. 2009. "Armed conflict, international linkages, and women's parliamentary representation in developing nations." *Social Problems* 56(1):174–204.

Hughes, Melanie M. and Aili Mari Tripp. 2015. "Civil war and trajectories of change in women's political representation in Africa, 1985–2010." *Social forces* 93(4):1513–1540.

Inter-Parliamentary Union. 2012. *Women in Parliaments: World and Regional Averages.* URL: `http://www.ipu.org/wmn-e/world.htm`.

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.". **URL:** *http://www.npr.org/2013/09/28/226903227/mexican-state-s-anti-corruption-plan-hire-women-traffic-cops*

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

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2010. *The Worldwide Governance Indicators: Methodology and Analytical Issues.* SSRN. URL: http://ssrn.com/paper=1682130.

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.

Kenworthy, Lane and Melissa Malami. 1999. "Gender inequality in political representation: A worldwide comparative analysis." *Social Forces* 78(1):235–268.

Kleibergen, Frank and Richard Paap. 2006. "Generalized reduced rank tests using the singular value decomposition." *Journal of Econometrics* 133(1):97–126.

Lindberg, Staffan I. 2004. "Democratization and Women's Empowerment: The Effects of Electoral Systems, Participation and Repetition in Africa." *Studies in Comparative International Development* 39(1):28–53.

Marshall, Monty, Ted Robert Gurr and Keith Jaggers. 2014. "Polity IV Project Political Regime Characteristics and Transitions, 1800-2013: Dataset Users' Manual.".

Mauro, Paolo. 1995. "Corruption and growth." *The Quarterly Journal of Economics* 110(3):681–712.

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.
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.

Norris, Pippa. 1985. "Women's legislative participation in Western Europe." *West European Politics* 8(4):90–101.

Oakes, Ann and Elizabeth Almquist. 1993. "Women in national legislatures." *Population Research and Policy Review* 12(1):71–81.

Paxton, Pamela and Melanie M. Hughes. 2007. *Women, politics, and power: A global perspective*. Pine Forge Press.

Paxton, Pamela and Sheri Kunovich. 2003. "Women's political representation: The importance of ideology." *Social Forces* 82(1):87–113.

Quinones, Sam. 1999. "Stop!" *Ms.* (December):24.

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

Roeder, Philip G. 2014. "Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985." Online. URL: `http://pages.ucsd.edu/~proeder/elf.htm` accessed by the QoG Data Set creators on 3-6-2014.

Rule, Wilma. 1981. "Why women don't run: The critical contextual factors in women's legislative recruitment." *Political Research Quarterly* 34(1):60–77.

Schwindt-Bayer, Leslie A. 2005. "The incumbency disadvantage and women's election to legislative office." *Electoral Studies* 24(2):227–244.

Schwindt-Bayer, Leslie A. 2010. *Political Power and Women's Representation in Latin America*. New York: Oxford University Press.

Schwindt-Bayer, Leslie A., ed. 2018. *Gender and Representation in Latin America*. New York: Oxford University Press.

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. Political Corruption and Governance 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.

Shleifer, Andrei and Robert W. Vishny. 1993. "Corruption." *The Quarterly Journal of Economics* 108(3):599–617.

Staiger, Douglas and James H. Stock. 1997. "Instrumental variables regression with weak instruments." *Econometrica* 65(3):557.

Stock, James H. and Motohiro Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. Donald W.K. Andrews and James H Stock. New York: Cambridge University Press chapter 5, pp. 80–108.

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. **URL:** *http://www.jstor.org/stable/116920*

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.

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

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

Teorell, Jan, Nicholas Charron, Marcus Samanni, Sören Holmberg and Bo Rothstein. 2015. "The Quality of Government Dataset, version Jan15." 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.

Tripp, Aili. 2001. "Women's Movements and Challenges to Neopatrimonial Rule: Preliminary Observations from Africa." *Development and Change* 32(1):33–54.

UNESCO Institute for Statistics. 2014. "UIS Data Center." Online. URL: `http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=136&IF_Language=eng&BR_Topic=0` accessed by the QoG Data Set creators on 3-12-2014.

Uslaner, Eric M. and Bo Rothstein. 2016. "The historical roots of corruption: State building, economic inequality, and mass education." *Comparative Politics* 48(2):227–248.

Wängnerud, Lena. 2012. Why women are less corrupt than men. In *Good Government: The Relevance of Political Science*, ed. Sören Holmberg and Bo Rothstein. Northhampton, MA: Edward Elgar chapter 12, p. 368.

Watson, David and Amy Moreland. 2014. "Perceptions of Corruption and the Dynamics of Women's Representation." *Politics & Gender* 10(03):392–412.

Watson, John and Mark McNaughton. 2007. "Gender differences in risk aversion and expected retirement benefits." *Financial Analysts Journal* 63(4):52–62.
URL: *http://www.jstor.org/stable/10.2307/4480862*

World Bank. 2014. *World Development Indicators*. The World Bank. URL: `http://data.worldbank.org/data-catalog/world-development-indicators` accessed by the QoG Data Set creators on 5-19-2014.

World Bank Group. 1997. "Corruption and Economic Development." Online. URL: `http://www1.worldbank.org/publicsector/anticorrupt/corruptn/cor02.htm` accessed 9/16/2018.

Yoon, Mi Yung. 2004. "Explaining Women's Legislative Representation in Sub-Saharan Africa." *Legislative Studies Quarterly* 29(3):447–468.

Żemojtel-Piotrowska, Magdalaena A., Alison Marganski, Tomasz Baran and Jarosław Piotrowski. 2016. "Corruption and Sexual Scandal: The Importance of Politician Gender." *Anales de Psicologì/Annals of Psychology* 33:133–141.

Appendix A: LATE of women's representation on corruption, by democratic consolidation

Esarey and Chirillo (2013) and Esarey and Schwindt-Bayer (2018) argue that electoral accountability to voters is a key reason why female representatives are disproportionately deterred from engaging in corruption. If this is true, we may expect to find that increased representation of women in government only lowers corruption in the most consolidated democracies, where electoral accountability is strongest.

