Women’s Representation, Accountability and Corruption in Democracies

JUSTIN ESAREY AND LESLIE A. SCHWINDT-BAYER*

At the turn of the twenty-first century, an important pair of studies established that greater female representation in government is associated with lower levels of perceived corruption in that government. But recent research finds that this relationship is not universal and questions why it exists. This article presents a new theory explaining why women’s representation is only sometimes related to lower corruption levels and provides evidence in support of that theory. The study finds that the women’s representation–corruption link is strongest when the risk of corruption being detected and punished by voters is high – in other words, when officials can be held electorally accountable. Two primary mechanisms underlie this theory: prior evidence shows that (1) women are more risk-averse than men and (2) voters hold women to a higher standard at the polls. This suggests that gender differences in corrupt behavior are proportional to the strength of electoral accountability. Consequently, the hypotheses predict that the empirical relationship between greater women’s representation and lower perceived corruption will be strongest in democracies with high electoral accountability, specifically: (1) where corruption is not the norm, (2) where press freedom is respected, (3) in parliamentary systems and (4) under personalistic electoral rules. The article presents observational evidence that electoral accountability moderates the link between women’s representation and corruption in a time-series, cross-sectional dataset of seventy-six democratic-leaning countries.

Keywords: gender; corruption; democracies; accountability; representation

Fifteen years ago, two important articles by Dollar, Fisman and Gatti and Swamy et al. established a curious observational link: greater representation of women in government is associated with lower levels of perceived corruption1 in that government.2 The impact of these studies was substantial. In academia, the articles are extremely well cited and have inspired a

* Department of Political Science, Rice University (email: justin@justinesarey.com); Department of Political Science, Rice University (email: schwindt@rice.edu). We would like to thank the participants and audiences at the numerous workshops and departmental colloquia where we presented this paper, including ‘Why is Gender Equality Good for Governance?’ at Freie Universitat, Berlin; the University of Tennessee; the University of Maryland; the Center for Women’s Leadership at Portland State University; the Quality of Government Institute at the University of Gothenburg; the European Conference on Politics and Gender in Uppsala, Sweden; and ITAM, Mexico City. We also thank Margit Tavits for sharing the data that she and Leslie Schwindt-Bayer collected. Data replication sets including logs, analysis scripts and data files are available at http://dataverse.harvard.edu/dataverse/BJPolS and online appendices are available at https://doi.org/doi:10.1017/S0007123416000478.

1 By corruption, we mean the appropriation of public authority for personal or private benefit. This definition includes the solicitation of bribes, embezzling public money and other forms of graft. Due to the difficulty of directly observing these usually hidden behaviors, we measure corruption using the perceptions of country experts and business professionals (among others). Our definition and measure is are consistent with those used in most empirical studies of country-level corruption; see the Data and Variables section for a more detailed discussion.

still-growing literature. In the policy world, the findings justified governments enthusiastically bringing women into political offices and bureaucratic positions, such as police forces and the civil service, as an anti-corruption measure.

In the years since the publication of these studies, two important challenges to this finding have emerged. First, research has found that the relationship between women’s representation and corruption is not universal – it holds in some countries but not others. Esarey and Chirillo, for example, find that the relationship is specific to democracies; it does not hold in autocracies. Secondly, scholars have criticized Dollar, Fisman and Gatti’s explanation for the finding – that women are simply more honest and trustworthy and therefore less likely to be corrupt – and offered alternative explanations, such as that women have had less opportunity to engage in corruption because they are often excluded from power and patronage or 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.

These challenges call into question our understanding of the relationship between women’s representation and corruption, as well as the basis for some public policies.

In this article, we ask: Why does the relationship between women’s representation and corruption vary across countries? We argue that women’s representation is linked to corruption through the moderating pathway of electoral accountability, which we define as voters’ ability to identify corrupt officials and punish them at the ballot box. Where electoral accountability is high, corruption is a risky behavior; where electoral accountability is low, corruption is less risky. We expect the strength of the relationship between women’s representation and perceived corruption to be proportional to the risk of being held accountable for corruption, with the strongest relationship in places where the risk of accountability is greatest. We offer two mechanisms to explain why accountability influences the women’s representation–corruption relationship. First, experimental and observational evidence indicates that women tend to be more risk averse than men (on average) when confronting identical situations. If women are more risk averse, they should be less likely to engage in corruption in high-accountability contexts because of the risks involved. Secondly, evidence suggests that voters hold female elected officials to a higher standard than men. If this is true, then the consequences for corruption disproportionately fall on women, which may deter them from participating in corruption when the risk of getting caught and punished is high. At the aggregate level, this translates into a strong relationship between women’s representation and corruption in political systems with high accountability and a weaker relationship in systems with low accountability.

3 Alhassan-Alolo 2007; Barnes and Beaulieu 2014; Esarey and Chirillo 2013; see, for example, Sung 2003; Wangnerud 2012; Watson and Moreland 2014. According to Google Scholar, Dollar, Fisman, and Gatti’s (2001) article has more than 379 citations and Swamy et al. (2001) has over 477 as of 22 September 2014.

4 Kahn 2013; Karim 2011; McDermott 1999; Moore 1999; Quinones 1999. Although Dollar, Fisman, and Gatti (2001) focused only on women’s representation in parliament and corruption, Swamy et al. (2001) studied the effect of women’s parliamentary representation, their presence in senior bureaucratic posts and their labor force participation on corruption. The findings have been used to justify increasing women’s presence in many areas of government, not just parliaments and legislatures.


6 Esarey and Chirillo 2013.


This article studies countries with democratic-leaning institutions, where the concept of electoral accountability for corruption is most relevant. We expect that the observed relationship between women in government and perceived corruption should be strongest in democracies, where institutions allow voters to hold government officials individually accountable for corruption by punishing them at the polls (and weakest where they do not). Specifically, there are four contexts in which we expect greater levels of women’s representation in the legislature to be more strongly associated with lower levels of perceived corruption: (1) where corruption is not an institutional norm, (2) where freedom of the press is respected, (3) in parliamentary rather than presidential systems and (4) under personalistic rather than party-centered electoral rules. As we explain below, each of these settings is associated with high levels of electoral accountability. We test these hypotheses empirically with a time-series, cross-sectional dataset of seventy-six democratic-leaning countries. We present a set of bivariate correlations, multivariate statistical models and substantive marginal effects plots to show that all four hypotheses have strong empirical support, providing compelling new evidence that electoral accountability moderates the relationship between women’s representation and corruption.

The goals of this study are (a) to demonstrate that the empirical link between women’s representation in government and perceived corruption is sensitive to the strength of electoral accountability and (b) to articulate a theory that explains our finding and the pattern of past results. This article is an important contribution because it makes sense of a somewhat confusing pattern of findings and sets a theoretically driven agenda for future research, but it poses at least as many questions as it answers. Future research examining the micro-level mechanisms of differential risk aversion and differential treatment by voters and empirically studying the direction of causality would not be justified if we cannot establish the context sensitivity of the gender–corruption relationship. We return to a more detailed discussion of extensions of the theory and future empirical analyses that we think are suggested by our study in the conclusion.

A THEORY OF GENDER, CORRUPTION AND ACCOUNTABILITY

Why would electoral accountability produce a stronger relationship between women’s representation and reduced corruption? Our theory hinges on gender differences in how elected officials respond to the increased risk of engaging in corruption in governments with strong electoral accountability. The risk of being held accountable for corruption by voters is determined by two factors: the likelihood of corruption being detected and the severity of punishment upon detection. Increases in the probability of detection and/or the severity of...
punishment make the prospect of corruption riskier. It is riskier for both women and men, but we argue that women are disproportionately more discouraged by the higher risk of engaging in corruption in high-accountability systems for two reasons. First, significant research shows that women are more risk averse than men, and if this is the case, then women will react more strongly to the greater risk associated with high-accountability systems. Secondly, research shows that voters perceive of and treat female representatives differently than male representatives, which could lead to women being more likely to be caught and more severely punished by voters than men. This risk increases in systems with higher electoral accountability. For both of these reasons, women should be disproportionately less likely to engage in corruption, and this gender difference should be larger in high-accountability systems than low-accountability systems.

**Mechanism 1: Differential Risk Aversion**

A recent review of the economic literature by Croson and Gneezy presents the following summary of the relationship between gender and risk taking:

The robust finding is that men are more risk prone than women. Previous surveys of economics and psychology report the same conclusions: women are more risk averse than men in the vast majority of environments and tasks.

