Women and Corruption: What Positions Must They Hold to Make a Difference?*

Chandan Kumar Jha[1] and Sudipta Sarangi[2]

[1]Department of Economics, Louisiana State University, Baton Rouge LA 70803.
e-mail: cjha1@tigers.lsu.edu

[2]Department of Economics, Louisiana State University, Baton Rouge LA 70803; National Science Foundation, Arlington VA 22230 USA, and DIW Berlin, 10117 Berlin, Germany.
e-mail: sarangi@lsu.edu

February 2015

Abstract

This paper examines in what precise role – as bribe takers, decision makers or as policy makers do women have an impact on corruption. Since much of the corruption literature is plagued either by the lack of instruments or weak instruments, this paper makes a methodological contribution by drawing inferences based on Moreiras (2003) conditional likelihood ratio approach. We provide robust evidence that women's presence in parliament has a causal and negative impact on corruption while other measures of female participation in economic activities are shown to have no effect. Further, this negative relationship between women's presence in government and corruption is also found to hold in a regional analysis of 17 European countries alleviating concerns that the relationship is driven by unobservable country-fixed characteristics. Finally, we show that this relationship does not disappear when women gain similarity in social status.

JEL classification codes: D73; J16; K42
Keywords: Corruption; Gender; Public Policy; Politics; Women; Labor Force

*We would like to thank Louis-Philippe Beland, Areendam Chanda, Briggs Depew, Ozkan Eren, Sukriye Filiz, Kaj Gittings, Gautam Hazarika, Carter Hill, Munechika Katayama, Douglas McMillin, Naci Mocan, Luiza Pogorelova, Christian Raschke, Seunghwa Rho, and Gregory Upton for their comments. We also thank seminar participants at the 9^{th} Annual Conference on Economic Growth and Development at the Indian Statistical Institute, Louisiana State University, Southern Economics Association Meetings (2013), and "The Economics of Corruption Workshop 2013" at the University of Passau for their helpful suggestions. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

1 Introduction

Corruption remains an important issue both in poor countries and advanced economies because of its negative impact on economic outcomes such as investment, economic growth, and per capita income.[1] Little over a decade ago a gender dimension was added to this topic through two classic papers by Swamy et al. (2001) and Dollar et al. (2001), both drawing on the notion that women behave differently from men in many economic circumstances.[2] The latter study found a negative correlation between women's presence in parliament and corruption, while the former reported lower corruption to be correlated with both women's presence in the labor force as well as in parliament using cross-country analysis. Subsequently, however, a number of studies have voiced concerns that this observed negative association between gender and corruption was not causal and driven by the omission of other factors that might be correlated with women's participation and/or corruption in a country. In this paper, we address the concerns raised in this literature by *first* looking for a causal relationship between gender and corruption using instrumental variable (IV) analysis and *second* by taking a more nuanced approach to this problem by identifying different economic roles women can take vis-a-vis corruption and investigating the impact of each on corruption.

We start by pointing out that the term "labor force" which has been found to be negatively correlated with corruption in earlier studies is a very broad measure and does not make clear how women affect corruption. For example, women may affect corruption if they are less corrupt and accept fewer bribes than men. Alternatively, women can affect corruption when they are in positions of power such as heads of their organizations, by designing and implementing stringent anti-corruption laws within their organizations or making the

[1] For instance, a higher level of corruption is associated with lower levels of GDP per capita (World Bank, 2001); lower rates of investment and economic growth (Mauro, 1995); high inequality and poverty (Gupta et al., 2002).

[2] A number of studies support this hypothesis (see Eckel and Grossman (1998) and references therein).

existing laws more strictly enforceable. Since female participation in the labor force consists of women in both the roles – the bribe-taking role as well as the decision-making role, it is important to distinguish which of these roles (or a combination of the two) is associated with lower corruption. In order to capture these roles, we introduce two additional measures of female participation in economic activities: (i) the share of women in clerical positions, and (ii) the share of women as legislators and managers. While the first measure indicates the presence of women in potential bribe-taking positions, the second measure indicates their presence in decision-making positions. Finally, in keeping with the earlier literature, we narrow down this somewhat broader measure of women's presence in positions of power to only their presence in policy-making positions, and examine the relationship between the share of women in parliament and corruption. The investigation of the relationship between these four different measures of female involvement and corruption enables us to identify the exact role in which women are able to effectively reduce corruption.

A possible reason behind the lack of studies identifying a causal relationship between gender and corruption could be the fact that a panel study on corruption is not possible due to the invariability of corruption indices over time, and finding instruments that are both valid and strong is a daunting task as well. Moreover, the determinants of women's presence in different occupations are likely to be different, and hence, an instrument that works well for women's presence in one occupation need not work well for their presence in other occupations giving rise to the need of finding more than one instrument. In an attempt to establish causality, we take up this challenge by looking at some of the recent studies that discover historical and linguistic determinants of women's presence in different occupations. We identify such potential instruments that have predictive powers for women's presence in different positions, yet there is little reason to expect a direct effect of these variables on corruption. We experiment with multiple instruments to explore the causal relationship between the share of women in parliament and corruption. While using more than one

instrument for one endogenous variable allows us to check for the validity of our instruments conditional on at least one of our instruments being valid, our instruments tend to be weak in some specifications which may lead to invalid inferences. We overcome this possibility by using the conditional likelihood ratio (CLR) approach developed by Moreira (2003) for hypothesis testing that provides for the robust inferences in the presence of weak instruments.

Another concern may be that our findings of women's presence in politics and its effect on corruption is driven by country fixed effects. To address this issue, we use data from 155 regions from 17 European countries and find that there is a negative association between women's share in the local government and corruption.

The next question that this paper addresses is regarding the *persistence* of the observed association between women's presence in different positions and corruption. It has been argued that women are not actually less corrupt, and the observed association between different measures of female participation and corruption is actually driven by gender differences in the social status limiting women's access to corruption.[3] To the best of our knowledge, this paper is the first to investigate this hypothesis that the rate of corruption among women will converge to that among men as gender-gap in social and economic status narrows down.

Women may have better access to corrupt practices and activities over the years as they get similar in status to men and may also have greater exposure to bribe-taking activities. As a result, it is possible that the negative relationship observed by previous studies may no longer be valid. Hence, quite aside from the fact that we have addressed the concerns of the previous studies and provided new insights, the present study is also a timely re-investigation of this topic to the extent allowed by the availability of data and empirical limitations.

Our main results are as follows. The role in which women have an impact on corruption

[3]Swamy et al. (2001) clarify *"... we do* not *claim to have discovered some essential, permanent or biologically determined differences between men and women. Indeed, the gender differences we observe may be attributable to socialization, or to differences in networks of corruption, or in knowledge of how to engage in corrupt practices, or to other factors."*

is through their presence in politics. Using an IV approach we show that this relationship is robust and causal. Moreover, our findings hold at both, the national and sub-national, levels. We also show that the observed negative relationship between female participation and corruption cannot entirely be explained by gender differences in social status.

The rest of the paper is organized as follows. In the next section, we discuss our sources of data and specify empirical strategy as well as establish the validity of our instruments. Section 3 reports cross-country OLS and IV results, and section 4 presents sub-national evidence. We check whether there is an evidence of "corruption convergence in gender" in section 5 and discuss the implications of our findings in section 6.

2 Data and Empirical Strategy

2.1 Data

The primary measure of corruption used in this paper is the Control of Corruption Index (CCI) published by the World Bank. The CCI lies in the range of -2.5 (most corrupt) to 2.5 (least corrupt). It is a continuous variable and takes values up to 2 decimal points. We use negative of the CCI in all our specifications such that a higher number indicates more severe corruption. The CCI has been constructed in a way that mean of the index is 0 and the standard deviation is equal to 1. The purpose of CCI, as described by Kaufmann et al. (2010), is – *"capturing perceptions of 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"*.[4] The index combines information available from a combination of surveys of firms and households, subjective opinion of business analysts, non-governmental organizations (NGOs) and public sector agencies assigning larger weights

[4]For details of data sources, questions and methodology used in computing the CCI, see Kaufmann et al. (2010) and also visit http://info.worldbank.org/governance/wgi/resources.htm.

to sources that have similar findings.

The data for the share of women in the labor force (WP) comes from the *International Labor Organization (ILO)*. The United Nations Statistics Division (UNSD) provides data for the share of women in clerical positions and the share of women in decision-making positions.[5] Data for the percentage of women in parliament is compiled by the *Inter-Parliamentary Union (IPU)* and has been taken from the World Bank. It provides the percentage of parliamentary seats held by women in a single or lower chamber. All measures of female participation used in this paper are the *percentage* of women in the respective category.

We use Gross National Income per capita (formerly Gross National Product (GNP) per capita) in US dollars obtained from the World Bank as a measure of income and refer to it as $GNPPC$. The Association of Religion Data Archive (ARDA) provides data on proportions of Christians ($Christian$) and Muslims ($Muslim$) in the total population.[6] Note that the latest year for which data is available for these variables is 2005. Data for the colonial history of countries has been taken from Treisman (2007). Freedom House awards a score of 1 through 7 for political rights – a score of 1 indicates that the citizens enjoy a wide range of political rights while a rating of 7 implies few or no political rights.[7] Summary statistics are presented in Table 1.