We create a variable for the long-term consolidation of democratic institutions, which $= 1$ if a state has been rated as a democracy by Chiebub, Gandhi and Vreeland (2010) every year between 1960 and 1989 and $= 0$ otherwise.[19] Chiebub, Gandhi and Vreeland (2010, quoting p. 69) code a country as a democracy in a particular year if:

1. the chief executive is chosen by popular election or by a body that was itself popularly elected;
2. the legislature is popularly elected;
3. there is more than one party competing in the elections; and
4. an alternation in power under electoral rules identical to the ones that brought the incumbent to office has taken place.

[19] Countries in the sample for Model 1 of Table 3 and classified as consolidated democracies are: Australia, Austria, Belgium, Canada, Colombia, Costa Rica, Denmark, Finland, France, Germany, Iceland, India, Ireland, Israel, Italy, Jamaica, Japan, the Netherlands, New Zealand, Norway, Papua New Guinea, Sweden, Switzerland, Trinidad and Tobago, the United Kingdom, the United States, and Venezuela. Jamaica, New Guinea, and Trinidad and Tobago are included despite achieving independence after 1960 because they were administered by consolidated democracies (Australia and the United Kingdom) until independence. Countries in the sample and coded as non-consolidated democracies are: Albania, Argentina, Bangladesh, Bolivia, Botswana, Bulgaria, Chile, Croatia, Cyprus, Czech Republic, Dominican Republic, Ecuador, El Salvador, Estonia, Ghana, Greece, Guatemala, Guyana, Honduras, Hungary, Latvia, Lithuania, Malawi, Malaysia, Mali, Mexico, Mongolia, Mozambique, Namibia, Nicaragua, Panama, Paraguay, Peru, the Philippines, Poland, Portugal, Romania, Slovakia, Slovenia, South Africa, South Korea, Spain, Sri Lanka, Thailand, Turkey, the Ukraine, and Uruguay. Model 3 of Table 3 adds Brazil and Zambia to the sample, both classified as non-consolidated.

We get the original Chiebub, Gandhi and Vreeland measure from the QoG data set. We measure consolidated democracy in this way to minimize endogeneity between corruption in government and democracy level. Most importantly, corruption is not directly implicated in any aspect of the measure. Moreover, the long time span of the measure (unlike the Polity score, which varies from year to year) makes it less likely that fluctuations in the corruption level between 1990 and 2010 are responsible for classification as a consolidated democracy.

Table 3 shows the results of our models of the ICRG corruption score using an interaction of the consolidated democracy dummy with women's representation in government for non-lag instruments (models 1 and 2) and lagged instruments (model 3); note that we do not estimate a country fixed effect model in this case because of perfect collinearity with the consolidated democracy variable.[20]

While it may not be immediately apparent from the coefficients, Models 1 and 3 find a substantial and negative relationship between the proportion of women in government and corruption in consolidated democracies, but no statistically significant relationship in non-consolidated democracies. Model 2, however, finds a statistically significant and negative relationship in both contexts. To clarify our finding, Figure 3 displays the marginal effect of % women in parliament on the ICRG score at varying values of the Polity score for all three models with 95% confidence intervals. As the Figure shows for Model 1, in consolidated democracies a 10 percentage point increase in women's representation in government causes a 1.2 point decline in the ICRG corruption score, representing 17.1% of the total change possible in the corruption scale. In non-consolidated democracies, this effect is much smaller and statistically insignificant. Qualitatively comparable results are found in Model 3. However, in Model 2 (including region and year fixed effects), the causal relationship between women's

[20]Note that Model 1 in Table 3 excludes the instrument of interaction between female secondary school enrollment and consolidated democracy. We omit this instrument because the `ranktest` package detects excessive collinearity between the instruments when calculating the underidentification test statistic for this model. The estimates corresponding to the first stage of the model are in Appendix Tables 12 and 13.

Table 3: IV/GMM2S Estimate, Effect of Women's Representation in Government on ICRG Corruption by Dem. Consolidation

	(1)	(2)	(3)
% women in lower house	-0.0362	-0.0803*	-0.000378
	(-0.95)	(-2.55)	(-0.05)
consolidated democracy	0.526	-0.0606	0.122
	(0.72)	(-0.14)	(0.64)
% women * democracy	-0.0841	-0.00387	-0.0133
	(-1.78)	(-0.12)	(-1.08)
lag ICRG			1.030***
			(28.82)
lag (2) ICRG			-0.155***
			(-5.39)
Observations	1177	1177	1341
Countries	74	74	76
Years	21	21	19
Region FE	No	Yes	No
Country FE	No	No	No
Time FE	No	Yes	Yes
Hansen's J	1.872	1.376	6.390
Hansen's J p-value	0.171	0.503	0.0410
1st stage F-stat (Cragg-Donald)	36.60	43.85	25.25
1st stage F-stat (Kleibergen-Paap)	4.410	4.746	3.331
endog. test	14.19	9.453	0.262
endog. p-value	0.000828	0.00886	0.609

t statistics in parentheses

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

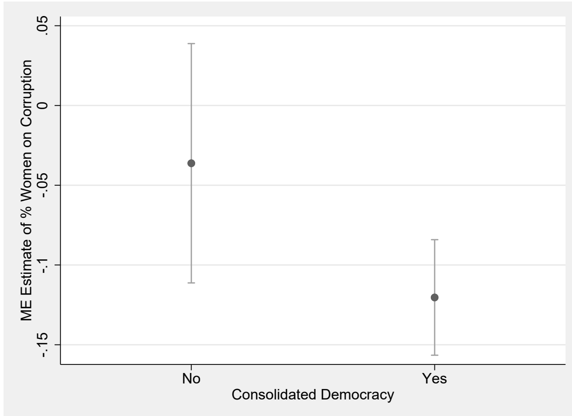
Instrumental variables for models (1) and (2): gross enrollment ratio of females in secondary school, proportion females in the labor force, and interaction of labor force participation with consolidated democracy. Model (2) adds interaction of female secondary school enrollment and consolidated democracy as an instrument. Instrumental variables for model (3): lag and second lag of % women in the lower house, and interactions with consolidated democracy. Estimates and standard errors are clustered on country. Constants and fixed effects omitted from table. First stage results in Appendix Tables 12 and 13.

representation in government and corruption score is estimated to be similar in magnitude in both consolidated and non-consolidated democracies, and is statistically significant in both contexts.