Much of the evidence of women’s greater risk aversion in economics comes from laboratory experiments. Subjects in these experiments make a series of choices between lotteries offering a different combination of risks and rewards; the lotteries are structured to determine a subject’s risk aversion. The experimental findings are bolstered by observational research on differential risk taking in investment portfolios managed by men and women. In psychology, evidence of gender differences in risk taking comes from a combination of survey experiments with hypothetical choices, self-reported risky behavior from surveys (for example, unsafe sex) and directly observed risky behaviors, such as dangerous traffic maneuvers monitored by researchers.

The explanation for women’s greater risk aversion is unclear. Based on recent evidence indicating that there is no gender difference in risk aversion in traditional societies,
we speculate that it results from the social, cultural and institutional environments in which women are socialized and operate. For the purpose of our research, the reason why women are, on average, more risk averse than men is less important than building from the empirically grounded assumption that they are, on average, and determining how and when that risk aversion translates into different behavior. Experimental research on gender and bribe taking lends insight into this question: it finds that women will only be less likely to take bribes than men when their behavior is being monitored and there is a chance of it being detected – in other words, when that bribe taking (that is, corruption) is risky.25

If women are more averse to the risks presented by corruption than men, then women should be less likely than men to participate in corruption when it is risky. The risks of corruption come from the likelihood of corruption being detected and punished in systems with high levels of electoral accountability. Increases in the probability of detection or the severity of punishment for corruption will more strongly decrease women’s propensity to engage in corruption compared to men. This translates into an empirical expectation: the relationship between women in government and corruption gets stronger as corruption gets riskier.26 This occurs because women respond more strongly than men to an increased possibility of getting caught and punished.

**Mechanism 2: Differential Treatment by Gender**

A second reason why the relationship between women’s representation and corruption may be moderated by the strength of accountability is that the mechanisms of accountability may be biased against women. That is, it is possible that women are proportionally more likely than men to be investigated and caught for engaging in corruption and more likely to be blamed and more harshly punished for corruption. This argument is rooted in recent research findings that women are perceived and treated differently while running for and holding of office.

Research has found that voters evaluate male and female candidates through the filter of gender stereotypes.27 Women have been perceived to be less likely to win elections than men,28 even though research shows that women are as likely to win as men in settings with relatively gender equal cultures;29 surveys suggest that many citizens still think that men make better political leaders than women.30 Evidence of the importance of these stereotypes in evaluations of candidate choice is mixed,31 but stereotypes about differences between male and female political leaders clearly exist. Research has also found that these gendered perceptions of elected officials translate into different behaviors by women in office. One line of research argues that because voters hold female candidates to a higher standard than their male counterparts, women are less likely to run for office.32 Another suggests that women actually perform better than their male counterparts in direct response to gender stereotypes about women in politics.33

---

25 Armantier and Boly 2011; Schulze and Frank 2003.
26 As explained in Footnotes 15 and 16, our theory makes no empirical prediction about the relationship between women in government and corruption when electoral accountability is low or about how men will respond to a greater risk of corruption.
30 Inglehart and Norris 2003; Morgan and Buice 2013.
31 Dolan 2014.
32 Fox and Lawless 2004; Lawless and Fox 2005.
33 Anzia and Berry 2011.
If women in office are viewed differently than men and adapt their behavior accordingly, then it is logical that they may avoid a risky activity (like corruption) while in office. This could occur because they are more at risk of getting caught and/or because they are more at risk of being punished harshly. Their higher risk of being caught derives from the fact that they are more likely to be under a microscope while in political office. Recent research on female candidates for executive office and women serving as presidents and prime ministers reports that the novelty of women in politics leads the media and voters to pay extra close attention to women’s actions and behaviors. A recent study of the US Senate finds that female voters are not blindly loyal to women in office simply because they are women but are, in fact, more likely to evaluate women in office more carefully based on the policies they promote while in office and hold them accountable.

Women’s higher probability of being punished for corruption results from the higher standard to which women are held. If women in office are stereotypically thought to be less corrupt than men, then they are likely to be more severely punished if they are accused of engaging in (or are perceived to be engaging in) corruption. Recent anecdotal evidence of female presidents (Laura Chinchilla in Costa Rica (2010–14) and Michelle Bachelet in Chile (2014 to present)) shows how quickly and severely their approval ratings have fallen in response to corruption scandals.

Summary
In sum, we argue that electoral accountability makes corruption risky, and therefore accountability should moderate the relationship between women’s representation and corruption for two reasons. First, women are more averse to the risks of engaging in corruption than men. Secondly, women may be more likely than men to be held accountable for corruption due to unequal treatment. If female legislators are less likely to engage in corruption than male legislators when accountability is high, then we should see this reflected in an aggregate relationship between women’s representation in legislatures/parliaments and corruption levels in a country. We expect a negative relationship between women’s representation and corruption when electoral accountability is strong, and this relationship will get weaker as electoral accountability gets weaker.

HYPOTHESES FOR ACCOUNTABILITY AND CORRUPTION
We identify four contexts in which voters should be able to hold elected representatives accountable for corruption – in other words, when they can more easily perceive corruption in government and punish corrupt officials at the polls – and, in turn, make corruption more risky: (1) when corruption is not a pervasive norm, (2) where press freedom is respected, (3) in parliamentary systems (as compared to presidential systems) and (4) when electoral rules establish direct and personalistic linkages between voters and elected legislators or members.

34 Bauer and Tremblay 2011; Murray 2010.
35 Jones 2014.
36 Another way to test this theory empirically would be to directly measure the extent to which elected officials will engage (or not) in corrupt activities. Convincing elected officials to participate in an experiment on corruption or even trying to survey them about their corrupt behavior or potential willingness to engage in corrupt behavior is challenging because of social desirability bias (among other reasons). Additionally, what is driving our study is not so much empirically evaluating the behavior of individual legislators, but trying to explain why an aggregate relationship between women’s representation and corruption varies across settings. Thus we think it is appropriate under these circumstances to test the aggregate country-level implications of our theory (which also has individual-level implications).
If our theory is correct, the empirical relationship between women’s representation in legislatures/parliaments and corruption should be statistically significant and negative in these settings of high accountability; the empirical relationship should be substantially smaller, and perhaps statistically insignificant, in low-accountability settings. In this section, we explain our reasoning for the link between our theory and these observable relationships.

**Corruption Norms**

Although corruption occurs in countries all over the world, research has found that democracies are less corrupt, on average, than non-democracies. But even within democracies, corruption is present in (and in some cases endemic to) the political system. Countries where corrupt behaviors (such as bribery and graft) are ‘rooted in widely shared expectations among citizens and public officials’ and become a normal part of doing government business have strong corruption norms. Measuring the presence of corruption norms is a challenge, but one proxy for it could be the (perceived) pervasiveness of corruption in politics and society. Where corruption is endemic and pervasive, corruption norms develop because corruption becomes the accepted and expected way that politics is done. Corruption norms do not develop, however, where corruption is not pervasive.

We use the pervasiveness of corruption as a proxy for corruption norms and one of the institutions (albeit an informal institution) of electoral accountability. In countries with pervasive corruption, the risk of corruption being detected and punished (that is, accountability) must be low in order for corruption to flourish. By comparison, a country with less corruption has (*ipso facto*) demonstrated a tendency to remove or exclude corrupt persons from government. The pervasiveness of corruption can moderate the relationship between women’s representation and corruption because less pervasive corruption (stronger corruption norms) increases the risk of engaging in corruption. Because women are more risk averse and aware of the differential treatment they may receive as officeholders, less pervasive corruption creates

---

37 An anonymous reviewer suggested two other possible contexts in which electoral accountability is higher: (a) in countries with stronger democratic institutions (see Esarey and Chirillo (2013) for relevant theoretical arguments along this line), as measured by Polity IV’s polity2 score (Marshall, Gurr, and Jaggers 2014), and in countries that lack electoral quotas for gender and therefore women are not guaranteed seats in parliament (Childs and Krook 2012; for related theory and evidence, see Franceschet and Piscopo 2008). Although we do not describe these contexts in detail in our article, we did empirically test each hypothesis and confirmed that the relationship between women in government and corruption is stronger in the presence of stronger democracy and in the absence of electoral gender quotas; see Appendix Table S6 for details.


