Response to Reviewer's Comments
Does Corruption Deter Female Leadership in Firms?
Manuscript JOIE-2025-0056

July 14, 2025

We thank the six referees for their comments, concerns, and suggestions. In this document we detail our responses and the revisions that have been made to the paper. We hope that they are satisfactory and believe that the paper is now improved.

All quoted sections of the revised text that appear below are indented. (Since the revised text is "copied-and-pasted" into this document, please note that footnotes are numbered differently here than in the paper.)

Referee comments are grouped according to topic. Given the number and length of referee reports, we do not reproduce each of their comments below but we note which comment we are specifically referring to. We then copy all reviewer comments at the end of this document. We label each reviewer R# according to how they were presented in the system. We put their specific comment number in parentheses (e.g., **R1(3)** implies reviewer #1, comment number 3).

Concerns Regarding Data Clarity

Reviewers 1, 3, 4, 5, and 6 raised important questions regarding the data description. This section of our response refers to **R1(3)**, **R3(1)**, **R3(2)**, **R3(5)**, **R4(1)**, **R5(3)**, and **R6(8)**.

The first concern in this group relates to what was previously Table 1 (now Table 2) and the construction of the outcome variables. For example, **R1(3)** is concerned with the minimum and maximum of the first set of outcome variables being at 0 or 1; **R3(1)** is concerned with the different sample sizes; **R4(1)** believes we should emphasize the employer variable more and highlights somethings that we were not clear on. All three comments can be addressed by better clarifying the measures within the table and in the paper. We have therefore separated what was previously Table 1 into two different tables - now Table 2 for outcomes and Table 3 for independent variables. Table 1 is now a table listing each outcome variable and their detailed definitions. We have altered the subtext in the tables to better explain what each group of outcomes measures.

We have also added to the notes in Table 2. Note 1 explains that the observation numbers are smaller for the first group of outcomes because some municipalities have *no* leadership roles. It also clarifies that the denominator is different for each category of leadership in this first group and that the denominator does not change when looking at the other outcomes. Now that the denominator is more clearly defined given the aforementioned changes, it is easier to explain that given the lack of any leadership positions the ratio of female leaders to total leaders is undefined for these municipalities. This also helps clarify why the minimum and maximum here is 0 and 1.

The new tables and associated notes are copied immediately below.

Table 1: Outcome Variable Definitions

Full Sample

Variable	Definition
Female Presence in Leadership Positions	
Employer	Total female employers divided by total employers
Managers, Directors, or Executives	Total female managers, directors, or executives (MDE) divided by total MDE
Leadership	Total female employers, managers, directors, and executives (leaders) divided by total leaders
Female Labor Force Participation	
Female Labor Force Participation	Female workers divided by total number of working age women
Female Labor Force Job Type	
Employer	Total female employers divided by total women in the labor force
Managers, Directors, or Executives	Total female managers, directors, or executives (MDE) divided by total women in the labor force
Leadership	Total female employers, managers, directors, and executives (leaders) divided by total women in the labor force

"Corrupt" Sectors Only

Variable	Definition
Female Presence in Leadership Positions	
Employer	Total female employers in "corrupt" sectors divided by total employers in "corrupt" sectors
Managers, Directors, or Executives	Total female managers, directors, or executives (MDE) in "corrupt" sectors divided by total MDE in "corrupt" sectors
Leadership	Total female employers, managers, directors, and executives (leaders) in "corrupt" sectors divided by total leaders in "corrupt" sectors
Female Labor Force Participation	
Female Labor Force Participation	Female workers in "corrupt" sectors divided by total number of working age women
Female Labor Force Job Type	
Employer	Total female employers in "corrupt" sectors divided by total female workers in corrupt sectors
Managers, Directors, or Executives	Total female managers, directors, or executives (MDE) in "corrupt" sectors divided by total female workers in corrupt sectors
Leadership	Total female employers, managers, directors, and executives (leaders) in "corrupt" sectors divided by total female workers in corrupt sectors

Notes: All outcome variables are calculated using the 2010 Census.

Table 2: Summary statistics for outcome measures.