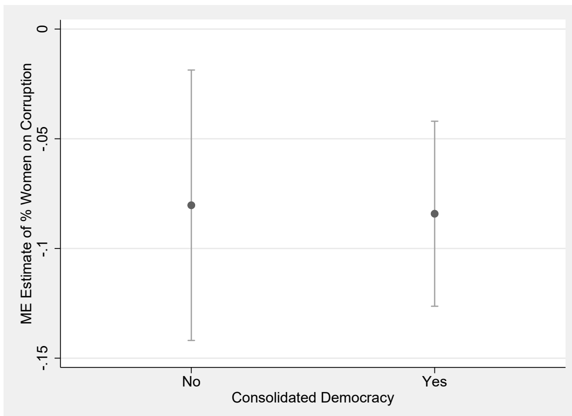
Results from models using the Transparency International CPI dependent variable are reported in Appendix Table 7 and Appendix Figure 7. The substantive conclusions of this analysis are similar to those we derived from the ICRG analysis of Table 3. In this case, the estimated causal effect of women's representation on TI CPI score is statistically significant for consolidated democracies and insignificant for non-consolidated democracies in two models, using $\alpha = 0.05$, two-tailed (though very close to statistical significance in both cases). But these effects are statistically insignificant in both contexts for the model using lagged instruments.

Our overall conclusion is that our evidence neither cleanly supports nor refutes the hypothesis that increased women's representation in the legislature lowers corruption in consolidated democracies but not in non-consolidated democracies. Although the total evidence leans somewhat toward supporting the hypothesis, our findings depend on modeling choices where the best choice is not evident. Consequently, we rate our findings as inconclusive on this question.

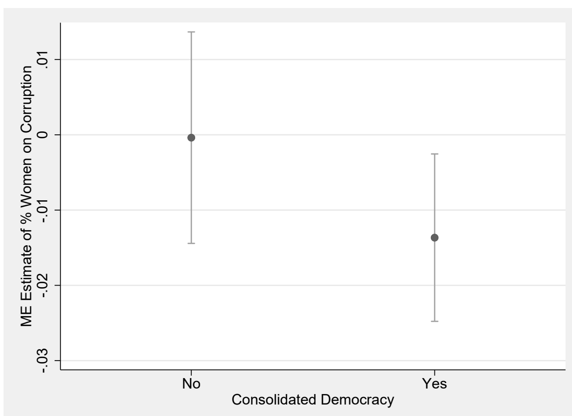
Figure 3: Marginal Effect of Women's Representation, by Democracy with 95% Confidence Intervals