39 Mishra 2006.

40 Helmke and Levitsky 2004.

41 Fisman and Miguel 2007.

42 We recognize that this argument appears tautological: if corruption norms were the only way in which accountability operated in democracies, then citizens would never be able to hold elected officials accountable in settings of high accountability. Similarly, the only time they could hold elected officials accountable would be when corruption is already low. However, we know that other mechanisms of accountability exist – we discuss three others in this article – and as a result, corruption norms are rarely operating in isolation from other forces for accountability. We do think that high levels of corruption make it difficult to hold elites accountable, and moving from a high-corruption to a low-corruption environment via accountability will be a slow process. But norms change, albeit slowly, and the presence of other institutions that increase electoral accountability can help destroy the norms of corruption present in various countries. The critical distinction is between (a) existing levels of corruption and (b) the degree to which corruption responds to changes in women in government as a function of that level: the lower the level of corruption, the more responsive that corruption will be to the proportion of women in government. Our operationalization reflects that distinction.
a stronger disincentive for women to engage in corruption than men. As a result, we expect a stronger link between women’s representation in government and corruption in countries with weaker corruption norms.43

HYPOTHESIS 1: The relationship between female share of the legislature and corruption level will be more negative in states with low prior levels of corruption compared to states with high prior levels of corruption.

Some evidence supporting this hypothesis has already been presented in prior work. For example, Chaudhuri44 reviews multiple experimental studies of the propensity to commit various corrupt behaviors (such as offering or accepting bribes).45 He finds that there is substantial heterogeneity in female behavior across multiple experiments. In some experiments, women are less likely to offer a bribe than men, but in others women are statistically indistinguishable from men. He suggests that one of the key contextual factors may be the degree to which corruption is endemic to its political and economic culture: ‘evidence for greater incorruptibility on the part of women comes primarily from developed nations. We do not find strong differences in developing countries where the problem of corruption is far more endemic’.46

Press Freedom
A second contextual factor that could affect the relationship between women’s representation and corruption in a democracy is the freedom of the press. The ability of citizens to identify corrupt officials is at least partly conditional on the ability of the media to investigate and report on allegations of corruption. Brazil’s now-infamous mensalão scandal, for example, came to light when several newspapers and news magazines produced a series of news stories alleging that the governing Worker’s Party (PT) was paying opposition legislators monthly salaries to support the governing party’s legislative agenda.47 In the aftermath of the scandal, several deputies were forced from office, and the PT lost eight seats in the 2006 Chamber of Deputy elections – the first time since the transition to democracy in 1985 that it lost seats rather than gained them.

We argue that corruption is riskier in countries with a freer press compared to those where the government restricts press freedom because the risk of detection, and consequently punishment, is higher where journalists are free to investigate and expose corruption.48 The greater risk of detection and punishment in countries with a free press should in turn lead women in office to be proportionally less likely to engage in corruption compared to men, resulting in a stronger relationship between female participation in government and corruption.

HYPOTHESIS 2: The relationship between female share of the legislature and corruption level will be more negative in countries with a free press than in those with an unfree press.

Parliamentary Governance
A third contextual factor influencing accountability for corruption in a democracy is the nature of the separation of powers. Research on the differences between parliamentary and presidential...
systems has long debated the strengths of each in terms of accountability. Scholars concerned about the fragility of democracy in presidential systems often argue that parliamentary systems are better for democracy because the fixed terms inherent to presidential systems make it impossible to bring an end to unpalatable governments in any way other than the breakdown of democracy.\textsuperscript{49} The ability to call a vote of no confidence in parliamentary systems, in contrast, gives voters an opportunity to preserve democracy but turn over the government more quickly. Linz notes one of the key drawbacks\textsuperscript{50} of the fixed terms of presidential systems: ‘It breaks the political process into discontinuous, rigidly demarcated periods, leaving no room for the continuous readjustments that events may demand’.\textsuperscript{51} He later explicitly relates this to corruption, saying ‘parliamentary systems, precisely by virtue of their surface instability, often avoid deeper crises. A prime minister who becomes embroiled in scandal or loses the allegiance of his party or majority coalition and whose continuance in office might provoke grave turmoil can be much more easily removed than a corrupt or highly unpopular president’.\textsuperscript{52}

We build on Linz’s logic and argue that the absence of fixed terms in parliamentary systems should strengthen accountability for corruption. Indeed, there is already empirical evidence that parliamentary systems have lower levels of perceived corruption than presidential ones, although the causal pathway identified varies.\textsuperscript{53} In parliamentary systems, the chief executive, cabinet and parliament’s terms in office are not fixed and elected officials constantly face the threat of being held to account by voters at any time. When a corruption scandal breaks, the absence of fixed terms for parliament, the threat of a vote of no confidence, and the fact that a no confidence vote not only causes the member of parliament (MP) to suffer defeat but can bring down the entire government mean that the punishment for an MP and a party is severe, and thus corruption is risky. In presidential systems, fixed terms mean that punishment may be delayed to the end of the term in office, giving elites time to rebuild their images prior to being held to account by voters, and the separation of powers means that actions in the legislature do not necessarily threaten the government itself. Thus, we argue that corruption is riskier in parliamentary systems. Because of women’s greater behavioral response to this risk (attributable to greater risk aversion and/or differential treatment by voters), the link between women’s representation and lower levels of corruption should be strongest in parliamentary systems.

**HYPOTHESIS 3:** The relationship between female share of the legislature and corruption level will be more negative in parliamentary systems when compared to presidential systems.

**Personalism**

Finally, we directly examine the strength of the link between elected representatives and voters – the degree of personalism produced by the electoral system. Existing research has produced mixed

\textsuperscript{49} Linz 1990, 1994.

\textsuperscript{50} Defenders of presidentialism have pointed out some of the strengths of accountability in presidential systems: for example, voters have the opportunity to hold the executive and legislature independently accountable for government (Hellwig and Samuels 2008; Mainwaring and Shugart 1997; Persson, Roland, and Tabellini 1997; Samuels and Shugart 2003; Shugart and Carey 1992). However, this also means that voters may have a more difficult time assigning blame due to the separation of powers inherent in presidential systems (Samuels and Shugart 2003; Shugart and Carey 1992); each branch of government can blame the other. In this article, we cannot empirically distinguish corruption in the executive branch from corruption in the legislative branch, making it impossible to test this angle of the accountability argument.

\textsuperscript{51} Linz 1990, 54.

\textsuperscript{52} Linz 1990, 64.

\textsuperscript{53} Gerring and Thacker 2004; Lederman, Loayza, and Soares 2005; Treisman 2007; but see Persson and Tabellini 2002.
findings regarding the effects of electoral rules on corruption. Persson, Tabellini and Trebbi\textsuperscript{54} and Kunicová and Rose-Ackerman\textsuperscript{55} link electoral rules to voters’ ability to monitor elected officials and find that stronger ties between constituents and individual elected representatives produce lower levels of corruption. In contrast, Chang\textsuperscript{56} and Chang and Golden\textsuperscript{57} find that electoral systems that produce incentives to cultivate personal votes (measured as open-list proportional electoral systems with high district magnitudes) have higher levels of corruption, which they argue results from candidates having greater incentives to seek illegal funds for their campaigns in more personalistic systems. Attempting to mediate these divergent findings, Treisman found that the relationships between electoral rules and corruption were often indeterminate\textsuperscript{58}.

We argue that more personalistic rules should strengthen the effect of women’s representation on corruption. Personalistic electoral rules create tighter ties between voters and their elected representatives, while less personalistic rules emphasize the mediating role of parties in the voter–representative linkage.\textsuperscript{59} The risk of being punished for corrupt behavior is therefore greater in personalistic systems because voters can individually identify their representative and hold them directly accountable. In less personalistic (more party-centric) systems, elites may be able to hide inside the party organization and deflect direct punishment at the polls. Voters may be willing to swallow one bad egg the party wants to defend if they are supportive of the party more generally. Parties may even collaborate to conceal the individual guilt of one member to preserve their collective electoral viability.

Because of the stronger electoral accountability created by personalistic systems, we claim that the individual risk of corrupt behavior is greater in these systems. Our theory predicts that this risk deters women in office from engaging in corruption more strongly than men, and as a result, the link between female representation in government and corruption is stronger than in party-centered systems.

**HYPOTHESIS 4:** The relationship between female share of the legislature and corruption level will be more negative in personalistic systems than in party-centric systems.