[5]The share of women in decision-making positions, as per the International Standard Classification of Occupations 1988 (ISCO-88), includes: 11. legislators and senior officials; 12. corporate managers; and 13. general managers. According to ISCO-88, the clerical positions include: 41. Office clerks; 42. Customer service clerks. In the sample, some countries also report employment statistics according to the earlier edition, ISCO-1968. Visit http://laborsta.ilo.org/applv8/data/isco68e.html for the details on ISCO-1968 and http://laborsta.ilo.org/applv8/data/isco88e.html for the details on ISCO-1988 classification of other job-categories included under these broader groups.

[6]The data was downloaded from http://thearda.com. The principal investigators in this data collection are Jaime Harris, Robert R. Martin, Sarah Montminy, and Roger Finke of the ARDA.

[7]Visit http://www.freedomhouse.org/report/freedom-world-2012/methodology for the details on how political rights index is computed.

2.2 Empirical Specification

We estimate the following equation using OLS which is our baseline specification

$$Corruption_i = \alpha + \beta WP_i + \gamma_1 GNPPC_i + \gamma_2 Political\,Rights_i + \gamma_3 Christian_i + \gamma_4 Muslim_i + \gamma_5 Past\,UK\,Col_i + \gamma_6 Never\,Colonized_i + \varepsilon_i \quad (1)$$

where $Corruption_i$ is the index of corruption in country i, and WP_i stands for the share of women in different positions in country i, depending on the specification. The dummy variable $Past\,UK\,Col_i$ takes a value of 1 if country i is a former British Colony, and 0 otherwise. $Never\,Colonized_i$ takes a value of 1 if country i was never colonized, and 0 if country i has a colonial past. As discussed earlier, we use the negative of the CCI in all our regressions, and therefore, the coefficient β is expected to be negative.

GNP per capita has been added as a control variable in all the specifications because it has been argued that countries with higher incomes may be able to constrain corruption more effectively than poor countries (Treisman, 2000). Countries with long democratic histories are less corrupt (Treisman, 2000). Therefore, in countries with strong political and democratic institutions, corruption is likely to be lower. So we include 'political rights' published by Freedom House as an additional regressor in the model. Cultural factors (Alatas et al., 2009) and social norms (Fisman and Miguel, 2007) have been found to have an impact on corruption. To capture these aspects, following the literature, we include proportions of Christians and Muslims in the total population as additional regressors (Swamy et al., 2001; Treisman, 2000).[8] Recent studies find that the colonial past (Swamy et al., 2001) and legal

[8]Treisman (2000) argues that the objections raised against corrupt activities by office holders may be less in countries with a large proportion of population belonging to hierarchical religion compared to the countries where the population belonging to more egalitarian or socialistic religions such as Protestantism is higher. La Porta et al. (1997) classify Roman Catholic, Eastern Orthodox, and Muslim religions as hierarchical religions, and show that corruption is positively associated with hierarchical religions. They attribute this association to the lack of trust caused by the hierarchical structure. Due to the unavailability of data, we cannot control for the Catholic proportion. However, our objective is to capture the cultural aspects, and

origin (Treisman, 2000) of a country may affect corruption in a country via its impact on economic and political institutions. Several studies find that former British colonies have better property rights, economic and political institutions, and more developed financial markets compared to former French, Portuguese, and Spanish colonies (La Porta et al., 1998; North et al., 2000). It is argued that a colonized country inherits the institutional set-up from its colonizer which is likely to persist.[9] Following this, we include a 'Former British Colony' dummy, and 'never colonized' dummy in the model.

2.3 Instruments

Recognizing that the women participation variables are potentially endogenous, we use an IV approach to establish causality. In our quest for valid instruments, we appeal to the recent literature that reports a robust link between the grammatical structure of a language and various economic outcomes. For instance, Chen (2013) finds that languages that grammatically associates the present with the future are correlated with the speaker's health behavior and financial decisions, both within and across countries. We instrument women's presence in the labor force by a dummy variable that assumes a value 1 for the countries having a dominant language with two genders (masculine vs. feminine) and value 0 if the country's dominant language has either 0, 3 or more genders. Gender distinction is more pronounced, and hence, female participation in economic activities is lower in countries in which the dominant language has two genders as opposed to countries with dominant language having no gender or three or more genders (Gay et al., 2014). The gender marking of a language is, therefore, a valid instrument for the share of women in the labor force as

we find that our results are robust to the inclusion of the proportion of population belonging to other major religious faiths including Hinduism, Buddhism, Confucianism, and Atheism.

[9] Acemoglu and Robinson (2001) list three reasons why the institutions are unlikely to change. First, changing existing institutions are costly. Second, Europeans often delegated power to a small group of elite who may have an incentive to protect the existing extractive institutions set up by the former. Third, the irreversible investments made by the agents to complement the existing institutions will make them favor these existing institutions.

there is no direct effect of this variable on corruption. Moreover, the gender marking of a country's dominant language has strong predictive power for women's presence in the labor force even after controlling for income, institutions and cultural variables making it a strong instrument for the share of women in the labor force.

We present our IV results of the impact of women's presence in parliament on corruption by experimenting with three potential instruments.[10] Our first instrument is women's exposure to democratic rights as measured by the year when they were granted suffrage. The rationale for this instrument is that an early exposure of women to voting rights will affect women's presence in parliament today, while there is little reason to believe that an early voting rights to women will have a direct effect on corruption. We recognize that this instrument may not be perfect as in certain scenarios, the exclusion restriction may be violated. For instance, it is possible that institutions may affect voting rights which also have an impact on corruption. Furthermore, as aptly noted by Murray (2006a), in IV estimation, the omitted variable bias arises in a new form – IV estimates are biased if an omitted variable that belongs to the model is either correlated with other explanatory variables or with the instruments. To alleviate these concerns, we control for the likely sources of correlation between our instruments and the error term by including cultural, historical and contemporaneous controls, as well as colonial and continent dummies.

Moreover, we employ a second instrument that allows us to check whether or not our instruments are valid, conditional on either one of the instruments being valid. Our second

[10]However, we only report results with two instruments. We check the consistency of our estimates by using a third instrument which is a dummy variable that takes a value of 1 if the dominant language of the country is sex-based, and value 0 if the country's dominant language has either non-sex-based gender system or no gender system. There are examples of languages that have non-sex-based gender system. Such systems might be based on the distinction between human and non-human, or animate versus non-animate (Gay et al., 2014). The rationale for our third instrument stems from the fact that the gender distinction tends to be more pronounced among the speakers of such languages, and women's participation in economic positions is lower in countries where such languages are spoken dominantly (Gay et al., 2014). On the other hand, we do not expect that whether or not a language has sex-based-gender system should directly affect corruption in the country, except, of course, via affecting women's presence in parliament. The results are in agreement with the findings with the above two instruments and are obtainable from authors on request.

instrument, years since transition to agriculture, comes from a recent study (Hansen et al., 2012) that finds that the societies that have long agricultural histories have more unequal gender roles and lower participation of women in economic and political arenas including the labor force and parliament. This is a valid instrument as we find that years since transition to agriculture is indeed associated with lower participation of women in parliament, and at the same time, there is no reason to expect that it can affect corruption directly. Mauritius is the last country that adopted agriculture in our sample 375 years ago, while there are countries that adopted agriculture as early as 10,500 years ago.

Though our instrument for the share of women in the labor force is strong, our instruments for the share of women in parliament tend to be weak in some specifications which may lead to invalid inferences. To counter this possibility, we use the Conditional Likelihood Ratio (CLR) approach proposed by Moreira (2003) for hypothesis testing. Furthermore, while under homoskedasticity, the CLR test is the most powerful test for hypothesis testing in the presence of one endogenous variable and weak instruments, this result remains to be established for other IV-type estimators (Murray, 2006a; Finlay and Magnusson, 2009). Hence, we also report $p-$values for alternative approaches that provide robust inferences in the presence of weak instruments such as LM-J (a combination of Kleibergen-Moreira Lagrange Multiplier (LM) and the overidentification (J)-tests) (Kleibergen, 2002), and Anderson-Rubin (AR) tests (Anderson and Rubin, 1949) against the null that the coefficient of the instrumented variable, the share of women in parliament, is zero. In case of an over-identified equation, all these three statistics test both the structural parameters and the overidentification restrictions simultaneously by combining the LM statistic and J statistic, and provide inferences that are robust to the presence of weak instruments.

3 Cross-country results

3.1 Cross-country OLS evidence

A. Women in the labor force and corruption

First, we investigate the relationship between the share of women in the labor force and corruption. The first column of Table 2 presents the result of the baseline specification with the variable of interest being the share of women in the labor force. The coefficient on the share of women in the labor force is negative and significant at the 5% level. However, when we include continent dummies in column 2, coefficient of the share of women in the labor force is no longer significant though it has expected sign.

B. Women in potential bribe-taking positions and corruption

As shown in Table 2 (column 3), we do not find any significant association between the share of women in clerical positions and corruption which suggests that the bribe-taking role of women is not significant in determining the relationship between female participation in the labor force and corruption.