(a) Model 1 in Table 3



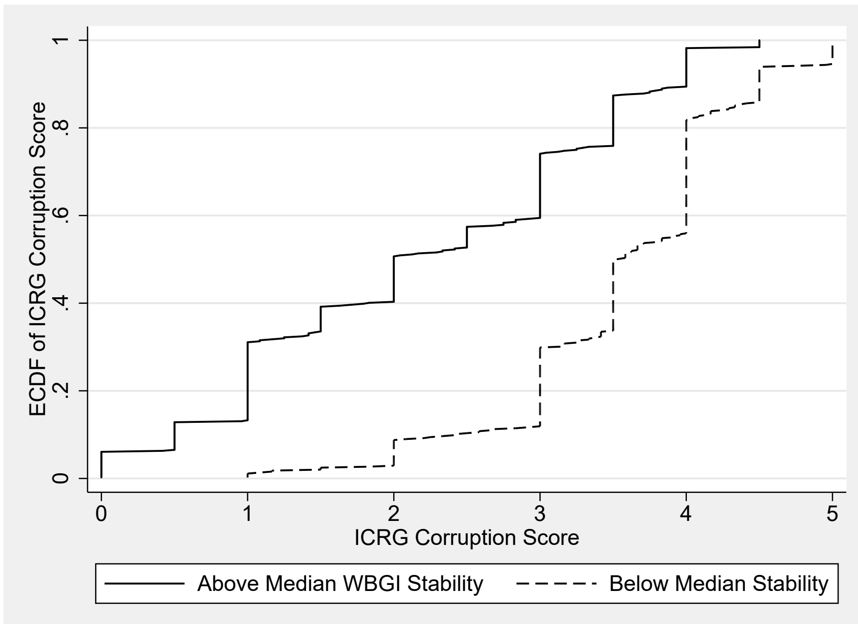
(b) Model 2 in Table 3



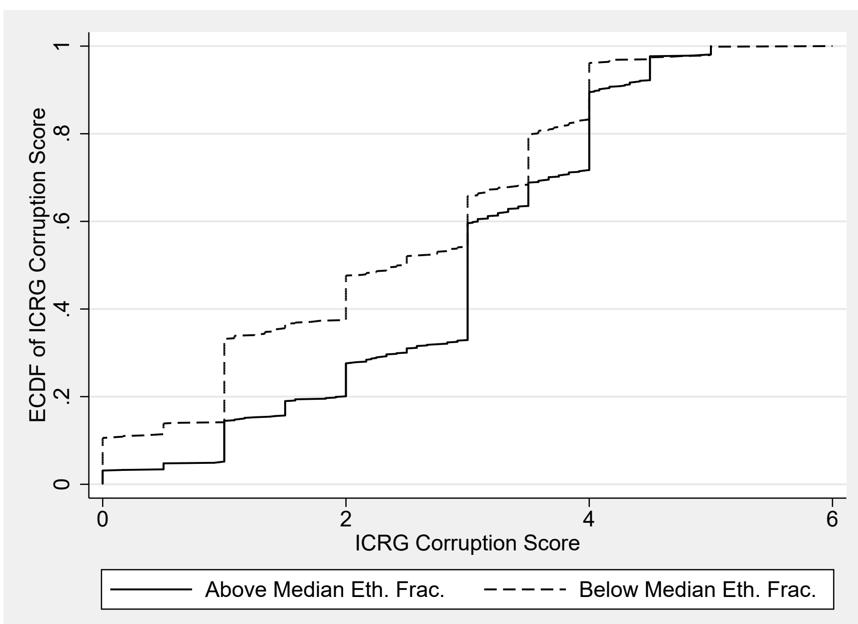
(c) Model 3 in Table 3

Appendix B: Additional Models and Checks

Figure 4: Instrument Monotonicity Check, ICRG Corruption Score

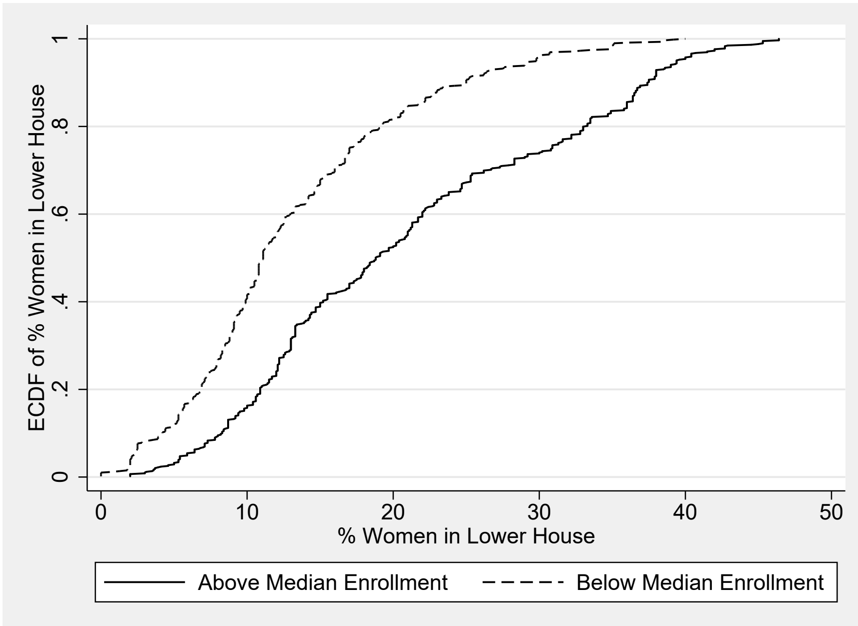


(a) Political Stability

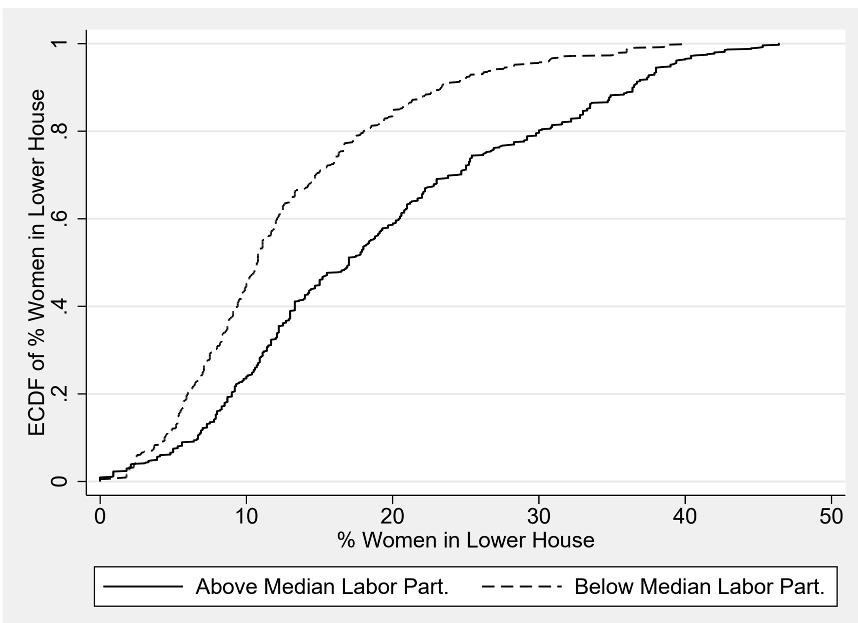


(b) Ethnolinguistic Fractionalization

Figure 5: Instrument Monotonicity Check, % Women in the Lower House



(a) Gross Enrollment Ratio of Females in Secondary Schools



(b) Proportion of Women in the Labor Force

Table 4: FIML/SEM Estimate, Relationship Between Women's Representation in Government and ICRG Corruption Score, w/o Political Stability Instrument

	(1)	(2)	(3)
DV: ICRG Corruption Score			
% women in lower house	-0.110***	-0.0964***	-0.00661***
	(-5.99)	(-4.71)	(-4.97)
Ethnolinguistic Fractionalization	0.931*	-0.0425	
	(2.06)	(-0.10)	
lag ICRG			1.050***
			(33.92)
lag (2) ICRG			-0.137***
			(-4.71)
DV: % women in lower house			
ICRG Corruption Score	13.84	6.520*	-0.202**
	(0.69)	(2.44)	(-2.72)
Female School Ratio	0.480	0.244***	
	(1.07)	(4.11)	
Female Labor Force Ratio	0.877	0.750*	
	(1.09)	(2.23)	
lag % women in lower house			0.888***
			(22.89)
lag (2) % women in lower house			0.0805
			(1.95)
Observations	1116	1116	1341
Countries	71	71	76
Years	21	21	19
Region FE	No	Yes	No
Country FE	No	No	No
Time FE	No	Yes	Yes

t statistics in parentheses

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

Models estimated with `sem` command in Stata. Model (2) specifies "diagonal" correlation between error terms; Models (1) and (3) specify "unstructured." Model (2) fails to estimate with "unstructured" covariance. Constants and fixed effects omitted from table.

Table 5: FIML/SEM Estimate, Relationship Between Women's Representation in Government and ICRG Corruption Score, with Political Stability Instrument