**DATA AND VARIABLES**

The dataset that we use is from Schwindt-Bayer and Tavits, and it contains measures of corruption perceptions, women’s representation in the legislature, accountability indicators and control variables for seventy-six democratic-leaning countries from 1990–2010,\textsuperscript{60} summary statistics are reported in Table 1. The dataset includes all countries and years for which Freedom House’s average Civil Liberties and Political Rights scales\textsuperscript{61} was 5 or lower and Polity IV’s polity2 score was greater than 0 for twelve years or more.\textsuperscript{62} The dataset also requires that, during this 12+ year period, countries have a consistent executive structure (presidential or...
parliamentary) and to not be missing all (or nearly all) data for any variable. These selection criteria have three main advantages: (1) they exclude countries that do not function according to the rules and norms of minimal democracy, (2) they include both semi-democracies and full democracies to allow generalization across degrees of democracy and (3) they allow sufficient time points and data availability to conduct a panel analysis.

The dependent variable is the perceived level of corruption in countries as determined by three widely accepted country-level measures of corruption: Transparency International Corruption Perceptions Index (TI CPI), which measures ‘the abuse of public office for private gain’; the World Bank Governance Indicators Control of Corruption measure (WBGI), which measures ‘the extent to which public power is exercised for private gain, including both petty and grand forms of corruption as well as “capture” of the state by elites and private interests’; and the Political Risk Services’ International Country Risk Guide’s (ICRG) corruption risk measure, which measures ‘bribery […] excessive patronage, nepotism, job reservations, “favor-for-favors,” secret party funding, and suspiciously close ties between politics and business’. Because corruption is notoriously difficult to assess, cross-national research often relies on corruption perceptions as a measure of underlying corruption; we believe these measures are advantageous because of their comprehensive nature and their wide availability over space and time. All three measures are created from surveys and expert

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Dataset Summary Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>TI CPI</td>
<td>4.790</td>
</tr>
<tr>
<td>ICRG</td>
<td>2.576</td>
</tr>
<tr>
<td>WBGI</td>
<td>2.175</td>
</tr>
<tr>
<td>FH press freedom</td>
<td>−31.097</td>
</tr>
<tr>
<td>Presidentialism</td>
<td>0.412</td>
</tr>
<tr>
<td>Personalism</td>
<td>4.856</td>
</tr>
<tr>
<td>% women in lower house</td>
<td>15.731</td>
</tr>
<tr>
<td>FH freedom score</td>
<td>−2.093</td>
</tr>
<tr>
<td>log GDP per capita</td>
<td>8.567</td>
</tr>
<tr>
<td>% protestant</td>
<td>23.320</td>
</tr>
<tr>
<td>Trade imbalance (% of GDP)</td>
<td>78.649</td>
</tr>
<tr>
<td>Women’s economic rights</td>
<td>1.590</td>
</tr>
</tbody>
</table>

(From note continued)

the regulation of political participation and the competitiveness of participation (Marshall, Gurr, and Jaggers 2014).

63 Transparency International 2011, 2.

64 Kaufmann, Kraay, and Mastruzzi 2010, 4.

65 Political Risk Services Group 2012. The ICRG measures the risk that corruption presents to foreign business and investment (Political Risk Services Group 2012, 5–6). This means that the ICRG index does not just capture raw levels of corruption, but the degree to which the state’s institutions convert this corruption into a threat to businesses (e.g., by threatening the stability of a government). A country’s democratic accountability and tolerance of corruption might therefore influence the ICRG rating because states with greater accountability or less tolerance might be more likely to experience political turmoil as a result of corruption scandals (Lambsdorff 2006, 82–3).

66 The use of perception-based corruption measures has been hotly debated in recent years (Donchev and Ujhelyi 2014; Provost 2013). The primary concern is that the subjective perception of corruption is not necessarily identical with its reality. However, alternative ‘objective’ measures of corruption are also subject to criticism: ‘since corruption
assessments of country-level corruption, and each measure has strengths and weaknesses.\textsuperscript{67} By examining all three, we strengthen the robustness of our conclusions. The three measures correlate very highly with one another as well as with several alternative measures of corruption, which bolsters their validity.\textsuperscript{68}

We focus on the TI CPI (available from 1995–2010) in presenting our results, but our primary findings are similar regardless of whether we use the ICRG (available from 1990–2010) or the WBGI (available from 1996–2010, with biannual measurements between 1996 and 2002). The TI CPI measure is a scale of 0 to 10, the ICRG measure is a scale of 0 to 6\textsuperscript{69} and the WBGI measure is a scale of −2.5 to 2.5. The original coding of all of these variables is such that higher numerical values indicate less perceived corruption (or more perceived government control of corruption). However, for ease of interpretation, we have recoded all three variables so that higher values equal more perceived corruption.\textsuperscript{70}

Our main independent variables are the percentage of the lower house of the legislature/parliament\textsuperscript{71} that is female\textsuperscript{72} and four measures of accountability in the political system:

1. a one-year time lag of the dependent variable (specific to the corruption measure under analysis) to capture corruption norms in a country,\textsuperscript{73}
2. the Freedom House’s Freedom of the Press measure, which we recode to range from −80 to 0 in order of increasing freedom,\textsuperscript{74}

\textsuperscript{(F note continued)}

is clandestine, it is virtually impossible to come up with precise objective measures of it. … There should be no presumption that objective data is necessarily more informative than reports from experts, citizens, or firms on the ground – irrespective of their extent of perception or subjectivity’ (Kaufmann, Kraay, and Mastruzzi 2007, 4). For example, consider two ‘objective’ alternatives: contract-intensive money (CIM) and the Global Corruption Barometer (GCB) survey measure of bribes paid to legal and judiciary institutions. Although not dependent on subjective perceptions or available for many countries and time periods, CIM does not solely measure corruption. Specifically, CIM is ‘the ratio of non-currency money to the total money supply’ (Clague et al. 1999, 188), as compiled by Mark Souva (Johnson, Souva, and Smith 2013), and therefore is a measure of citizens’ willingness to hold non-cash monetary assets. It is designed to measure ‘the enforceability of contracts and the security of property rights’ (p. 185) including ‘not only the risk of government expropriation of financial assets (for example, through bank nationalization), but the expropriation through arbitrary regulation or outright confiscation of any type of fixed asset’ (p. 203).

Freedom from corruption constitutes only one aspect of secure property rights. The GCB legal/judicial bribery variable is more narrowly defined than any comprehensive definition of corruption would imply: it is the proportion of respondents in a country-year indicating that someone in their household paid a bribe to the legal/judicial system (Teorell et al. 2015 codebook, 254; Transparency International 2015). While bribery of these officials is one aspect of corruption, corruption can take many other forms and involve many other government and non-government officials. Additionally, this measure is available for a relatively limited number of countries and time periods compared to the TI CPI, WBGI and ICRG.

\textsuperscript{67} Knack 2007; Lambsdorff 2006; Treisman 2007.
\textsuperscript{68} WBGI and TI CPI correlate at $r = 0.98$; ICRG correlates with WBGI at $r = 0.87$ and with TI CPI at $r = 0.86$. In Appendix Figure S1, we show strong associations between TI CPI and the two ‘objective’ measures noted in the preceding footnote, CIM and GCB Legal/Judicial Bribery.
\textsuperscript{69} ICRG data were monthly up through mid-2009. In those cases, we use the twelve-month average score.
\textsuperscript{70} The TI CPI measure is recoded by 10 minus the original value of the dependent variable. The ICRG measure is recoded by 6 minus the original value of the dependent variable. The WBGI measure is recoded by 2.6 minus the original value of the dependent variable.
\textsuperscript{71} We believe that focusing on women in the legislature in this analysis is appropriate because our accountability measures are focused on accountability to voters.
\textsuperscript{72} Inter-Parliamentary Union 2012.
\textsuperscript{73} Our results are robust to using a two- or three-year lag instead of a one-year lag in this model; see Appendix Table S3 for details.
\textsuperscript{74} Freedom House assesses freedom of the press in all countries every year. Their measure assesses freedom in print, broadcast and internet media by creating a sub-score for each media type of the following ways in which media freedom can be restricted: laws and regulations that influence media content, political pressures and controls on media content, economic influences over media content and repressive actions (www.freedomhouse.org). These are
(3) a dichotomous coding of whether a country’s form of government is presidential (coded as 1) or parliamentary (coded as 0) and (4) a measure of the degree of personalism produced by a country’s parliamentary or legislative electoral system. Personalism ranges from 1 to 13 in order of increasing levels of personalism. Each of these four measures of accountability is interacted with the percentage of women in the lower house of the parliament/legislature to allow the relationship between female participation in government and corruption to be conditional on the accountability variable.