C. Women in decision-making positions and corruption

Next, we investigate whether decision-making ability allows women to impact corruption. This position captures both the bribe-giving and demanding role: While women in the positions of senior managers and officials are likely to be bribe-givers; women as legislators and senior government officials are likely to be bribe-takers. We, however, find no association between the share of women in decision-making positions and corruption as the coefficient on the former is not significant (column 4 of Table 2).[11]

D. Women in parliament and corruption

Finally, we investigate if women can have an impact on corruption by being in the

[11] We also do not find any significant impact of the share of women in clerical positions and share of women in decision-making positions on corruption when the continent dummies are included in the model. These results are omitted for the sake of brevity and are available from the authors on request.

role of policy makers. Consistent with the findings of the previous studies, we present evidence of a significant and negative association between the share of women in parliament and corruption. The coefficient of the share of women in parliament is found to be highly significant with the expected sign (column 5). Moreover, this relationship is robust to the inclusion of the continent dummies in column 6. Finally, column 7 controls for both the variables – women's share in the labor force and their presence in parliament. As we can see, the coefficient of the share of women in the labor force is very small and insignificant. On the other hand, the coefficient of women's participation in parliament remains significant.[12]

E. Inclusion of additional variables

Next, we control for a number of variables in order to minimize the possibility of omitted variable bias as well as to address the concerns of some of the previous studies that hypothesize that the relationship between female participation variables and corruption is spurious and is driven by the omission of relevant variables. These results are presented in Table 3.

Liberal democracy: Sung (2003) argues that, in liberal democracies, women's participation is higher and corruption is lower, and it is the omission of the *liberal democratic institutions* variable that may be responsible for the relationship between female participation and corruption. We address this concern by replacing the *political rights* variable with the *civil liberties index* published by Freedom House. It takes a value from 1 (high civil liberties) through 7 (low civil liberties) and is a broader measure of liberal democracy than political rights.[13] The index takes into account, among other things, the personal and social freedom of women including their choice of marriage partners and say in the family size. The coefficient of both the female participation variables remains significant when this variable

[12] Notice that the regressions with women in decision-making positions and clerical positions as variables of interest have a considerably smaller sample size of 91 countries because of the data unavailability. Restricting the regressions to only those countries for which data is available for all the women's participation variables, we find similar results. These results are presented in the Appendix.

[13] For details on the differences between the two indices and how they are computed, visit Freedom House website: http://www.freedomhouse.org/report/freedom-world-2012/methodology.

is controlled for in columns 1 and 7.

Power structure and corruption: Different cultures have varied levels of tolerance for an unequal distribution of power. Hofstede's Power Distance Index (PDI) measures this tolerance providing a score in the range of 0 to 120, with the higher value indicating tolerance for a hierarchical order while a lower value implies that people strive to equalize the distribution of power.[14] We control for the PDI as an alternative measure of cultural differences among countries, and find that while women's share in parliament remains highly significant (column 8), women's share in the labor force is no longer significant (column 2).

Gender-biased institutions and corruption: It has been hypothesized that social institutions that discourage female participation in political and economic spheres are also more corrupt (Branisa et al., 2013). Hence, to address the concerns that our results may have been driven by the omission of gender-biased institutions, we control for the Social Institutions and Gender Index (SIGI). SIGI is a measure of gender inequality attributed to institutions and was first launched in 2009 by the Organisation for Economic Co-operation and Development (OECD). It captures *"discriminatory social institutions, such as early marriage, discriminatory inheritance practices, violence against women, son preference, restricted access to public space and restricted access to land and credit."* The index takes a value from 0 to 1, with 1 representing high inequality.[15] The inclusion of SIGI causes the share of women in the labor force to be close to zero and insignificant in column 3. The coefficient of the share of women in parliament, however, remains sizable and significant in column 9.

Schooling, ethnic division, and corruption: Columns 4 and 10 control for two additional covariates – 'proportion in largest ethnic group' and 'average years of schooling'.[16]

[14] Visit `http://geert-hofstede.com/dimensions.html` for details.

[15] For details of the construction of SIGI, see Branisa et al. (2009) and Branisa et al. (2013).

[16] The proportion of the population belonging to the largest ethnic groups (*Ethnic*) is taken from Sullivan (1991). While the World Bank publishes data on schooling, the coverage of countries in Barro-Lee (Barro and Lee, 2013) data set is broader making it our preferred source for schooling data. We use average years of

Corruption may be higher in countries that are more ethnically divided, and lower in countries with higher human capital where people are aware of their legal and constitutional rights. The negative relationship between women's participation variables and corruption is significant in both the columns. However, once we add continent dummies along with these variables, the share of women in the labor force loses significance in column 5.

Openness to trade: It has been found that countries that are more open and have lower barriers to international trade are less corrupt (Ades and Di Tella, 1999; Treisman, 2000). Hence, we include the share of imports of goods and services in GDP as a measure of openness to trade. We take this data from the World Bank. We find that the share of women in parliament (column 12) is significant; while women's share in the labor force is not significant at conventional levels (column 6).[17]

3.2 Establishing Causality: IV Analysis

In the last section, we find that the relationship between female participation in the labor force and corruption though negative is not robust. On the other hand, the relationship between the share of women in parliament and corruption was found to be negative and significant across different specifications. Although we control for a number of variables in the previous section, the possibility of endogeneity cannot be entirely ruled out in a cross-country OLS specification. We now use instrumental variable analysis in order to determine if gender representation has a causal impact on corruption.[18]

schooling (*Education*) data for the year 2010 as the Barro-Lee educational attainment data is available only for 5-year intervals. Our results are, however, robust when we use the proportion of the population with the secondary education or tertiary education, instead of the years of the average years of schooling attained by the population.

[17]Notice that we have a smaller sample when the variable of interest is the share of women in parliament. In order to rule out the concern that sample selection is responsible for the differences in the significance of the two variables, we re-run the regression specifications in columns 1-6 restricting the sample to the countries that are included in columns 7-12. Results are similar and presented in the Appendix.

[18]The IV estimates of the other two variables – the share of women in clerical positions, and the share of women in decision-making positions – using voting rights and transition since agriculture as instruments indicate that women's presence in these positions do not have an impact on corruption. These results have

A. Women in the labor force and corruption

Our instrument for the share of women in the labor force in a country is *the number of genders present in its dominant language.*[19] The lower panel of Table 4 reports the first stage results while second stage results are reported in the upper panel. In the first stage, the language variable is a significant predictor of women's share in the labor force across all specifications. The coefficient of the share of women in the labor force in the second stage has the expected negative sign, but it is not significant in column 1. It is also not found to be significant when we control for cultural variables, colonial dummies and a set of contemporaneous variables. In all these specifications, while the coefficient of the share of women in the labor force has the expected sign, in none of these columns, it is significant at conventional levels. Also, note that the instrument is strong in each specification as indicated by $F-$statistic in the columns of the bottom row of panel 1. Overall, the IV results convey the same story as OLS – while a larger share of women in the labor force seems to be associated with lower corruption, this relationship is not robust.

B. Women in parliament and corruption

Although we control for a number of variables to minimize the possibility of omitted variable bias, and find that the relationship between the share of women in parliament and corruption is robust, the concerns for reverse causality remains. It is possible that not only does women's presence in parliament affect corruption, but in corrupt countries women are also discouraged from participating in politics. If this is true, our OLS estimates will be biased. To address this issue, we use instruments for women's presence in parliament.

First, we present IV estimates by instrumenting women's participation in parliament

been omitted for the sake of brevity and are available from authors on request.

[19]There are two dominant theories that trace the origin and evolution of language. The first theory credits the evolution of the grammatical structure to biological adaptation. The second theory considers languages as institutions that are shaped by a society's cultural heritage and links the emergence of grammatical structure to the cultural transmission of language to hundreds of generations of learners (for a review see Christiansen and Kirby, 2003).

with "the year women were granted voting rights". An initiative to include women in the political process should be positively correlated with their presence in politics and in national parliaments. Figure 1 shows a scatter plot of the association between the year women were granted suffrage and their presence in parliament. New Zealand is the first country to allow women to vote in 1893, and the State of Kuwait was the last country (in our sample) to grant voting rights to women in 2005.[20] However, there still might be some concerns about omitted variables in the IV regression. Hence, we discuss a number of potential factors from economics and political science studies that have explored the reasons leading to women being granted voting rights in different countries, and include these factors as explanatory variables in the empirical specification.

The timing of suffrage for women was affected by different factors throughout the world including threat of revolution, sociocultural factors, and the quality of institutions (Acemoglu and Robinson, 2000; Reynolds, 1999; also see Rule and Zimmerman, 1994). For example, Acemoglu and Robinson (2000) argue that the threat of revolution was the main contributing factor to political reforms and extension of voting rights to the poorer sections of the society in the West. In several countries like Germany and Sweden, this included women from the beginning. Another important determinant of enfranchisement of women is cultural factors. To the extent that cultural factors have an impact on corruption, our instrument will not satisfy the exclusion restriction if cultural factors are omitted from the model. However, since the proportions of Christians and Muslims in the total population are arguably exogenous to corruption (Treisman, 2000), we are able to control for this variable. Finally, it may also be institutions which may affect the timing of women being granted voting rights. Since institutions have also been found to impact corruption, exclusion of an institutional variable

[20]The State of Kuwait granted voting rights to women on the condition that they observe Islamic laws. See the link for additional details, http://www.cnn.com/2005/WORLD/meast/05/16/kuwait.women/. This underscores the impact of culture on voting rights for women which we address by controlling for the proportions of Christians and Muslims in the total population. We come to this issue later in the main text.

may lead to biased estimates. We address this concern by controlling for the colonial status of a country which is also exogenous to the level of corruption in a country. As argued above, colonial status of a country has important bearings on its institutions (see Acemoglu and Robinson, 2001).