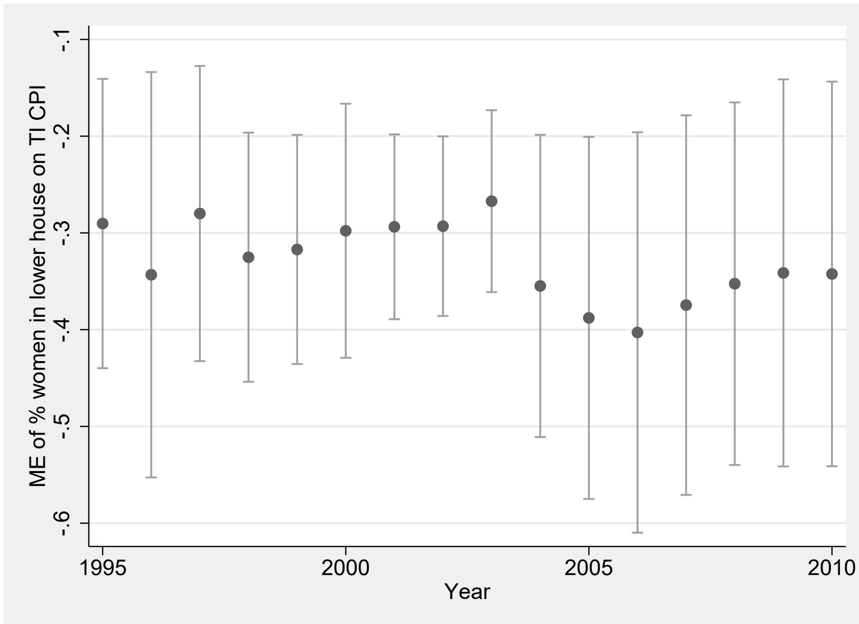
	(1)	(2)
DV: ICRG Corruption Score		
% women in lower house	-0.0368	-0.0699**
	(-1.17)	(-2.83)
WBGI Political Stability	-0.779**	-0.511*
	(-3.23)	(-2.56)
Ethnolinguistic Fractionalization	0.219	-0.413
	(0.52)	(-1.54)
DV: % women in lower house		
ICRG Corruption Score	-0.203	-3.576
	(-0.08)	(-1.44)
Female School Ratio	0.147	0.0976
	(1.48)	(1.84)
Female Labor Force Ratio	0.430	0.301*
	(1.93)	(2.05)
Observations	707	707
Countries	71	71
Years	12	12
Region FE	No	Yes
Country FE	No	No
Time FE	No	Yes

t statistics in parentheses

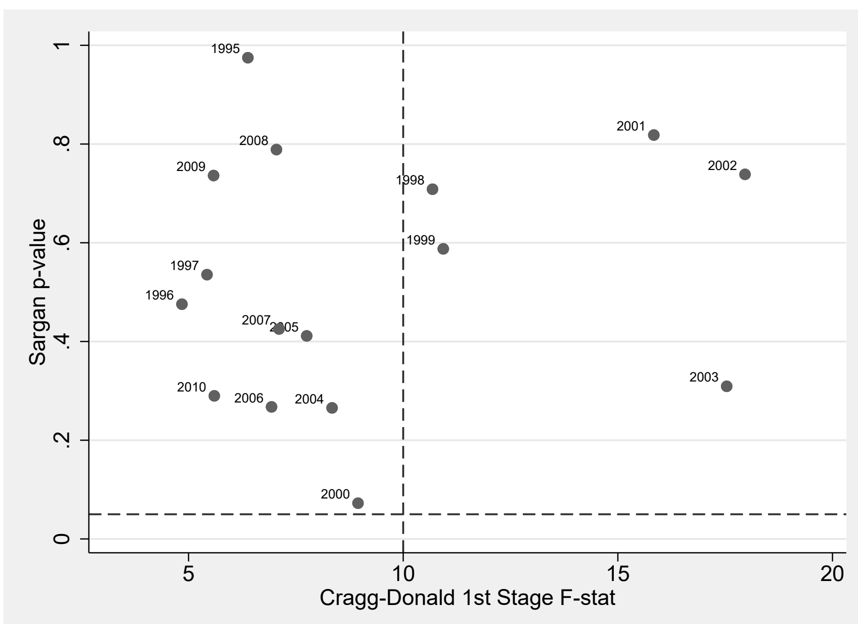
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Model estimated with `sem` command in Stata. Model (1) specifies the "diagonal" correlation option between the error terms; Model (2) specifies "unstructured." Model (1) fails to estimate with "unstructured" covariance. Constants and fixed effects omitted from table.

Figure 6: IV/2SLS Estimates of Marginal Effect of Women's Representation on TI CPI, with 95% Confidence Intervals



(a) Marginal Effects



(b) Sargan / F-Statistics

Marginal effect estimates in panel 6a and Sargan / F-statistics in panel 6b are from cross-sectional two-stage least squares models predicting TI CPI score using % women in the lower house of the legislature for each year of data between 1995 and 2010. Instrumental variables are gross enrollment ratio of females in secondary school and proportion females in the labor force.

Table 6: IV/GMM2S Estimate, Effect of Women's Representation in Government on TI Corruption Score

	(1)	(2)	(3)	(4)
% women in lower house	-0.316***	-0.211***	0.0717	-0.00154
	(-6.11)	(-3.57)	(1.20)	(-1.52)
lag TI CPI				1.030***
				(26.76)
lag (2) TI CPI				-0.0449
				(-1.19)
Observations	877	877	875	858
Countries	73	73	73	76
Years	16	16	16	14
Region FE	No	Yes	No	No
Country FE	No	No	Yes	No
Time FE	No	Yes	Yes	Yes
Hansen's J	0.788	0.0728	0.0673	0.0593
Hansen's J p-value	0.375	0.787	0.795	0.808
1st stage F-stat (Cragg-Donald)	137.2	71.40	7.594	4904.7
1st stage F-stat (Kleibergen-Paap)	15.75	7.315	1.308	5269.4
endog. test	20.39	9.248	1.154	1.759
endog. p-value	0.00000630	0.00236	0.283	0.185

t statistics in parentheses

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

Instrumental variables for models (1) and (2): gross enrollment ratio of females in secondary school and proportion females in the labor force. Instrumental variables for model (3): lag and second lag of % women in the lower house. Estimates and standard errors are clustered on country. Constants and fixed effects omitted from table.

Table 7: IV/GMM2S Estimate, Effect of Women's Representation in Government on TI Corruption Score by Dem. Consolidation