We also include a set of common control variables for these kinds of corruption models: the percentage of citizens who are Protestant; democratic freedom, measured as the average political rights and civil liberties Freedom House scores inverted such that higher scores indicate greater freedom; level of economic development, as measured by logged GDP per capita; trade imbalance, measured as imports minus exports as a percentage of GDP; and women’s economic rights, as measured in the Cingranelli-Richards Human Rights Dataset. These measures block possible sources of spurious correlation attributable to cultural, socioeconomic and political explanations for variation in levels of corruption across countries and over time.

STATISTICAL METHODS

Our approach to analyzing and presenting our empirical evidence is straightforward: we consider each of our four accountability variables in turn. For each one, we first use a scatterplot to examine the pooled bivariate relationship between the TI CPI dependent variable and the percentage of women in government. To determine whether this relationship changes with the strength of electoral accountability for corruption, as would be consistent with Hypotheses 1–4, we split the data into high and low values on the accountability variable and construct separate scatterplots for each.

Secondly, we verify the findings of the bivariate plot by constructing a multivariate linear regression model. We include a lagged dependent variable in all models because we believe

(Fe) note continued

aggregated into a scale that runs from 0 to 100 (in order of decreasing freedom). No country in the dataset had levels higher than 80 because it excludes non-democracies, where press freedom is likely to be most restricted.

The dataset authors coded semi-presidential systems as presidential or parliamentary depending on the powers of the president. Specifically, premier-presidential systems were coded as parliamentary systems in which the president has no power to dissolve the cabinet (only the assembly can) and president-parliamentary systems as presidential where the president has the power to dissolve the cabinet alongside the assembly (Elgie 2011; Samuels and Shugart 2010).

Johnson and Wallack 1997. Johnson and Wallack’s personalism score has become a common measure of how strongly certain configurations of electoral rules incentivize personalistic rather than party-centered behavior among candidates and elected representatives. They use Carey and Shugart’s (1995) schema for coding electoral systems according to the extent to which the ballot structure allows voters to disturb party lists, how votes are pooled across a ballot and the type of vote a voter places. Configurations of scores are then ranked by how much personalism they create, and the electoral system of a country is classified accordingly.


CIA World Factbook 2013.

Freedom House 2014.

World Bank 2013.

World Bank 2013.

Cingranelli and Richards 2010.

Random effects variants of these models are substantively no different from standard ordinary least squares regressions; the random effects explain no appreciable portion of variance when added. Fixed-effects (FE) models and system GMM dynamic panel data models produce weaker and inconsistent findings, albeit with some qualitative similarities to our main results (see Appendix Tables S4 and S5). We argue that the models we present are more credible; consider the comparison with FE models. First, FE models are inefficient in the presence of
that corruption is a path-dependent process, the presence of which is a function of its history; this variable also models temporal dependence in the data. We include year and geographical region dummies to account for additional temporal and spatial dependence in the data.

We plot the marginal effect of percentage women in the legislature on the dependent variable at different values of the accountability variable, as prescribed by Brambor, Clark and Golder, to determine whether the relationship between perceived corruption and women in government is stronger when individual accountability for corruption is stronger (as indicated by Hypotheses 1–4).

Some variables (including the TI CPI and WBGI dependent variables, press freedom, personalism, trade imbalance, and women’s economic rights variables) have missing observations in our dataset. Simply deleting the observations with partially missing data can lead to biased and inefficient estimates in cases where stochastic multiple imputation of the dataset would not. Consequently, when we estimate our model, we use multiple imputation with chained equations as implemented in Stata 14.2 to perform regression including the

\( \text{X} \) (note continued)

short panels thanks to an incidental parameters problem (Hsiao 2003, 48–9; Neyman and Scott 1948); in our dataset, even the dependent variable with the greatest availability (the ICRG) has just \( \approx 20 \) observations per panel. Secondly, FE models are inefficient for estimating the effect of slow-moving independent variables, and all of our main variables are very slow moving within panels; fixed effects alone explain 83 per cent of the variance in press freedom, 99 per cent of the variance in presidentialism and 86 per cent of the variance in personalism. Thirdly, from a theoretical perspective, we do not believe in (or wish to model) persistent country-level variation in corruption net of the path-dependent history of corruption being captured by the lagged dependent variable and the other institutional influences being captured by our control variables. For example, corruption in the United States was widespread in the latter nineteenth century but comparatively low by the end of the twentieth century; a fixed effect presumes that this characteristic is essentially permanent. Finally, FE models are known to be biased in short panels in the presence of lagged dependent variables (Judson and Owen 1999; Nickell 1981), and the lagged dependent variable is a theoretically relevant variable for the analysis. Dynamic panel data models address only this last problem: models with a lagged dependent variable are consistent in short panels when the number of panels is large (Roodman 2006). But a new problem is created in its place: dynamic panel data models are not supported after multiple imputation (using mi estimate) in Stata 14.2, and thus any problems inherent to the missing data problem reappear. Additionally, we have a relatively small number of panels to support an argument of consistency. We do not have confidence that this model is compatible with an interaction between the percentage of women in the legislature and a lag of the dependent variable, and so do not estimate this model.

The regions are Sub-Saharan Africa, South Asia, East Asia, South East Asia, Pacific Islands/Oceania, Middle East/North Africa, Latin America, Caribbean and non-Iberic America, Eastern Europe/Soviet Union and Western Europe.

For the ICRG dependent variable, our main regression models use datasets that have 1.3 per cent (lag DV and presidentialism models), 1.7 per cent (personalism model) or 13.2 per cent (press freedom model) of cases with missing observations. For the TI CPI dependent variable, our main regression models use datasets that have \( \approx 18.7 \) per cent of cases with missing observations. For the WBGI dependent variable, our main regression models use datasets that have \( \approx 25.7 \) per cent of cases with missing observations.

Multiple imputation by chained equations (MICE) generates multiple imputation datasets by (1) eliminating any observations with missing values for all variables; (2) substituting random values for missing values in any remaining observations; (3) imputing the values of a missing variable \( X_{i} \) using model predictions from a GLM model of all the other variables \( X_{-i} \) on the (non-missing) values of \( X_{i} \), where the model includes observations with imputed values of \( X_{-i} \) where \( X_{i} \) is non-missing; (4) repeating step 3 for all values of \( i = 1...k \) in sequence for the \( k \) independent variables; (5) repeating steps 3–4 a large number of times to refine the predicted missing values; and finally (6) repeating steps 2–5 with new initial values \( M \) times to generate \( M \) imputation datasets. The resulting datasets are analyzed and the results combined using the method of Rubin (1996). See Royston and White (2011) for more details of the implementation of MICE in Stata.

Royston and White 2011.
partially missing cases while incorporating uncertainty about the unknown true values of the missing variables.

**EVIDENCE: GENDER, ACCOUNTABILITY AND CORRUPTION**

As described above, we have four hypotheses about how accountability should influence the relationship between women’s political representation and corruption. In this section, we show evidence associated with each hypothesis in turn. Note that our statistical methods only look for a *correlational* relationship between women’s representation and corruption at different levels of accountability. A specific pattern of correlations is predicted by our theory (that women’s representation in government causes lower corruption only when electoral accountability is high) and we seek to match those predictions as evidence for our theory, but we cannot definitively determine a direction of causality. We discuss empirical causal modeling strategies in the conclusion as suggestions for future research, but in this article our goal is simply to establish whether there is an empirical relationship between women’s representation and corruption that is conditional on electoral accountability and consistent with our theoretical predictions.

**Hypothesis: Corruption Norms**

Our first hypothesis is that the relationship between the female share of the legislature and the perceived corruption level should be stronger (more negative) in democracies with low prior levels of corruption compared to democracies with high prior levels of corruption. As Figure 1 indicates, we find evidence for this relationship in our data. The simple bivariate scatterplots with the linear prediction included show that the percentage of the legislature/parliament that is female is not associated with perceived corruption in countries with high levels of prior perceived corruption. Where prior perceived corruption levels are low, greater levels of women’s representation in the lower house of parliament are strongly associated with lower levels of perceived corruption.

Table 2 confirms this pattern in a multivariate regression using all three measures of corruption. The interaction between the percentage of the legislature that is female and the lagged measure of corruption perceptions is positive and statistically significant in all three models. Figure 2 presents the marginal effect of women’s representation on perceived corruption.

---

90 As described in Footnote 37, and at the suggestion of an anonymous reviewer, we also find evidence that the negative relationship between women in government and perceived corruption is stronger in the presence of stronger democratic institutions (as measured by the Polity score) and in the absence of electoral gender quotas; these are both environments where electoral accountability may be stronger. See Appendix Table S6 for the results.