In several countries in our sample, women were granted voting rights at the time the country got independence. So, in these countries there was no independent initiative to grant voting rights to women. However, our instrument is still valid due to the fact that as long as women got voting rights earlier (because of early freedom of a country from its colonizer) women's presence in the parliament will be higher. The concern, however, is that in these countries women's presence in parliament may also be capturing the impact of a change in democratic status of these countries since the independence year and women's participation in parliament are correlated. Since the year of independence of a country is exogenous, we control for the year of independence to rule out the possibility of any bias arising from this. Data for the year of independence has been taken from Acemoglu et al. who set any year before 1800 as 1800 (see Acemoglu et al., 2008 for details).

In panel 2 of Table 5, we present the first stage regression results. Consistent with our expectations, we find that women's presence in parliament is higher in countries where women were granted voting rights earlier. The coefficient of the "year women granted suffrage" in column 1 suggests that the presence of women in parliament in a country is about 1 percentage point less compared to a country that granted voting rights to women about 6 years earlier. Voting rights remains a significant predictor of women's share in parliament when we control for additional variables in columns 2, 3 and 4.

Panel 1 of Table 5 presents IV estimates obtained from the two-stage least squares (2SLS) regressions.[21] The specification presented in column 1 controls for cultural variables, colo-

[21]The point estimates obtained from using limited information maximum likelihood (LIML) are very similar to the 2SLS estimates and are available from authors on request.

nial dummies and year of independence for the reasons discussed above. We also include continent dummies in order to rule out the possibility of any bias arising from the omission of continent fixed factors. The IV coefficient of the share of women in parliament is negative and highly significant in column 1. Column 2 controls for the contemporaneous variables – log (GNPPC), political rights, average years of schooling, openness to trade, and proportion in largest ethnic group. The coefficient of women in parliament from the second stage results indicates a strong and significant negative impact of this variable on corruption.

Columns 3 and 4 report the results with the instruments being the year women were granted voting rights and years since transition to agriculture. We control for a number of historical variables that may potentially be correlated with our instrument, transition to agriculture. In particular, we control for the use of plow, suitability of agriculture and the presence of tropical climates. These variables are likely to have played a role in the adoption of agriculture (see Alesina et al., 2013 and Hansen et al., 2012). Moreover, we also control for the year of independence, colonial dummies, and continent dummies. Both, the voting rights variable as well as years since transition to agriculture, are significant predictors of women's presence in parliament, though, they are weak predictors as indicated by an $F-$ statistic of 4.32. Finally, in column 4, in addition to the historical controls, we also control for the cultural variables and a set of contemporaneous variables. The voting rights variable remains a significant predictor of women's presence in parliament, while the transition to agriculture is no longer significant, though, it has expected sign. In both the columns, the coefficient of women's presence in parliament is statistically highly significant. A concern, however, remains – it is possible that the null hypotheses in both these columns may have been falsely rejected because of the presence of weak instruments.

To overcome this concern, we report the CLR, AR and LM-J statistics, and 95% confidence sets based on the CLR approach using the methods discussed in Finlay and Magnusson

(2009) that allow for robust standard errors.[22] In all our specifications, all the three statistics (the CLR and the AR in columns 1-4, and the LM-J in columns 3-4) reject the null hypothesis that the coefficient of women's presence in parliament is zero. Notice that though the Sargan-Hansen over-identification test fails to reject the null that our instruments are valid in both the columns (3 and 4), the J-statistic rejects the null that the instruments are valid in column 3, which could be because contemporaneous controls are not included. The $LM - J$ statistic, which combines the LM and J statistics, however, still rejects the null that the coefficient of women's share in parliament is zero at less than 5% level of significance. In column 4, however, when contemporaneous controls are added to the model, the J-statistic indicates that the instruments are valid with a p-value of 0.23.[23]

The IV findings can be summarized as follows. First, the coefficient of women's presence in parliament is always highly significant in all the specifications, irrespective of the instruments used. Second, across all the specifications, the CLR statistic rejects the null that the coefficient of the share of women in parliament is zero. Third, the IV coefficients of women's share in parliament are always greater than the OLS coefficient reported in the previous section indicating that the OLS estimates of the effect of women's presence in parliament on corruption may be the lower bound.[24]

[22]The Finlay and Magnusson (2009) Stata routine does not report the CLR statistic when there is only one instrumental variable. Since in case of an exactly identified equation, the CLR and AR statistics are equivalent (Murray, 2006b), in the first two columns, we report the 95% confidence sets and CLR $p-$values based on the AR statistics as reported by the *"rivtest"* (Finlay and Magnusson (2009)) command in Stata.

[23]While experimenting with the other instrument in conjunction with these two instruments, the J-statistic indicates that our instruments are valid with a $p-$value exceeding 0.53 in each case. Moreover, the coefficient of the share of women in parliament is always found to be negative and highly significant according to all the three (CLR, AR and LM-J) weak instrument robust statistics.

[24]We also experimented with both the endogenous variables in the same specification and a combination of instruments. In such cases, though the results are similar, the instruments become weak. Unfortunately, there are no weak-instrument robust tests available when there are more than one endogenous regressors, hence, we do not report these results (see page 28 of Murray, 2006b).

3.3 Robustness Check

We limit our robustness checks to the two main variables of interest that have been found to be significantly associated with corruption in the previous section: the share of women in the labor force and the share of women in parliament.

Sensitivity to an alternative measure of corruption

We check the sensitivity of our results with the use of an alternative measure of corruption – Corruption Perception Index (CPI).[25] The CPI is published by Transparency International which defines it as "the misuse of public power for private benefit." It takes values in the range of 0 to 10, with a higher value indicating a lower level of corruption. The index is created by using data from different surveys conducted by a number of independent sources making it reliable and one of the most widely used indices in empirical corruption literature. In order to ensure reliability and robustness of the index, CPI ranks only those countries which are covered by a minimum of three different sources. We use negative of the CPI so that a higher number indicates higher corruption. The results are presented in Table 7. The results show that female participation in parliament is again highly significant and has a negative impact on corruption regardless of the regression type and inclusion of the continent dummies. Share of women in the labor force, on the other hand, has expected negative sign, but this relationship is not significant at conventional levels.

Further robustness check

In addition, we also perform median and robust regressions, as well as check the non-linearity of the relationship by using ordered-probit regressions (see the Appendix for details). These results are similar and show that women's presence in parliament is negatively associated with corruption. We do not find a robust relationship between women's share in the labor force and corruption.

[25]The CPI method was developed by Johann Lambsdorff of the University of Passau. For the detailed methodology of computation of 2009 corruption perception index, visit: http://archive.transparency.org/policy_research/surveys_indices/cpi/2009/methodology.

4 Sub-national evidence

Our IV results show a robust and negative relationship between women's share in parliament and corruption. However, there may still be a concern that there could be some country-specific fixed factors that are associated with both a higher participation of women in government as well as lower corruption and it is these factors that may be driving the relationship between the two. Although this seems to be less likely, as a means to bolster our results further, we provide sub-national evidence from 17 European countries covering 155 regions.

The advantage of this exercise is manifold. First, the European countries constitute a relatively homogeneous sample compared to the world sample. Second, this analysis allows us to control for the country fixed effects ruling out the possibility that our results are driven by the omission of country fixed factors. Finally, the measure of corruption used in this exercise is *actual corruption* instead of perceived corruption and hence is free from the criticisms that perception indices are subjected to. The corruption or bribery data comes from the European Quality of Government Index (EQI).[26] The first round of the EQI survey was conducted in 2010 covering 172 regions of 28 member states of European Union and surveyed 34,000 citizen respondents in December 2009 (Charron et al., 2011). Our corruption measure is the response to the following question of the 2010 EQI survey:

In the past 12 months have you or anyone living in your household paid a bribe in any form to:

(a) Education services? (b) Health or medical services?

(c) Police? (d) Any other government-run agency?

The response "yes" is coded as 1 and "no" is coded as 2 and the percentage of population having paid a bribe is calculated. This number is then standardized with mean 0 and

[26]The EQI data (2013) has been used by Agerberg (2014) to explore the gender differences.

standard deviation 1 such that a higher value indicates lower bribe experience. We use the negative of the index such that a higher value implies a greater bribery experience. Sundström (2013) has collected the data on women's political representation at the lowest administrative tier. The data for women's share in the local level government in each region is collected in a way to maximize the comparability across countries. The unit of analysis in this section is NUTS 2 (the Nomenclature of Territorial Units for Statistics) – a standard developed by the European Union for subdivisions of countries for statistical purposes. The data for the regional controls by NUTS 2 region (income per capita, the share of women in the labor force and proportion of the population aged 25-64 with tertiary education attainment) come from Eurostat.

Table 6 provides the results. In the first 3 columns, instead of controlling for country fixed effects we control for country-level variables – income and political rights. In the first two columns, a higher share of women in the labor force and higher share of women in local government is found to be negatively associated with bribe experiences of citizens. Column 3 controls for the share of women in parliament and the national share of women in the labor force. While women's participation in politics at both levels, local and national, is associated with lower corruption, only women's share in the national labor force is found to have a negative association with corruption. The results, however, may have been driven by the omission of the country fixed factors. In order to rule out this possibility, we control for country fixed effects in the next 3 columns. While the share of women in local government is still negatively associated with corruption, the association between the share of women in the labor force and corruption disappears once country fixed factors are controlled for. These results suggest that the effect of women's presence in politics on corruption is not driven by country-specific fixed factors and that the relationship between the women's political participation and corruption is causal.