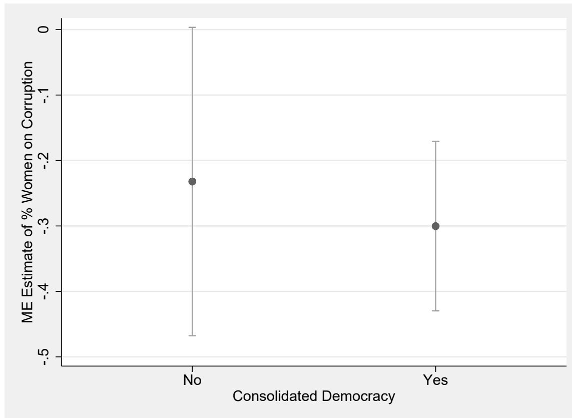
	(1)	(2)	(3)
% women in lower house	-0.232	-0.163	0.000508
	(-1.93)	(-1.78)	(0.17)
consolidated democracy	0.152	-0.631	0.0632
	(0.05)	(-0.50)	(0.63)
% women * democracy	-0.0680	-0.0234	-0.00422
	(-0.44)	(-0.29)	(-0.73)
lag TI CPI			1.044***
			(27.75)
lag (2) TI CPI			-0.0642
			(-1.73)
Observations	877	877	855
Countries	73	73	76
Years	16	16	14
Region FE	No	Yes	No
Country FE	No	No	No
Time FE	No	Yes	Yes
Hansen's J	0.582	1.733	4.024
Hansen's J p-value	0.446	0.420	0.134
1st stage F-stat (Cragg-Donald)	19.35	32.15	45.11
1st stage F-stat (Kleibergen-Paap)	2.375	3.100	3.685
endog. test	16.59	7.694	1.562
endog. p-value	0.000250	0.0213	0.211

t statistics in parentheses

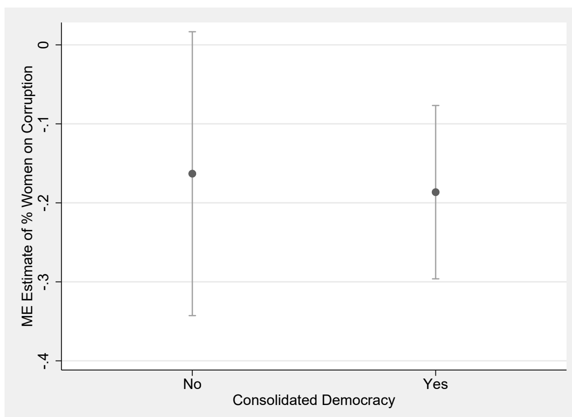
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Instrumental variables for models (1) and (2): gross enrollment ratio of females in secondary school, proportion females in the labor force, and interaction of each of these variables with consolidated democracy. Instrumental variables for model (3): lag and second lag of % women in the lower house, and interactions with consolidated democracy. Estimates and standard errors are clustered on country. Constants and fixed effects omitted from table.

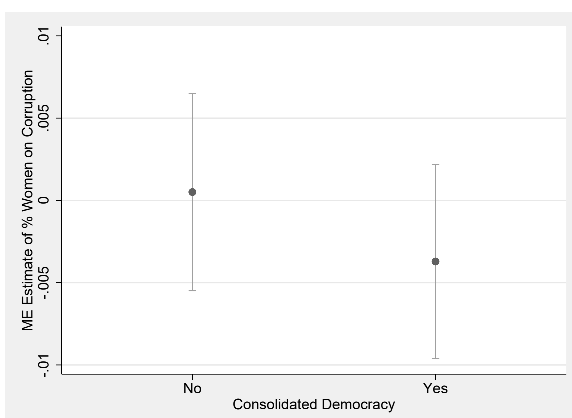
Figure 7: Marginal Effect of Women's Representation, by Democracy
with 95% Confidence Intervals



(a) Model 1 in Table 7

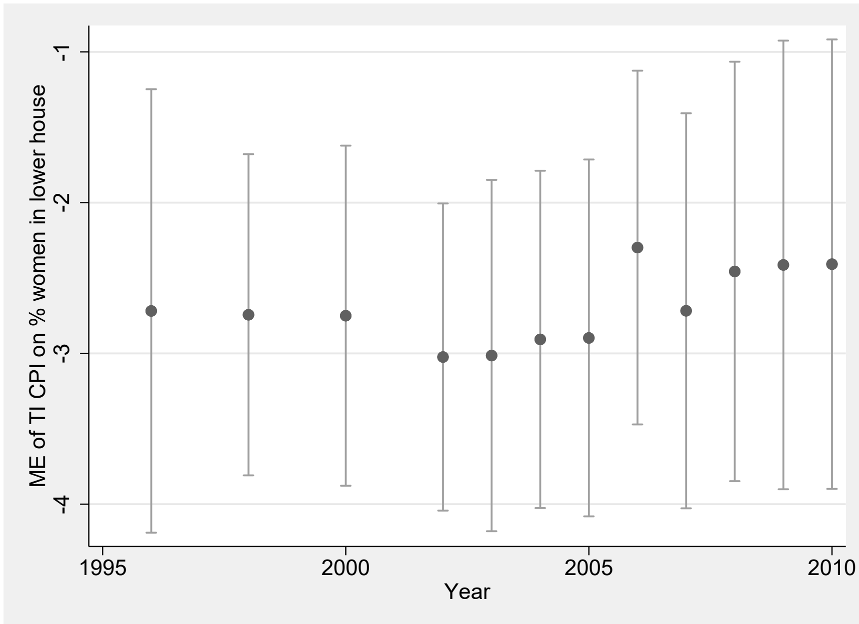


(b) Model 2 in Table 7

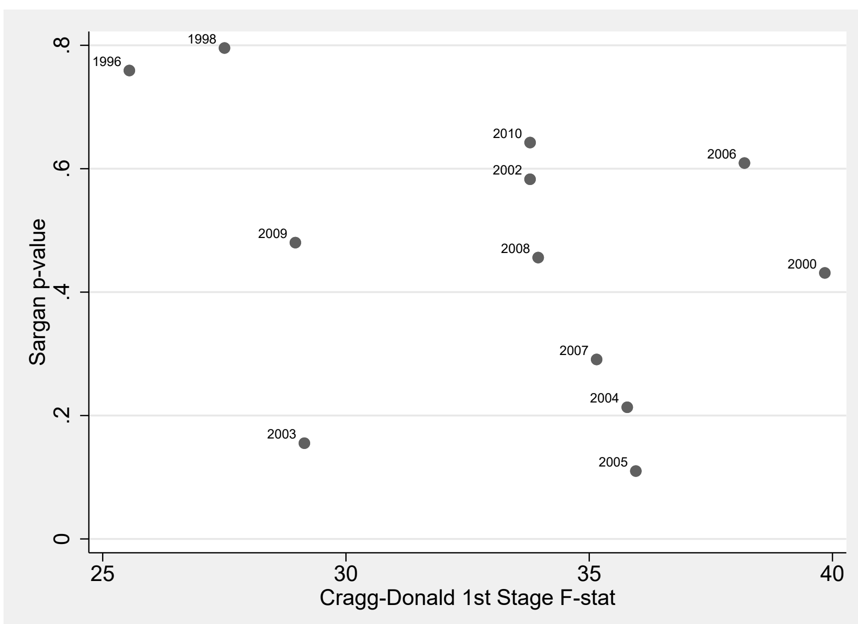


(c) Model 3 in Table 7

Figure 8: IV/2SLS Estimates of Marginal Effect of TI Corruption Perception Index on Women's Representation, with 95% Confidence Intervals



(a) Marginal Effects



(b) Sargan / F-Statistics

Marginal effect estimates in panel 2a and Sargan / F-statistics in panel 2b are from cross-sectional two-stage least squares models predicting % women in the lower house of the legislature using TI CPI score using for each year of available data between 1996 and 2010. Instrumental variables are ethnolinguistic fractionalization and political stability.