91 As an initial effort to establish some credible evidence for a causal effect of women’s representation on corruption, we estimate a two-stage least squares (2SLS) version of our TI CPI model in Tables 3, 4 and 5 by using two-period lags of the key independent variables (including interaction terms) as instruments; this identification strategy is suggested by Reed (2015). The results are shown in Appendix Table S2. The substantive findings from this model are similar to those from our OLS models, although the relationship between women’s representation and corruption becomes statistically insignificant in the presidentialism instrumental variable (IV)/2SLS model; this change might be due to efficiency loss because we can no longer use multiple imputation for the IV/2SLS model.

92 The TI CPI time series passes the Augmented Dickey-Fuller and Phillips-Perron unit root tests with $p < 0.01$ using the inverse $\chi^2$ transformation, which indicates that the series is stationary and the state-dependence model can be used (Esarey and DeMeritt 2014, 74–6).

93 When the lagged dependent variable is interacted with the percentage of women in the legislature, a state-dependent dynamic model is created (Esarey and DeMeritt 2014). In a state-dependent system, the effect of an
The past prevalence of corruption influences the relationship between gender and corruption. The figure shows the relationship between the TI CPI and the percentage of women in the lower house for seventy-six democratic-leaning countries between the years 1996–2010; the top panel shows countries with prior TI CPI scores > 5 and the bottom panel shows countries with TI CPI scores ≤ 5. The difference between the slopes is 0.070, which is statistically significant (p < 0.001).

Fig. 1. How does the past prevalence of corruption influence the relationship between gender and corruption?

Note: the figure shows the relationship between the TI CPI and the percentage of women in the lower house for seventy-six democratic-leaning countries between the years 1996–2010; the top panel shows countries with prior TI CPI scores > 5 and the bottom panel shows countries with TI CPI scores ≤ 5. The difference between the slopes is 0.070, which is statistically significant (p < 0.001).

At the suggestion of an anonymous reviewer, we conducted a robustness check for all our hypothesis tests using the TI CPI dependent variable and adding two additional control variables: years since women’s suffrage was granted without restrictions (Inter-Parliamentary Union n.d.) and official development assistance from the
When prior perceived corruption levels range from 0 to about 4.5, increasing women’s representation correlates with less perceived corruption to a statistically significant degree. At a prior corruption score of 2, the present corruption score would be ≈0.02 lower for every one-percentage-point higher value of women in parliament. This indicates that a state with a 40 per cent share of women in the legislature would have a 0.80 point lower present TI corruption score compared to a state with no women in parliament; this is about 8 per cent of the maximum difference possible on this perceived corruption scale. The finding is consistent with our theoretical argument that the gender–corruption relationship is sensitive to electoral accountability.

(For note continued)

Organisation for Economic Co-operation and Development recorded in replication data for Lebovic and Voeten (2009). The results, which are substantively similar to those presented in the main body of the article, are shown in Appendix Table S1.

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>How Does the Past Prevalence of Corruption Influence the Relationship Between Gender and Three Measures of Corruption?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>lag TI CPI</td>
<td>0.583***</td>
</tr>
<tr>
<td>lag ICRG</td>
<td>(15.00)</td>
</tr>
<tr>
<td>lag WBGI</td>
<td></td>
</tr>
<tr>
<td>% women in lower house</td>
<td>−0.0303***</td>
</tr>
<tr>
<td>% women × lag DV</td>
<td>0.00531***</td>
</tr>
<tr>
<td>FH freedom</td>
<td>−0.198****</td>
</tr>
<tr>
<td>log GDP per capita</td>
<td>−0.365****</td>
</tr>
<tr>
<td>% protestant</td>
<td>−0.00245*</td>
</tr>
<tr>
<td>Trade imbalance (% of GDP)</td>
<td>−0.000667</td>
</tr>
<tr>
<td>Women’s economic rights</td>
<td>−0.0168</td>
</tr>
<tr>
<td>N</td>
<td>1,176</td>
</tr>
</tbody>
</table>

Note: the table reports the output of ordinary least squares regression models using three dependent variables: (1) Transparency International Corruption Perception Index (TI CPI); (2) the International Country Risk Guide corruption rating (ICRG); and (3) the World Bank Governance Indicators Control of Corruption measure (WBGI). All three measures have been recoded so that higher values on each DV indicate more corruption. The data includes seventy-six democratic-leaning countries; the time dimension spans 1995–2010 for the TI CPI variable, 1996–2010 for the WBGI variable, and 1991–2010 for the ICRG variable. Year and region dummies are included in the models, though not reported in this table. Estimates are based on multiple imputation into fifty datasets using chained equations. R² for the models are: (1) 0.919, (2) 0.931, (3) 0.867. t statistics in parentheses. *p < 0.05, **p < 0.01, ***p < 0.001
Interestingly, the model also indicates that there is a statistically significant and positive relationship between women’s participation in government and perceived corruption at the highest lagged values of perceived corruption (between $\approx 7.5$ and $10$ on the TI CPI scale); however, only about 9 per cent of our observations lie in this range.

**Hypothesis 2: Press Freedom**

We also find evidence that press freedom is associated with the relationship between women’s representation and corruption in a way that is consistent with our theory of electoral accountability. The bivariate scatterplot shown in Figure 3 shows no relationship between gender and corruption perceptions when press freedom is restricted, but a strong negative relationship in countries with high levels of press freedom. This is consistent with the idea that the greater risk of detection and punishment for corruption that is created by a free press disproportionately affects the behavior of women.

Table 3 shows that this finding is supported by the results of a multivariate regression: a statistically significant interaction effect exists between women’s representation in parliament and press freedom for all three measures of corruption. The relationship is most clearly seen in Figure 4, which illustrates the marginal effect of women’s parliamentary representation on corruption perceptions as press freedom increases based on the TI CPI results from Table 3.

The estimated marginal effect of women’s representation on perceived corruption becomes negative and statistically significant when press freedom is in the top third of its range (about $-25$ to $0$). When press freedom is at $-10$, the marginal effect of women in parliament is $\approx -0.02$. Once again, this implies that countries with a 40 per cent female parliament are, on average, about 0.8 lower in the TI corruption perceptions measure compared to a country with no women in parliament. As Hypothesis 2 indicated, a larger share of women in parliament is associated with lower levels of corruption when the press is free, but not when the press is restricted.

**Hypothesis 3: Parliamentary Governance**

The relationship between women’s representation and corruption perceptions in our data is different across types of democratic government. In presidential systems, women’s
representation in legislatures has no discernible relationship with perceived corruption, whereas in parliamentary systems, greater women’s representation correlates with considerably lower levels of perceived corruption. These divergent patterns are striking in the bivariate relationships depicted in Figure 5.

As Table 4 shows, multivariate regression models support the bivariate findings: the interaction between the percentage of female legislators and the presidentialism dummy variable is statistically significant and positive in all three models. The marginal effect plot in Figure 6 shows the relationship between the percentage of the legislature/parliament that is female and corruption perceptions estimated in the TI CPI model in Table 4. While greater women’s representation has no statistically significant relationship with the level of perceived corruption in presidential systems, it has a strong and statistically significant negative relationship in parliamentary systems of $\approx -0.01$ – about half the substantive magnitude of the relationships in the prior two contexts.

Fig. 3. How does press freedom influence the relationship between gender and corruption?
Note: the figure shows the relationship between the TI CPI and the percentage of women in the lower house for seventy-six democratic-leaning countries between the years 1995–2010; the top panel shows countries with press freedom scores $\leq -30$ and the bottom panel shows countries with press freedom scores $> -30$. The difference between the slopes is 0.110, which is statistically significant ($p < 0.001$).
This finding supports our theoretical argument that parliamentary systems present a greater individual risk for corruption due to the threat of swift sanctioning by voters, which creates a larger gender difference in corruption behavior. This difference becomes manifest in a stronger negative relationship between the perceived level of corruption and the share of women in the legislature in parliamentary systems compared to presidential systems.