5 Will the relationship between gender and corruption disappear as women become more equal?

In this section, we explore the notion of "corruption convergence in gender", that is, whether over time women's participation in corrupt activities will be no different from that of men. This seems to be the only logical possibility since none of the previous studies claim that women are inherently less corrupt or that the observed gender differences in attitude towards corruption are permanent or biological. Ideally, this issue should be investigated using a panel specification, but as discussed earlier, it is not feasible when using corruption indices. Hence, we devise an alternative approach to deal with this problem.

As discussed in the introduction, several studies have suggested that the observed gender differences in attitude towards corruption could be the result of gender differences in social status or women's lack of knowledge regarding how to engage in corrupt activities or their ability to make decisions relating to corrupt activities or even for that matter being given access to corrupt activities. For instance, Goetz (2007) gives the example of Peru where the salary of traffic cops is insufficient to maintain a family. She observes that while men as traffic officers accept bribes, this is not yet the case for women. However, doubting that this will persist in the long run, she writes *"...women will not passively conform to the idealized notions of their finer moral nature when they have families to feed and if there is money to be made from public office."*

Consequently, one possibility is that there would be no difference in corrupt practices between men and women if women held similar status as men. We capture this issue by using the Gender Inequality Index (GII) computed by the United Nations Development Programme (UNDP) as a measure of the status difference between genders. The GII measures the disadvantages of being a woman in three dimensions: reproductive health, empowerment and labor market outcomes. The index can take a value in the range of 0 (implying more

equal gender-status) to 1 (when the gender-gap is large).[27]

Now, if the hypothesis of "corruption convergence in gender" were true implying that women will become as corrupt as men with equality of status, then the interaction term between the GII and female participation variables must assume a negative coefficient in Table 8 – as equality of status increases (or GII decreases), corruption must go up.

However, we find exactly the opposite result – the coefficient of the interaction term is positive and significant in all the four columns. In other words, the relationship between gender and corruption is not driven by gender differences in social status. This finding is in contrast with the hypothesis of "corruption convergence in gender".[28] However, note that we cannot rule out the possibility of endogeneity in this section, and these findings are just suggestive evidence – an acceptance or rejection of this hypothesis calls for the need of further research.

6 Concluding remarks

In investigating the role in which women can affect corruption, this paper finds that women's presence in the labor force, in clerical positions, and in decision-making (senior-level) positions is not significantly associated with corruption in a country. We show that women are able to have a systematic negative impact on corruption only if they are represented in parliaments and possibly through policy making. We use instruments to establish causality, and draw inferences based on the conditional likelihood ratio approach proposed by Moreira (2003), Anderson-Rubin Statistic (Anderson and Rubin, 1949), and $LM - J$ statistic

[27] Annual data for GII, however, is not available, hence, we use the 2008 index in all our regressions. For more details on the GII, visit http://hdr.undp.org/en/statistics/gii/

[28] This finding is also in line with the hypothesis that women's presence in parliament affects corruption via policy effects. In countries where women actually have a say in policy making rather than just being a member of a parliament with no actual power, women are more likely to be successful in lobbying for the policies they deem necessary, and hence, a bigger effect on corruption. Nevertheless, we cannot rule out other explanations, and therefore, this finding must be interpreted with caution and requires further investigation.

(Kleibergen, 2002). Our results are also shown to hold in a regional analysis of European countries where after controlling for the country fixed effects, we find that women's share in the local government is associated with lower bribery.

Our analysis also raises one question for future research – how do women reduce corruption by being in politics? One possible answer could be that they favor policies that are different from those favored by men. The recent literature has extensively explored the policy implications of gender representation in government. Women in local government in India have been reported to allocate a greater budget to public goods more closely associated with women's concerns (Chattopadhyay and Duflo, 2004), to the provision of basic infrastructural needs, and to be more concerned about whether the subsidies were provided to the targeted group without corruption (Kudva, 2003). Moreover, using close elections between men and women, women legislators have been found to invest more in health in the United States (Rehavi, 2007); and education and health in India (Clots-Figueras, 2011).[29]

At the same time, there are also studies which show that education reduces corruption. For instance, Glaeser and Saks (2006) find that corruption, as measured by *"the number of government officials convicted for corrupt practices through the Federal justice department"*, is less in more educated states in America. Moreover, our findings also refute the hypothesis that the observed gender differences in the attitude towards corruption is entirely due to gender differences in social status and will disappear over time when women acquire more equal socio-economic status. Indeed if women have distinct preferences for policies that are different from that of men and if policy-making is the channel through which women in government affect corruption, then there is little reason to think that the association between women's presence and corruption will vanish as they get similar as men in social

[29]The political science literature also documents that women in the government tend to introduce and pass more bills that are concerned with the welfare of women, children and families (Thomas, 1991), advocate women's right issues more strongly than men, and participate more frequently in debates on bills that address issues of women, children and the family than debates on other bills (Taylor-Robinson and Heath, 2003).

status. Clearly this question bears further investigation.

To sum up, gender inequality is still persistent around the globe. The gender gap exists in access to education, work and participation in economic and political activities. In the Millennium Development Goal Report (2012), United Nations recognizes the importance of women's empowerment for achieving its goals. The findings of this paper suggest that women's participation in politics should not only be encouraged for the sake of obtaining gender equality but also because it has positive externalities – a negative impact on corruption.

References

1. Acemoglu, D. and J A Robinson. 2000. "Why did the West extend the franchise? Democracy, inequality, and growth in historical perspective." *Quarterly Journal of Economics* 115.4, 1167–1199.
2. — . 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91.5, 1369–1401.
3. Acemoglu, Daron, Simon Johnson, James A Robinson, and Pierre Yared. 2008. "Income and Democracy". *American Economic Review* 98.3, 808–842.
4. Ades, Alberto and Rafael Di Tella. 1999. "Rents, competition, and corruption". *American Economic Review*, 982–993.
5. Agerberg, Mattias. 2014. "Gender Differences in the EQI Data". *QoG Working Paper Series* 2014.14, 14.
6. 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.
7. Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough". *Quarterly Journal of Economics* 128.2, 469–530.
8. Anderson, Theodore W and Herman Rubin. 1949. "Estimation of the parameters of a single equation in a complete system of stochastic equations". *Annals of Mathematical Statistics*, 46–63.
9. Barro, Robert J and Jong Wha Lee. 2013. "A new data set of educational attainment in the world, 1950–2010". *Journal of Development Economics* 104, 184–198.
10. Branisa, Boris, Stephan Klasen, and Maria Ziegler. 2009. "Background Paper: The Construction of the Social Institutions and Gender Index (SIGI)".
11. — . 2013. "Gender inequality in social institutions and gendered development outcomes". *World Development* 45, 252–268.
12. Brunetti, Aymo and Beatrice Weder. 2003. "A free press is bad news for corruption". *Journal of Public Economics* 87.7, 1801–1824.
13. Charron, Nicholas, Victor Lapuente, Bo Rothstein, A Varraish, and M Hernandez. 2011. "Measuring quality of government and sub-national variation". *Report for the EU Commission of Regional Development, European Commission. Directorate-General Regional Policy*.
14. Chattopadhyay, Raghabendra and Esther Duflo. 2004. "Women as policy makers: Evidence from a randomized policy experiment in India". *Econometrica* 72.5, 1409–1443.
15. Chen, M Keith. 2013. "The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets". *American Economic Review* 103.2, 690–731.

16. Christiansen, Morten H and Simon Kirby. 2003. *Language Evolution*. Oxford University Press.

17. Clots-Figueras, Irma. 2011. "Women in politics: Evidence from the Indian States". *Journal of Public Economics* 95.7, 664–690.

18. 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.

19. Eckel, Catherine C and Philip J Grossman. 1998. "Are women less selfish than men?: Evidence from dictator experiments". *Economic Journal* 108.448, 726–735.

20. Finlay, Keith and Leandro M Magnusson. 2009. "Implementing weak-instrument robust tests for a general class of instrumental-variables models". *Stata Journal* 9.3, 398.

21. Fisman, Raymond and Edward Miguel. 2007. "Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets". *Journal of Political Economy* 115.6, 1020–1048.

22. Gay, Victor, Estefania Santacreu-Vasut, and Amir Shoham. 2014. "Language and Gender Economics". *Working Paper*.

23. Glaeser, Edward L and Raven E Saks. 2006. "Corruption in America". *Journal of Public Economics* 90.6, 1053–1072.

24. Goetz, Anne Marie. 2007. "Political Cleaners: Women as the new anti-corruption force?" *Development and Change* 38.1, 87–105.

25. Gupta, Sanjeev, Hamid Davoodi, and Rosa Alonso-Terme. 2002. "Does corruption affect income inequality and poverty?" *Economics of Governance* 3.1, 23–45.

26. Hansen, Casper Worm, Peter S Jensen, and Christian Skovsgaard. 2012. "Modern Gender Roles and Agricultural History: The Neolithic Inheritance". *Available at SSRN 2170945*.

27. Kleibergen, Frank. 2002. "Pivotal statistics for testing structural parameters in instrumental variables regression". *Econometrica* 70.5, 1781–1803.