Table 8: IV/GMM2S Estimate, Effect of TI Corruption Score on Women's Representation in Government

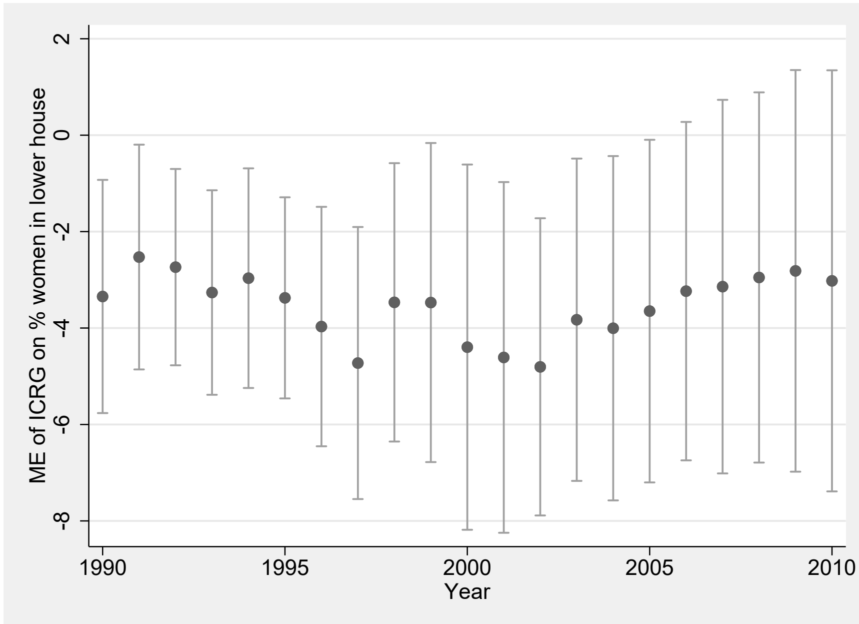
	(1)	(2)	(3)
TI Corruption Perception	-2.671***	-2.082*	-0.0239
	(-5.41)	(-2.44)	(-0.62)
lag % women in lower house			0.894***
			(32.02)
lag (2) % women in lower house			0.0897**
			(3.16)
Observations	787	787	858
Countries	73	73	76
Years	12	12	14
Region FE	No	Yes	No
Country FE	No	No	No
Time FE	No	Yes	Yes
Hansen's J	0.912	0.0564	0.0322
Hansen's J p-value	0.340	0.812	0.858
1st stage F-stat (Cragg-Donald)	407.9	253.6	16691.6
1st stage F-stat (Kleibergen-Paap)	34.43	16.37	27765.1
endog. test	0.386	0.0151	0.574
endog. p-value	0.535	0.902	0.449

t statistics in parentheses

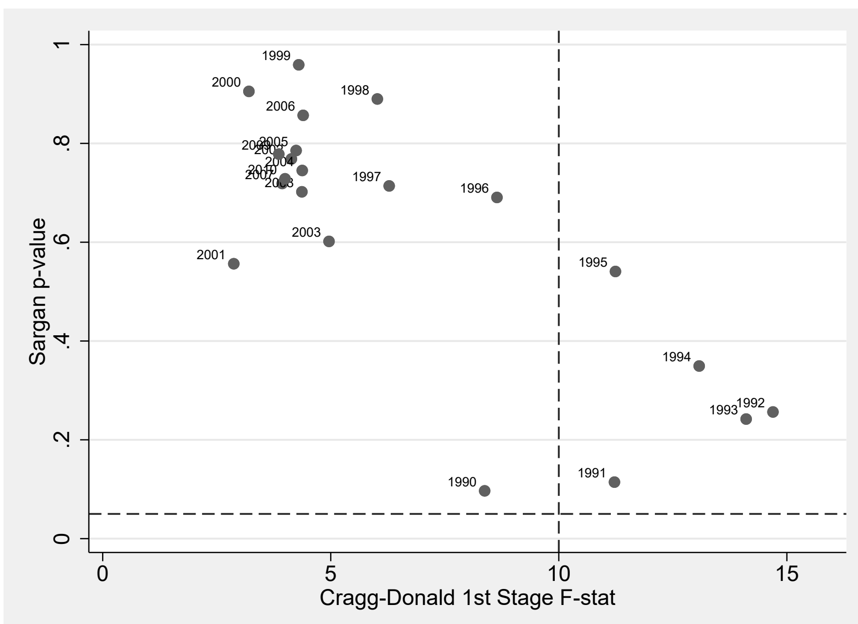
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Instrumental variables for models (1) and (2): ethnolinguistic fractionalization and political stability. Instrumental variables for model (3): lag and second lag of the TI CPI score. Estimates and standard errors are clustered on country. Constants and fixed effects omitted from table.

Figure 9: IV/2SLS Estimates of Marginal Effect of ICRG Corruption Score on Women's Representation using Colonial Dummy Instruments, with 95% Confidence Intervals



(a) Marginal Effects



(b) Sargan / F-Statistics

Marginal effect estimates in panel 9a and Sargan / F-statistics in panel 9b are from cross-sectional two-stage least squares models predicting % women in the lower house of the legislature using ICRG corruption using for each year between 1990 and 2010. Instrumental variables are ethnolinguistic fractionalization and Spanish, French, and British colonial heritage dummy variables.