Hypothesis 4: Personalism

Finally, we examine how the relationship between women’s representation and corruption is influenced by the personalism embedded in legislative or parliamentary electoral rules. We find that more personalistic rules are associated with a stronger negative relationship between the percentage of women in parliament and perceived corruption. Figure 7 shows the bivariate scatterplots and

<table>
<thead>
<tr>
<th>Table 3</th>
<th>How Does Press Freedom Influence the Relationship Between Gender and Three Measures of Corruption?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>TI CPI</td>
</tr>
<tr>
<td>lag TI CPI</td>
<td>0.681***</td>
</tr>
<tr>
<td></td>
<td>(21.07)</td>
</tr>
<tr>
<td>lag ICRG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>lag WBGI</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>% women in lower house</td>
<td>-0.0263***</td>
</tr>
<tr>
<td></td>
<td>(-4.46)</td>
</tr>
<tr>
<td>Press freedom</td>
<td>0.00951*</td>
</tr>
<tr>
<td></td>
<td>(2.22)</td>
</tr>
<tr>
<td>% women × press freedom</td>
<td>-0.000802***</td>
</tr>
<tr>
<td></td>
<td>(-4.20)</td>
</tr>
<tr>
<td>FH freedom</td>
<td>-0.181**</td>
</tr>
<tr>
<td></td>
<td>(-2.94)</td>
</tr>
<tr>
<td>log GDP per capita</td>
<td>-0.324***</td>
</tr>
<tr>
<td></td>
<td>(-7.15)</td>
</tr>
<tr>
<td>% protestant</td>
<td>-0.00284*</td>
</tr>
<tr>
<td></td>
<td>(-2.38)</td>
</tr>
<tr>
<td>Trade imbalance (% of GDP)</td>
<td>-0.000409</td>
</tr>
<tr>
<td></td>
<td>(-0.59)</td>
</tr>
<tr>
<td>Women’s economic rights</td>
<td>-0.0383</td>
</tr>
<tr>
<td></td>
<td>(-0.83)</td>
</tr>
<tr>
<td>N</td>
<td>1,176</td>
</tr>
</tbody>
</table>

Note: the table reports the output of ordinary least squares regression models using three dependent variables: (1) the Transparency International Corruption Perceptions Index (TI CPI); (2) the International Country Risk Guide corruption rating (ICRG); and (3) the World Bank Governance Indicators Control of Corruption measure (WBGI). All three measures have been recoded so that higher values on each DV indicate more corruption. The data includes seventy-six democratic-leaning countries; the time dimension spans 1995–2010 for the TI CPI variable, 1996–2010 for the WBGI variable and 1991–2010 for the ICRG variable. Year and region dummies are included in the models, though not reported in this table. Estimates are based on multiple imputation into fifty datasets using chained equations. R² for the models are: (1) 0.922, (2) 0.931, (3) 0.873. t statistics in parentheses. *p < 0.05, **p < 0.01, ***p < 0.001
linear predictions for democratic-leaning countries with more party-centered (less personalistic) electoral rules compared to more personalistic electoral rules. Both figures show a negative relationship, but the effect is slightly steeper in democratic-leaning countries with more personalistic rules (and the difference between the slopes is statistically significant at \( p < 0.001 \)).

Table 5 shows our multivariate statistical models with personalism interacted with the percentage of female legislators. The interaction terms are negative and statistically significant in all three models. Figure 8 shows the TI CPI model’s marginal effect for women in parliament on perceived corruption at varying levels of personalism. The effect is not statistically significant in the least personalistic systems (where the personalism score is less than about 2.5); this encompasses 33 per cent of the sample of country-years (that is, about two-thirds of the sample has more personalistic electoral rules). As personalism increases from 3 to 13, the effect of women’s representation on corruption perceptions is negative and statistically significant. At a personalism value of 13, the marginal effect of women in parliament is about \(-0.03\); this means that a country with 40 per cent women in parliament is expected to have a corruption score 1.2 points lower than a country with no women in parliament. This supports our argument that electoral rules that produce a stronger accountability link between individual representatives and voters disproportionately deter women from engaging in corruption.

In sum, we observe that the relationship between the level of women’s representation and the perceived level of corruption is indeed conditional upon the strength of individual accountability to voters in the political system. This finding corresponds to the implications of our theoretical argument: accountability moderates the relationship between women’s representation and corruption through the mechanisms of greater risk aversion and/or higher standards of accountability for women.

**CONCLUSION**

Corruption is a political threat that all countries fight, with varying degrees of success. In some countries, corruption levels are low and instances of suspected corruption are quickly brought to justice. The recent convictions of former Illinois governor Rod Blagojevich and former New Orleans mayor Ray Nagin in the United States exemplify this. Corruption is a risky activity for
political elites in these settings. In other countries, like Mexico and Venezuela, corruption levels are persistently high and individual cases of corruption rarely make headlines or produce negative consequences for those involved. Participating in corruption is therefore not particularly risky in these locales and may even be a way for elites to further their political careers. Previous studies have found that women’s representation in government is associated with lower levels of corruption, leading some to think that increasing women’s election to office will reduce corruption in countries. Yet this finding is not consistent across countries.

In this article, we asked: why are women’s representation and reduced corruption related in some countries but not others? We argued that greater women’s representation in parliaments and legislatures is more strongly associated with lower levels of corruption in countries with higher electoral accountability, that is, where voters can identify corrupt officials and punish them at the ballot box. We explained this conditional relationship with two theoretical mechanisms that relate the relationship between women’s representation and corruption to the
risks of corruption – women’s greater risk aversion and the different ways in which voters treat them. We generated four institutional hypotheses about the rules and norms that influence electoral accountability and tested these hypotheses with data from seventy-six democratic-leaning countries around the world. We found consistent evidence that, where accountability is high, a strong negative relationship exists between women’s representation and perceived corruption levels. Where accountability is low, a much weaker relationship exists. Strong electoral accountability appears to be the mechanism by which higher levels of women’s representation relate to reduced corruption perceptions.

Identifying and providing empirical evidence that electoral accountability is the key moderating factor in the relationship between women’s representation and perceived corruption is an important new finding. It answers a puzzling question – why does the relationship only exist in some countries and not others – and provides a critical caution to policy makers who

| Table 4 | How Does Separation of Powers (Accountability) Influence the Relationship Between Gender and Three Measures of Corruption? |
|------------------|------------------|------------------|
|                  | (1)              | (2)              | (3)              |
|                  | TI CPI           | ICRG             | WBGI             |
| lag TI CPI       | 0.683***         |                  |                  |
|                  | (20.76)          |                  |                  |
| lag ICRG         |                  | 0.855***         |                  |
|                  |                  | (68.99)          |                  |
| lag WBGI         |                  |                  | 0.376***         |
|                  |                  |                  | (9.31)           |
| % women in lower house | −0.0106**        | −0.00749***      | −0.00992***    |
|                  | (−2.89)          | (−4.65)          | (−4.53)         |
| Presidential system | −0.160           | −0.173***        | −0.165**        |
|                  | (−1.52)          | (−3.94)          | (−2.60)         |
| % women × presidentialism | 0.0140*          | 0.00844***       | 0.0106**        |
|                  | (2.50)           | (3.50)           | (3.24)          |
| FH freedom       | −0.178***        | −0.0437***       | −0.185***       |
|                  | (−4.37)          | (−2.75)          | (−7.07)         |
| log GDP per capita | −0.315***        | −0.0377***       | −0.262***       |
|                  | (−6.89)          | (−2.63)          | (−10.50)        |
| % protestant     | −0.00450***      | −0.000692        | −0.00308***     |
|                  | (−3.79)          | (−1.46)          | (−4.68)         |
| Trade imbalance (% of GDP) | −0.000422       | 0.000268         | 0.000177       |
|                  | (−0.58)          | (0.91)           | (0.44)          |
| Women’s economic rights | −0.0618          | 0.0465*          | −0.0436         |
|                  | (−1.31)          | (2.35)           | (−1.66)         |
| N                | 1,176            | 1,417            | 1,109           |

Note: the table reports the output of ordinary least squares regression models using three dependent variables: (1) the Transparency International Corruption Perceptions Index (TI CPI); (2) the International Country Risk Guide corruption rating (ICRG); and (3) the World Bank Governance Indicators Control of Corruption measure (WBGI). All three measures have been recoded so that higher values on each DV indicate more corruption. The data includes seventy-six democratic-leaning countries; the time dimension spans 1995–2010 for the TI CPI variable, 1996–2010 for the WBGI variable and 1991–2010 for the ICRG variable. Year and region dummies are included in the models, though not reported in this table. Estimates are based on multiple imputation into fifty datasets using chained equations. $t$ statistics in parentheses. *$p<0.05$, **$p<0.01$, ***$p<0.001$.
think bringing women into government can solve corruption problems. Increasing the proportion of women in government might reduce perceived corruption, but empirical evidence of an association between the two only exists in countries that already have high levels of electoral accountability. We find little reason to suspect that changing the proportion of women in government will change perceived corruption levels in countries with low electoral accountability.