28. Kudva, Neema. 2003. "Engineering elections: The experiences of women in Panchayati Raj in Karnataka, India". *International Journal of Politics, Culture, and Society* 16.3, 445–463.

29. La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Shleifer, and Robert W Vishny. 1997. "Trust in Large Organizations". *American Economic Review* 87.2, 333–338.

30. — . 1998. "Law and Finance". *Journal of Political Economy* 106.6, 1113–1155.

31. Mauro, Paolo. 1995. "Corruption and growth". *Qarterly Journal of Economics* 110.3, 681–712.

32. Moreira, Marcelo J. 2003. "A conditional likelihood ratio test for structural models". *Econometrica* 71.4, 1027–1048.

33. Murray, Michael P. 2006a. "Avoiding invalid instruments and coping with weak instruments". *Journal of Economic Perspectives* 20.4, 111–132.

34. — . 2006b. "The bad, the weak, and the ugly: Avoiding the pitfalls of instrumental variables estimation". *Available at SSRN 843185*.

35. North, Douglass C, William Summerhill, and Barry R Weingast. 2000. "Order, disorder, and economic change: Latin America vs. North America". *Governing for Prosperity* 19.

36. Rehavi, M Marit. 2007. "Sex and politics: Do female legislators affect state spending?" *Unpublished manuscript, University of Michigan*.

37. Reynolds, Andrew. 1999. "Women in the legislatures and executives of the world: Knocking at the highest glass ceiling". *World Politics* 51.04, 547–572.

38. Rule, Wilma and Joseph Francis Zimmerman. 1994. *Electoral systems in comparative perspective: Their impact on women and minorities*. 338. Greenwood Publishing Group.

39. Sullivan, Michael J. 1991. *Measuring global values: The ranking of 162 countries*. Greenwood Press New York.

40. Sundström, Aksel. 2013. "Womens Local Political Representation Within 30 European Countries". *QoG Working Paper Series* 2013.18, 18.

41. Sung, Hung-En. 2003. "Fairer sex or fairer system? Gender and corruption revisited". *Social Forces* 82.2, 703–723.

42. Swamy, Anand, Stephen Knack, Young Lee, and Omar Azfar. 2001. "Gender and Corruption". *Journal of Development Economics* 64.1, 25.

43. Taylor-Robinson, Michelle M and Roseanna Michelle Heath. 2003. "Do women legislators have different policy priorities than their male colleagues? A critical case test". *Women & Politics* 24.4, 77–101.

44. Thomas, Sue. 1991. "The impact of women on state legislative policies". *Journal of Politics* 53.04, 958–976.

45. Treisman, Daniel. 2000. "The causes of corruption: a cross-national study". *Journal of Public Economics* 76.3, 399–457.

46. — . 2007. "What have we learned about the causes of corruption from ten years of cross-national empirical research?" *Annual Review of Polititical Science* 10, 211–244.

47. United Nations. 2012. "The Millennium Development Goals Report 2012".

48. World Bank. 2001. "World development report 2002: building institutions for markets".

7 Figures and tables

Figure 1: Voting Rights and Presence of Women in Parliament

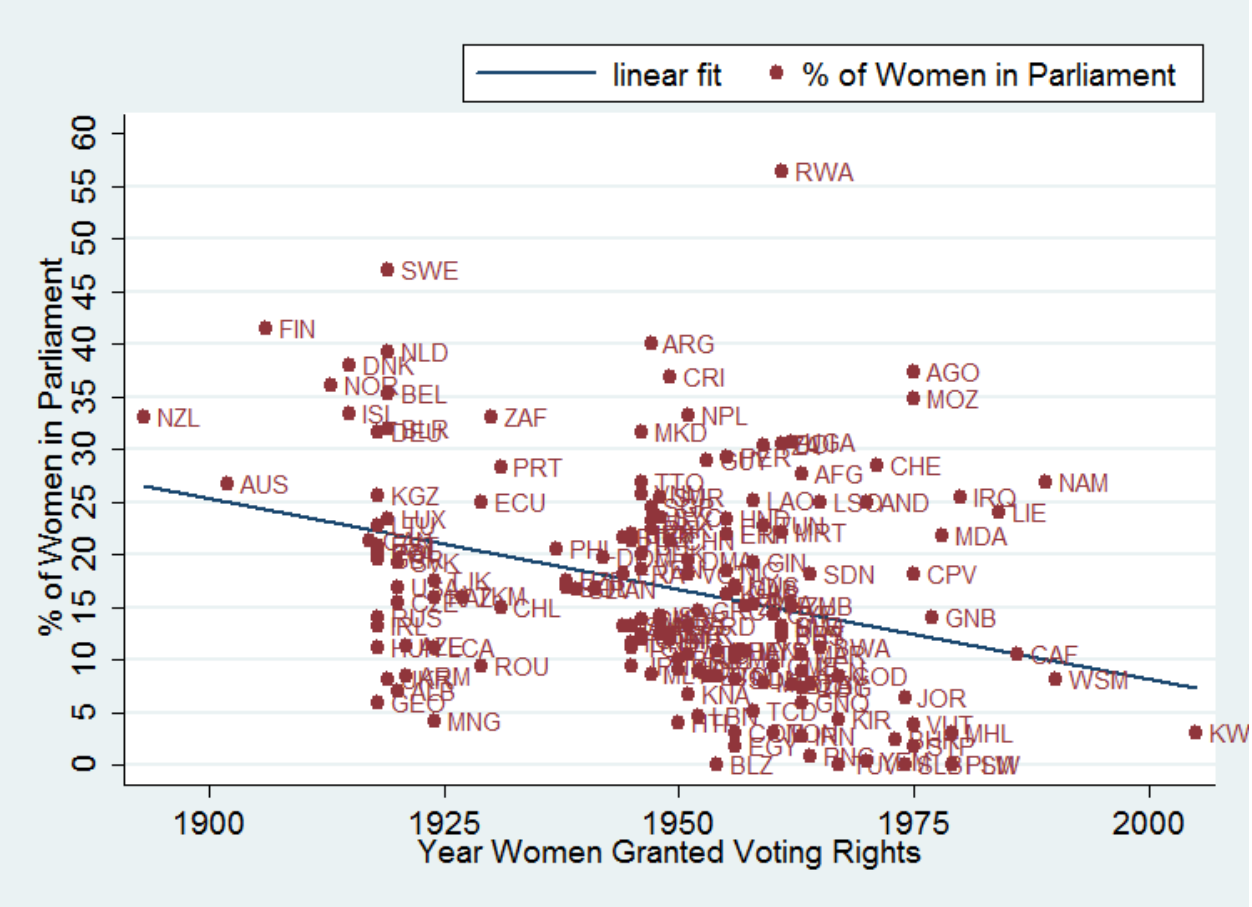

Table 1: Summary statistics

Variable	Mean	Std. Dev.	N
Control of Corruption Index	0.058	0.99	154
Corruption Perception Index	-4.734	2.349	85
Share of Women in Labor Force	41.179	8.707	154
Share of Women in clerical positions	59.218	19.277	91
Share of Women in decision making positions	27.502	10.73	91
Share of Women in Parliament	16.348	9.693	113
Gross National Product Per Capita	9154.460	13902.745	154
Years of Schooling (2010)	8.097	2.684	127
Christian Proportion (2005)	55.48	37.512	154
Muslim Proportion (2005)	23.905	34.064	154
Proportion in largest ethnic group*	68.643	24.515	133
Political Rights	-3.423	2.046	154
Civil Liberty	-3.29	1.667	154
Openness to Trade (2005)	47.543	23.563	128
Gender Inequality Index (2008)	2.487	1.645	76
Social Institutions and Gender Index (2009)	0.121	0.114	92
Power Distance Index	59.525	22.591	61

Variables – *Corruption Perception Index, share of women in the labor force, share of women in parliament, political rights, GNPPC, Civil liberty* are averaged over a period of 10 years (2001 - 2010). The Control of Corruption Index was not computed for year 2001, hence, 2001 index is replaced by 2000 index to compute the 10-year average. The data for share of women in clerical positions and decision making positions belong to the latest year (2000 - 2008) for which the data is available. The year indicated in brackets to the next of a variable indicates the year to which this data belongs. Averaging of the variables over a period of 10 years would ensure that the estimates are not disproportionately affected by any specific event in a given year.

Table 2: Women and Corruption. Dependent Variable: Corruption Index

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of women in the labor force	-0.0166**	-0.0103					-0.00828
	(0.00653)	(0.00656)					(0.00808)
Share of women in clerical positions			0.00275				
			(0.00421)				
Share of women in decision-making positions				0.00461			
				(0.00708)			
Share of women in parliament					-0.0299***	-0.0300***	-0.0293***
					(0.00477)	(0.00526)	(0.00524)
Continent Dummies	No	Yes	No	No	No	Yes	No
Observations	154	154	91	91	120	120	113
Adjusted R^2	0.742	0.754	0.763	0.763	0.801	0.802	0.807

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is negative of the Control of Corruption Index such that a higher value implies more corruption. All the specifications include baseline controls – log (GNPPC), political rights, proportion of Christians in total population, proportion of Muslims in total population, British colonial dummy, 'Never Colonized' dummy. Constant not reported.