Table 9: IV/GMM2S Estimate, Effect of ICRG Corruption on Representation of Women in Government using Colonial Heritage Dummy Instruments

	(1)	(2)
ICRG Corruption Score	-4.016***	-4.375
	(-3.46)	(-1.07)
Observations	1431	1431
Countries	73	73
Years	21	21
Region FE	No	Yes
Country FE	No	No
Time FE	No	Yes
Hansen's J	1.031	2.614
Hansen's J p-value	0.794	0.455
1st stage F-stat (Cragg-Donald)	99.53	30.23
1st stage F-stat (Kleibergen-Paap)	26.30	2.459
endog. test	0.943	0.0607
endog. p-value	0.332	0.805

t statistics in parentheses

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

Instrumental variables are ethnolinguistic fractionalization and Spanish, French, and British colonial heritage dummy variables. Constants and fixed effects omitted from table.

Table 10: IV/GMM2S First Stage of Table 1, Effect of Women's Representation in Government on ICRG Score

	(1)	(2)	(3)	(4)
Female School Ratio	0.151***	0.156***	-0.00668	
	(3.74)	(3.84)	(-0.17)	
Female Labor Force Ratio	0.452*	0.472*	0.701*	
	(2.25)	(2.05)	(2.24)	
lag % Women				0.889***
				(22.59)
lag (2) % Women				0.0790
				(1.89)
Observations	1177	1177	1177	1341
Countries	74	74	74	74
Years	21	21	21	21
Region FE	No	Yes	No	No
Country FE	No	No	Yes	No
Time FE	No	Yes	Yes	Yes

t statistics in parentheses

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

Coefficients for Lag ICRG and Lag (2) ICRG included in the model but excluded from this table. Constants and fixed effects omitted from table.

Table 11: IV/GMM2S First Stage of Table 2, Effect of ICRG Score on Women's Representation in Government

	(1)	(2)	(3)
WBGI Stability	-0.877***	-0.812***	
	(-5.75)	(-4.67)	
Ethnolinguistic Fractionalization	0.259	-0.236	
	(0.70)	(-0.60)	
lag ICRG			1.051***
			(33.96)
lag (2) ICRG			-0.137***
			(-4.73)
Observations	853	853	1341
Countries	73	73	73
Years	12	12	12
Region FE	No	Yes	No
Country FE	No	No	No
Time FE	No	Yes	Yes

t statistics in parentheses

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

Coefficients for Lag % Women and Lag (2) % Women included in the model but excluded from this table. Constants and fixed effects omitted from table.

Table 12: IV/GMM2S First Stage of Table 3 (% Women), Effect of Women's Representation in Government on ICRG Score by Consolidation of Democracy

	(1)	(2)	(3)
Consolidated Democracy	-13.86	-14.93	0.386
	(-0.91)	(-0.93)	(0.86)
Female School Ratio	0.105*	0.122	
	(2.47)	(1.90)	
School Ratio X Democracy		0.0461	
		(0.62)	
Female Labor Force Ratio	0.325	0.235	
	(1.56)	(0.91)	
Labor Ratio X Democracy	0.437	0.377	
	(1.20)	(0.81)	
lag % Women			0.888***
			(22.39)
lag % Women X Democracy			-0.0923
			(-0.15)
lag (2) % Women			0.0767
			(1.82)
lag (2) % Women X Democracy			-0.0867
			(-0.15)
Observations	1177	1177	1341
Countries	74	74	74
Years	21	21	21
Region FE	No	Yes	No
Country FE	No	No	No
Time FE	No	Yes	Yes

t statistics in parentheses

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

Coefficients for Lag ICRG and Lag (2) ICRG included in the model but excluded from this table. Constants and fixed effects omitted from table.

Table 13: IV/GMM2S First Stage of Table 3 (% Women X Democracy), Effect of Women's Representation in Government on ICRG Score by Consolidation of Democracy

	(1)	(2)	(3)
Consolidated Democracy	-18.90	-28.41	20.55***
	(-1.36)	(-1.83)	(9.10)
Female School Ratio	0.0643*	-0.00537	
	(2.14)	(-0.32)	
School Ratio X Democracy		0.166***	
		(3.40)	
Female Labor Force Ratio	-0.0334	-0.105	
	(-0.61)	(-1.06)	
Labor Ratio X Democracy	0.910**	0.799*	
	(2.70)	(1.99)	
lag % Women			0.290***
			(3.55)
lag % Women X Democracy			-1.027
			(-0.99)
lag (2) % Women			0.180***
			(3.49)
lag (2) % Women X Democracy			-1.158
			(-1.62)
Observations	1177	1177	1341
Countries	74	74	74
Years	21	21	21
Region FE	No	Yes	No
Country FE	No	No	No
Time FE	No	Yes	Yes

t statistics in parentheses

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

Coefficients for Lag ICRG and Lag (2) ICRG included in the model but excluded from this table. Constants and fixed effects omitted from table.

Table 14: IV/GMM2S Estimate, Effect of Women's Representation in Government on ICRG Corruption Score, with Gender Quota Control

	(1)	(2)	(3)	(4)
% women in lower house	-0.102***	-0.0943***	-0.0382	-0.00689***
	(-6.91)	(-4.99)	(-0.67)	(-4.80)
gender quotas in lower house	1.370***	0.569**	-0.0199	0.0486
	(7.11)	(2.87)	(-0.09)	(1.68)
lag ICRG				1.050***
				(33.69)
lag (2) ICRG				-0.143***
				(-4.89)
Observations	1177	1177	1177	1341
Countries	74	74	74	76
Years	21	21	21	19
Region FE	No	Yes	No	No
Country FE	No	No	Yes	No
Time FE	No	Yes	Yes	Yes
Hansen's J	4.039	0.230	0.304	0.165
Hansen's J p-value	0.0445	0.631	0.582	0.685
1st stage F-stat (Cragg-Donald)	239.1	122.8	25.69	5951.8
1st stage F-stat (Kleibergen-Paap)	19.84	10.65	2.472	5800.1
endog. test	7.429	9.313	0.739	0.184
endog. p-value	0.00642	0.00228	0.390	0.668

t statistics in parentheses

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

Instrumental variables for models (1), (2), and (3): gross enrollment ratio of females in secondary school and proportion females in the labor force. Instrumental variables for model (4): lag and second lag of % women in the lower house. Estimates and standard errors are clustered on country. "Gender quotas in lower house" is a binary variable (1 = yes) indicating whether the country has electoral quotas or reserved seats by gender in that year. Constants and fixed effects omitted from table.