Our empirical finding also highlights two areas for further research on the gender–corruption link. First, this article does not empirically establish the direction of causality in the relationship between women’s representation and corruption. Our theoretical argument is that having more women in legislatures and parliaments will reduce overall corruption levels because women are less likely to engage in corruption than men. Our empirics support a correlation between women in government and perceived corruption that is consistent with this theory, but they do not prove a direction of causality or establish that the relationship applies to directly observed (as opposed to perceived) corruption. Concordantly, other interpretations of our evidence are conceivable. For example, women could be more likely than men to avoid running for public office in high-corruption environments, but only when electoral accountability is high; this is a strategy of avoiding participation in corruption, and is therefore generally consistent with our theory that women avoid corruption in high-accountability contexts, but it has different causal implications. In this alternative

Note that this argument is quite different than saying that networks of corrupt officials collude to suppress female participation in government (which often involves newcomers to governance) as a part of ensuring and increasing the benefits that they derive from corrupt governance; this alternative argument has been made by others in the literature who believe that corruption suppresses women’s participation in government (Bjarneğard 2013; Goetz 2007; Grimes and Wängnerud 2012; Johnson, Einarsdóttir, and Pétursdóttir 2013; Stockemer 2011; Sundström and Wängnerud 2016). It is conceivable that corrupt officials exert more effort to suppress female participation in high-accountability environments, where women are less likely to co-operate with corruption. Yet heightened accountability to voters for corruption would presumably make it easier for women to gain office in spite of corruption, as prior research argues that many women come to participate in politics through social movements that (among other activities) work against corruption and serve as the basis for independent political networks (Rodríguez 2003). Corruption fighting can even become a signature issue for female candidates (Goetz 2002, 566).
interpretation, increasing the number of women in office may not reduce corruption; it depends on how women who would not have ordinarily run for office behave once they are elected. Another possibility is that greater female representation in government changes how observers perceive the degree of state corruption, but not the real rate of corrupt practices, when electoral accountability is high. Our current study is not designed to empirically disentangle these and other possibilities, but rather to identify that accountability is an important contextual factor that must be considered in future work that is designed to do so.

Two strategies could be particularly useful to empirically estimate the causal effect of increased women’s representation on corruption. First, survey and laboratory experiments could be designed to investigate the degree to which people select themselves out of positions that involve corruption, or choose to accept these positions but resist corruption once there. Experiments could also focus on how and why voters hold politicians accountable for
corruption, which would be particularly useful in helping us establish what causes voters to punish corruption (and what causes politicians to avoid it) at the micro level. As a bonus, experiments can allow a researcher to directly observe corrupt behaviors rather than the indirect perception of corruption. Secondly, instrumental variables techniques may allow us to directly measure the local average treatment effect of a program designed to increase female representation in government on corruption in that government.96

### Table 5

**How Does Personal Accountability Influence the Relationship Between Gender and Three Measures of Corruption?**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TI CPI</td>
<td>ICRG</td>
<td>WBGI</td>
</tr>
<tr>
<td>lag TI CPI</td>
<td>0.674***</td>
<td>(20.69)</td>
<td></td>
</tr>
<tr>
<td>lag ICRG</td>
<td>0.854***</td>
<td>(68.08)</td>
<td></td>
</tr>
<tr>
<td>lag WBGI</td>
<td></td>
<td></td>
<td>0.368***</td>
</tr>
<tr>
<td>% women in lower house</td>
<td>−0.00230</td>
<td>−0.000174</td>
<td>−0.00167</td>
</tr>
<tr>
<td></td>
<td>(−0.53)</td>
<td>(−0.09)</td>
<td>(−0.67)</td>
</tr>
<tr>
<td>Personalism</td>
<td>0.0111</td>
<td>0.0158**</td>
<td>0.0152</td>
</tr>
<tr>
<td></td>
<td>(0.73)</td>
<td>(2.68)</td>
<td>(1.74)</td>
</tr>
<tr>
<td>% women × personalism</td>
<td>−0.00229**</td>
<td>−0.00108**</td>
<td>−0.00179***</td>
</tr>
<tr>
<td></td>
<td>(−2.60)</td>
<td>(−2.96)</td>
<td>(−3.55)</td>
</tr>
<tr>
<td>FH freedom</td>
<td>−0.191***</td>
<td>−0.0435**</td>
<td>−0.190***</td>
</tr>
<tr>
<td></td>
<td>(−4.59)</td>
<td>(−2.72)</td>
<td>(−7.15)</td>
</tr>
<tr>
<td>log GDP per capita</td>
<td>−0.311***</td>
<td>−0.0359*</td>
<td>−0.258***</td>
</tr>
<tr>
<td></td>
<td>(−6.71)</td>
<td>(−2.49)</td>
<td>(−10.29)</td>
</tr>
<tr>
<td>% protestant</td>
<td>−0.00539***</td>
<td>−0.00131**</td>
<td>−0.00372***</td>
</tr>
<tr>
<td></td>
<td>(−4.38)</td>
<td>(−2.71)</td>
<td>(−5.44)</td>
</tr>
<tr>
<td>Trade imbalance (% of GDP)</td>
<td>−0.00109</td>
<td>0.000282</td>
<td>−0.000181</td>
</tr>
<tr>
<td></td>
<td>(−1.42)</td>
<td>(0.93)</td>
<td>(−0.44)</td>
</tr>
<tr>
<td>Women’s economic rights</td>
<td>−0.0403</td>
<td>0.0527**</td>
<td>−0.0309</td>
</tr>
<tr>
<td></td>
<td>(−0.85)</td>
<td>(2.63)</td>
<td>(−1.17)</td>
</tr>
<tr>
<td>N</td>
<td>1,176</td>
<td>1,417</td>
<td>1,109</td>
</tr>
</tbody>
</table>

**Note:** the table reports the output of ordinary least squares regressions using three dependent variables: (1) the Transparency International Corruption Perceptions Index (TI CPI); (2) the International Country Risk Guide corruption rating (ICRG); and (3) the World Bank Governance Indicators Control of Corruption measure (WBGI). All three measures have been recoded so that higher values on each DV indicate more corruption. The data includes seventy-six democratic-leaning countries in each model; the time dimension spans 1995–2010 for the TI CPI variable, 1996–2010 for the WBGI variable and 1991–2010 for the ICRG variable. Year and region dummies are included in the models, though not reported in this table. Estimates are based on multiple imputation into fifty datasets using chained equations. $R^2$ for the models are: (1) 0.920, (2) 0.931, (3) 0.869. $t$ statistics in parentheses. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

96 See Appendix Table S2 for an initial instrumental variable model that uses two-period lagged values of the independent variables as instruments; future research would presumably offer instruments for electoral accountability that do not depend on assumptions about dynamics.
A second important priority for future research is to distinguish between the two micro-level theoretical mechanisms – risk aversion and differential treatment – that we argued might explain why accountability moderates the relationship between women’s representation and corruption. Again, survey and laboratory experiments may help us separate these mechanisms and determine the extent to which each produces the greater responsiveness to electoral accountability that we see in observational data. Additionally, collecting panel data on individual voter attitudes and behavior toward women in government and corruption may give us traction on this question. As an added benefit, closer empirical examination of these theoretical mechanisms may also uncover other reasons why electoral accountability moderates the corruption and women’s representation relationship, some of which may be related to men’s responses to women’s increased levels of political participation.

This article takes an important first step towards understanding when women’s representation in government is associated with political corruption and why. While early work suggested that there is a clear and relatively simple link – namely, that more women in government means less corruption because women are intrinsically less corrupt – our findings support a subtler relationship that runs through electoral accountability. These findings matter for scholars hoping to better understand the causes and consequences of women’s political representation, and they have important implications for policymakers who think that increased women’s representation is a direct solution for endemic and pervasive corruption. Our findings support Goetz’s assertion that ‘To expect that women’s gender alone can act as a magic bullet to resolve a corruption problem that is much bigger than they are, that is systemic, is unrealistic to say the least. It reflects not just wishful but almost desperate thinking.’ At the same time, our findings suggest that countries considering the anti-corruption benefits of increasing gender parity in government should consider simultaneously implementing institutional reforms to catch and punish officials who are guilty of corruption. Women’s representation is much more likely to be associated with reduced corruption when accountability is high.

97 Goetz 2007, 102.
REFERENCES


Quinones, Sam. 1999. Stop! Ms, December, 24.