Table 3: Women and Corruption. Dependent Variable: Corruption Index. Inclusion of Additional Variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share of women	-0.0137**	-0.00555	-0.000889	-0.0265***	-0.0131	-0.0147						
in the labor force	(0.00642)	(0.0135)	(0.00795)	(0.00780)	(0.00851)	(0.00905)						
Share of women							-0.0296***	-0.0325***	-0.0264***	-0.0330***	-0.0291***	-0.0293***
in parliament							(0.00452)	(0.00708)	(0.00623)	(0.00494)	(0.00528)	(0.00557)
Civil liberty	Yes						Yes					
PDI		Yes						Yes				
SIGI			Yes						Yes			
Schooling				Yes	Yes	Yes				Yes	Yes	Yes
LEG				Yes	Yes	Yes				Yes	Yes	Yes
Openness						Yes						Yes
Continent Dummies	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
Observations	154	61	92	114	114	112	120	48	62	85	85	84
Adjusted R^2	0.762	0.777	0.422	0.790	0.802	0.811	0.824	0.848	0.552	0.863	0.869	0.876

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is negative of the Control of Corruption Index higher value implies more corruption. All the specifications include baseline controls: log (GNPPC), political rights, proportion of Christians in total population, proportion of Muslims in total population, British colonial dummy, 'Never Colonized' dummy. In columns 1 and 7, *political rights* variable has been replaced with *civil liberty*. Constant not reported.

Table 4: Women in the Labor Force and Corruption: IV estimates. Dependent Variable: Corruption Index

	(1)	(2)	(3)	(4)
	Second-stage regression. Dependent variable: Control of Corruption Index			
Share of women in the labor force	-0.00744	-0.0137	-0.0254	-0.0211
	(0.0157)	(0.0225)	(0.0185)	(0.0190)
F-stat (excluded. inst.)	47.006	23.276	16.659	13.617
	First-stage regression. Dependent variable: Share of women in the labor force			
Number of genders = 2 in country's dominant language	-12.12***	-7.272***	-6.506***	-6.841***
	(1.768)	(1.507)	(1.594)	(1.854)
Continent dummies	Yes	Yes	Yes	Yes
Cultural variables	No	Yes	Yes	Yes
Colonial dummies	No	Yes	Yes	Yes
Basline contemporaneous controls	No	No	Yes	Yes
Extended contemporaneous controls	No	No	No	Yes
Observations	125	120	98	84

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Cultural controls: Proportion of Christians in total population, proportion of Muslims in total population. Colonial dummies: Former British colonies, Nvever colonized. Baseline contemporaneous controls: Log (GNPPC), average years of schooling. Extended contemporaneous controls: Political rights, proportion in largest ethnic group, openness to trade.

Table 5: Women in Parliament and Corruption: IV estimates. Dependent Variable: Corruption Index

	(1)	(2)	(3)	(4)
	Second-stage regression. Dependent Variable: Control of Corruption Index			
Share of women in parliament	-0.0627***	-0.0570***	-0.0907***	-0.0462***
	(0.0228)	(0.0185)	(0.0335)	(0.0135)
	[-0.111, -0.002]	[-0.112, -0.0196]	[-0.222, -0.047]	[-0.088, -0.013]
F-stat (excluded. inst.)	10.166	7.910	4.321	4.841
CLR (p-value)	0.0491	0.0108	0.0006	0.0156
AR (p-value)	0.0491	0.0108	0.0008	0.0263
LM-J (p-value)			≤ 0.05	≤ 0.05
$J-$stat (p-value)			0.0369	0.2327
Sargan $J-$stat (p-value)			0.2223	0.2494
	First-stage regression. Dependent Variable: % of women in parliament			
Excluded Instruments				
Year women granted suffrage	-0.156***	-0.203***	-0.101*	-0.194**
	(0.0490)	(0.0722)	(0.0566)	(0.0826)
log(Years since transition to agriculture)			-4.382*	-4.349
			(2.238)	(3.161)
Controls				
Independence year	Yes	Yes	Yes	Yes
Continent dummies	Yes	Yes	Yes	Yes
Colonial dummies	Yes	Yes	Yes	Yes
Cultural	Yes	Yes	No	Yes
Contemporaneous	No	Yes	No	Yes
Historical	No	No	Yes	Yes
Observations	111	83	97	81

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Figures in square brackets correspond to the 95% confidence sets based on conditional likelihood ratio (CLR) approach developed by Moreira (2003), and implemented by the Finlay and Magnusson (2009) Stata routine. The $p-$ values for the CLR, AR and LM-J statistics is against the null that the coefficient of the endogenous variable, *i.e.*, the share of women in parliament, is zero. $J-$statistic $p-$value reflects the over-identification test results with the null that instruments are valid. This $J-$statistic is evaluated at the null hypotheses, as opposed to Sargan-Hansen $J-$ statistic which is evaluated at the parameter estimate. Colonial dummies: British colony, never colonized. Cultural controls: Proportion of Christians in total population, proportion of Muslims in total population. Contemporaneous controls: log (GNPPC), political rights, average years of schooling, openness to trade, proportion in largest ethnic group. Historical controls: Plow use, agricultural suitability, tropical climate. Constant not reported.

Table 6: Women and Corruption: Sub-national Evidence.

	(1)	(2)	(3)	(4)	(5)	(6)
Share of Women in the labor force (regional)	-0.0683*** (0.0247)		-0.0369 (0.0288)	-0.0350 (0.0317)		-0.0317 (0.0305)
Share of Women in local govt.		-0.0313*** (0.0119)	-0.0144* (0.00793)		-0.0126* (0.00672)	-0.0121* (0.00648)
Share of Women in the labor force (national)			-0.0864** (0.0410)			
Share of Women in parliament			-0.0459*** (0.00789)			
Regional Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Controls	Yes	Yes	Yes	No	No	No
Country Fixed Effects	No	No	No	Yes	Yes	Yes
Countries	17	17	17	17	17	17
Observations	155	155	155	155	155	155
Adjusted R^2	0.584	0.615	0.716	0.824	0.827	0.828

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. A higher value of the dependent variable indicates the higher corruption. Regional controls: log(GDPPC), education. Country Controls: log (GNPPC), political rights. Constant not reported. Countries included in this analysis are: Austria, Belgium, Bulgaria, Czech Republic, Denmark, France, Germany, Hungary, Italy, Netherlands, Poland, Portugal, Romania, Slovakia, Spain, Sweden, and United Kingdom.

Table 7: Women and Corruption. Dependent Variable: Corruption Perception Index.

	OLS Regression				**Median Regression**		**Robust Regression**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of women	-0.0311	-0.00705			-0.0164		-0.0319	
in the labor force	(0.0199)	(0.0266)			(0.0294)		(0.0225)	
Share of women			-0.0775***	-0.0751***		-0.0865***		-0.0838***
in parliament			(0.0151)	(0.0159)		(0.0184)		(0.0138)
Continent								
Dummies	No	Yes	No	Yes	No	No	No	No
Observations	85	85	66	66	85	66	85	66
Adjusted R^2	0.775	0.779	0.859	0.867			0.751	0.865

In first 4 columns, Heteroskedasticity-robust standard errors in parentheses. In last 4 columns, standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is negative of Corruption Perception Index such that a higher value implies more corruption. All the specifications include balseline controls: $\log(\text{GNPPC})$, political rights, proportion of Christians in total population, proportion of Muslims in total population, British colonial dummy, 'Never Colonized' dummy. Constant not reported.

Table 8: "Corruption Convergence in Gender?" Dependent Variable: Corruption Index

	(1)	(2)	(3)	(4)
Share of women in the labor force	-0.0253***	-0.0214***		
	(0.00771)	(0.00725)		
Share of women in the the labor force × GII#	0.0217**	0.0554***		
	(0.00984)	(0.0129)		
Share of women in parliament			-0.0442***	-0.0483***
			(0.00618)	(0.00682)
Share of women in parliament × GII			0.0402**	0.0687***
			(0.0172)	(0.0215)
Continent Dummies	No	Yes	No	Yes
Observations	124	124	95	95
Adjusted R^2	0.786	0.824	0.870	0.885

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is negative of the Control of Corruption Index higher value implies more corruption. #GII = Gender Inequality Index. All the specifications include baseline controls: log (GNPPC), political rights, proportion of Christians in total population, proportion of Muslims in total population, British colonial dummy, 'Never Colonized' dummy. Constant not reported.

Appendix for "Women and Corruption: What Positions Must They Hold to Make a Difference?"

Chandan Kumar Jha & Sudipta Sarangi

Contents

S.1 Ordered-probit estimation

Though the corruption indices used in this paper are continuous variables they are bounded. Consequently the linear relationship estimated by ordinary least squares (OLS) may not be most appropriate in this context. Most of the papers in corruption literature, however, use OLS following the fact that corruption index is a continuous variable and *not* an ordering. Brunetti and Weder (2003) argue that OLS is an appropriate specification and perform ordered-probit regression only as sensitivity check. Note that the corruption index used by them takes discrete values, and hence, it could be argued that their dependent variable could be construed as an ordering.[30] On the other hand, the index we use is continuous and does not allow for such criticisms. Nevertheless, we check the sensitivity of our results using order-probit.

We divide our sample of countries in to 4 categories: very clean (category 1), clean (category 2), corrupt (category 3) and highly corrupt (category 4).[31] We assign these categories in the following manner. First, we divide countries in two groups – below mean *i.e.* below

[30]The primary measure of corruption used by Brunetti and Weder (2003) is an indicator of corruption produced by the International Country Risk Guide (ICRG) that takes only discrete values in the range of 0 to 6. They use the average of the index, and use OLS specification. They argue that since corruption indices are based on numerical ratings and not on country ratings, corruption indices do not actually show the ordering (or ranking) of a country. See Brunetti and Weder (2003) for detailed discussion on why OLS is the most appropriate specification.

[31]There could be several other possible ways to define the ordering. We did check the robustness of the results under other classifications such as dividing the countries in 4 quartiles and found the similar results. These results could be obtained from the authors on request.

zero (relatively less corrupt countries), and above mean (relative more corrupt countries). Next we divide these groups into two sub-groups where there is a natural break in data points. Table S.2 describes the assignment of categories and dummies, and the number of countries that fall under each category for different regression specifications.

Table S.1: Women and Corruption: Ordered-probit categories

Dummy assigned	Control of Corruption Index	Corruption Classification	Number of countries		
			Col. 1	Col. 2	Cols. 3 & 4
1	$CCI < -1.82$	Very clean	14	13	13
2	$-1.82 < CCI < 0$	Clean	39	40	35
3	$0 < CCI < 0.75$	Corrupt	61	42	40
4	$CCI > 0.75$	Highly corrupt	40	25	25

Please note that the CCI values in the Table is *negative* of the actual CCI. Col. 1, 2, 3 and 4 refers to the columns in Table S.2.

Then we estimate an ordered-probit regression with the corruption variable defined as an ordering. In this context, ordered-probit estimation ensures that the corruption index does not exceed the maximum possible value (*i.e.* 2.5) for any country, when the variable of interest (such as the share of women in parliament) is very large. The results are presented in panel 1 of the Table S.2. The results are similar to those obtained from OLS specification – while the share of women in the labor force is not significantly associated with corruption (columns 1 and 3), the coefficient of the share of women in parliament is significant at less than 1% level in columns 2 and 4.

Panel 2 of the Table S.2 reports the marginal effects of change in regressors (only female participation variables reported) computed at their means for specifications in columns 1 and 2. If our hypothesis is correct, then a higher level of female participation should be associated with a higher probability of being in clean categories, and hence the marginal effect is expected to be positive for the first categories, and should change sign as we move from very clean to highly corrupt. This is exactly what panel 2 shows. While the marginal

Table S.2: Panel 1: Women and Corruption. Dependent Variable: Corruption Index

	Panel 1: : Ordered-probit estimates.			
	(1)	(2)	(3)	(4)
Share of Women in the Labor Force	-0.0139		-0.0158	
	(0.0161)		(0.0184)	
Share of Women in Parliament		-0.0789***		-0.0809***
		(0.0142)		(0.0150)
Observations	154	120	113	113
Pseudo R^2	0.434	0.550	0.458	0.546
	Panel 2: Marginal effects evaluated at the mean			
	Category 1 (Very clean)	Category 2 (Clean)	Category 3 (Corrupt)	Category 4 (Highly corrupt)
Share of women in the labor force	0.0000558	0.0042465	-0.0023942	-0.001908
	(0.0000777)	(0.0049845)	(0.002975)	(0.0021742)
Share of women in parliament	0.0000451	0.0312035***	-0.0290816***	-0.002167*
	(0.000063)	(0.0056669)	(0.0058838)	(0.0012644)

Panel 1: Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Countries have been divided in 4 catogeries in the Control of Corruption Index. Each category has been assigned a dummy – very clean (1), clean (2), corrupt (3) and highly corrupt (4). Dependent variable is the dummy assigned to each category. Columns 3 and 4 present the results of the same set of countries for both the variables of interest in order to check if the differences in the significance for our variables of interest is driven because of sample issues. All the specifications include baseline controls: log (GNPPC), political rights, proportion of Christians in total population, proportion of Muslims in total population, British colonial dummy, 'Never Colonized' dummy. Constant not reported.
Panel 2: Delta method standard errors in parentheses. Marginal effects reported here correspond to columns 1 and 2 of Table S.2

effect is consistent with the hypothesis for both the female participation variables, it is not significant for the share of women in the labor force variable for any of the four categories. For the share of women in parliament, however, the marginal effect is positive in first two columns, and is highly significant for category 2 indicating that countries which have a higher presence of women in parliament, are more likely to be in "clean" category. In the last two columns, marginal effect is negative and statistically significant indicating that countries, where the share of women in parliament is higher, have the lower probability to fall in "high corruption categories" (categories 3 and 4).

S.2 Median and robust regressions

Table S.3 presents the results of median and robust regressions.[32] The advantage of these regressions is that the estimates are not affected disproportionately because of the presence of outliers.[33] As we can see, the coefficient of the share of women in the labor force is significant at conventional levels in column 1 (median regression) and column 5 (robust regression). However, once continent dummies are added to the model, the coefficient on the share of women in the labor force loses its significance in each regression (columns 2 and 6). On the other hand, the coefficient on the share of women in parliament is always significant at conventional levels (columns 3, 4 and 7, 8).

[32]The robust regression, as performed by "rreg" command in stata 12, drops the influential observations (Cook's distance > 1), and performs the Huber iterations. It also down-weighs the observations with large absolute residuals (Introduction to STATA. UCLA: Statistical Consulting Group. from http://www.ats.ucla.edu/stat/stata/dae/rreg.htm accessed March 22, 2013).

[33] The least squares estimators are sensitive to outliers especially in the small samples. A class of robust estimators which are unaffected or less sensitive by the presence of outlying observations such as least absolute deviations (LAD) estimators might therefore be preferable in small samples. The median regression is a type of LAD estimator and is a robust alternative to the least squares regression. See Greene (2012) for a discussion on robust estimator.

Table S.3: Women and Corruption (Median and Robust Regression). Dependent Variable: Corruption Index.

	Median Regression				**Robust Regression**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of women in the labor force	-0.0137**	-0.00732			-0.0109*	-0.00607		
	(0.00667)	(0.00803)			(0.00603)	(0.00660)		
Share of women in parliament			-0.0275***	-0.0295***			-0.0305***	-0.0306***
			(0.00538)	(0.00647)			(0.00432)	(0.00463)
Continent Dummies	No	Yes	No	Yes	No	Yes	No	Yes
Observations	154	154	120	120	154	154	120	120
Adjusted R^2					0.779	0.789	0.838	0.839

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is negative of the Control of Corruption Index higher value implies more corruption. All the specifications include baseline controls: log (GNPPC), political rights, proportion of Christians in total population, proportion of Muslims in total population, British colonial dummy, 'Never Colonized' dummy. Constant not reported.

S.3 Restricted sample regressions

There are variations in the number of observations in the different specification in Table 2 in the main text. This is due to the fact that the data for women's share in the labor force is available for a larger number of countries than it is available for women's share in parliament. And, the data for the latter is available for a greater number of countries than it is available for the share of women in clerical positions and the share of women in decision-making positions. In order to rule out the concern that sample selection is responsible for the differences in the significance of our variables of interest, we run regression restricted to the sample of countries for which data is available for all the female participation variables, and find similar results as reported in the Table 2 in the main text. The results are presented in Table S.4.

Table S.4: Women and Corruption (Restricted Sample). Dependent Variable: Corruption Index

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share of women in the labor force	-0.0195*	-0.0241**					-0.0103
	(0.0112)	(0.0114)					(0.0105)
Share of women in clerical positions			0.00115				
			(0.00580)				
Share of women in decision-making positions				-0.00316			
				(0.00745)			
Share of women in parliament					-0.0328***	-0.0296***	-0.0315***
					(0.00672)	(0.00821)	(0.00665)
Continent Dummies	No	Yes	No	No	No	Yes	No
Observations	71	71	71	71	71	71	71
Adjusted R^2	0.788	0.810	0.779	0.779	0.847	0.850	0.847

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is negative of the Control of Corruption Index higher value implies more corruption. Regressions restricted to the sample of countries for which data is available for all women participation variables. All the specifications include baseline controls: log (GNPPC), political rights, proportion of Christians in total population, proportion of Muslims in total population, British colonial dummy, 'Never Colonized' dummy. Constant not reported. Sample restricted to the set of countries for which data is available for all the women's participation variables.

The number of observations in the specification in which the variable of interest is women's share in parliament is smaller compared to the specifications in which the variable of interest is women's presence in the labor force in Table 3. If the loss of observation is non-random, it may cause the differences in the significance of the two variables of interest. To address this issue, we present the results of the regressions of women's presence in the labor force on corruption restricted to the countries for which this analysis has been reported in the main text. The results are similar, and presented in Table S.5.

Table S.5: Robustness check: Women and Corruption (Restricted Sample). Dependent Variable: Corruption Index

	(1)	(2)	(3)	(4)	(5)	(6)
Share of women in the Labor Force	-0.0197** (0.00829)	0.0133 (0.0184)	-0.0119 (0.0115)	-0.0271*** (0.00974)	-0.0175 (0.0117)	-0.0181 (0.0123)
Continent Dummies	No	No	No	No	Yes	Yes
Observations	113	48	62	85	85	84
Adjusted R^2	0.778	0.791	0.458	0.808	0.816	0.820

Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is negative of the Control of Corruption Index higher value implies more corruption. Regressions restricted to the sample of countries which are included in columns 7-12 in Table 3 in the main text. All the specifications include baseline controls: log (GNPPC), political rights, proportion of Christians in total population, proportion of Muslims in total population, British colonial dummy, 'Never Colonized' dummy. Constant not reported.