Variable	Obs	Mean	Std. Dev.	Min	Max
Outcomes					
Female Presence in Leadership Positions[1]					
Female leaders divided by total leadership positions in each category.					
Employers	878	0.263	0.184	0	1
Managers, Directors, or Executives	930	0.362	0.167	0	1
Leadership	933	0.331	0.138	0	1
Female Labor Force Participation					
Female workers divided by total number of working age women.					
Female Labor Force Participation	935	0.330	0.115	0.116	0.728
Female Labor Force Job Type					
Female leaders divided by total number of female workers.					
Employers	935	0.010	0.010	0.000	0.062
Managers, Directors, or Executives	935	0.026	0.014	0.000	0.072
Leadership	935	0.036	0.019	0.000	0.118
Outcomes - "Corrupt" Sectors Only[2]					
Female Presence in Leadership Positions[1]					
Female leaders divided by total leadership positions (per category) in corrupt sectors.					
Employer	553	0.188	0.242	0	1
Managers, Directors, or Executives	639	0.207	0.259	0	1
Leadership	719	0.197	0.240	0	1
Female Labor Force Participation					
Female workers in corrupt sectors divided by total number of working age women.					
Female Labor Force Participation	935	0.021	0.021	0.001	0.181
Female Labor Force Job Type					
Female leaders in corrupt sectors divided by total number of female workers in corrupt sectors.					
Employer	935	0.001	0.002	0	0.024
Managers, Directors, or Executives	935	0.001	0.003	0	0.046
Leadership	935	0.002	0.005	0	0.070

Notes: [1]The observation numbers in this group are lower because some municipalities are small and have *no* leadership positions, making the denominator zero as the denominator is different for each category of leadership (i.e., total employers, total managers, directors, or executives, and total leaders). The denominator is constant across all other categories of leaderships for the other outcome measures. All municipalities have women and working women and thus there are no undefined observations when using the other measures. It is likely that these small municipalities also comprise the minimum (0% of leadership positions held by women) and maximum (100% of leadership positions held by women) of these measures. [2] "Corrupt" sectors are extractive industries, manufacturing, construction, and transportation and communication, following Bologna and Ross (2015).

Table 3: Summary statistics for main independent variable (corruption) and controls.

Variable	Mean	Std. Dev.	Min	Max
Independent Variables				
Corruption				
(Log) Corruption per-capita	0.429	0.067	0.181	0.631
Municipal Level Controls (Baseline Controls)				
GDP per-capita (R$)	11727.39	17367.88	2261.63	298819.8
Population Density	81.360	341.494	0.225	6140.697
Size of Informal Sector (%)	0.597	0.191	0.128	0.970
College Degree (%)	0.050	0.030	0.003	0.235
Male (%)	0.505	0.015	0.465	0.658
Working Age (18-65 years) (%)	0.595	0.046	0.392	0.698
Urban (%)	0.621	0.217	0.050	0.999
Instrumental Variables[1]				
Number of Councils[2]	3.157	0.702	1	6
Number of Councils[2] installed	2.804	0.762	1	5
Management Capacity Index	2.151	1.237	1	6
Has Local Judge	0.318	0.466	0	1

Notes: GDP per-capita and population density enter regression in logged form. Summary statistics given for full sample of 935 municipalities. [1]For instrumental variables, they refer to 929 municipalities; data is unavailable for 6 municipalities emancipated in the early 2000s, after the creation of the Management Capacity Index (IQIM - *Indicador de Qualidade Institucional Municipal*). [2] In the context of Brazil, municipal councils (*conselhos municipais*) are commissions established by law to propose or advise on policy initiatives in a specific area (e.g., health, education) and oversee their implementation. These councils typically include representatives from the local public administration and civil society organizations. Notably, they differ from city councils (*câmaras municipais*), which serve as the local legislative branch. The measures report indexes on the number of councils and number of *active* councils, scaled from 1 (least councils) to 6 (most councils). A council is coded as active if they have individuals appointed in positions.

We have also added some discussion to Section 3.3 to clarify the first set of outcomes. This is in response to **R4(1)**. We have clarified that our outcome is exactly what the reviewer wishes to see us use. We apologize for the lack of clarity in the previous draft. We have also added some discussion throughout as to why looking at employers separately from other types of leadership roles is particularly interesting. We have copied the relevant part of Section 3.3 below.

> "Our main measure is the share of all leadership positions - i.e., employees that are in a manager, general director, or executive role *and* employers - that are occupied by women. To avoid ambiguity, this measure is calculated as:

$$y_{ms} = \frac{\sum_{i=1}^{I} WomanInLeadershipRole}{\sum_{i=1}^{I} LeadershipRole} \qquad (1)$$

> where i indexes the individual and I represents the last individual in the

sample. We also consider the share of positions in each category held by women: employees in management roles (managers, directors, or executives) and employers. Note that in these latter two cases, the denominator focuses only on total leadership positions in the respective category. For example, for employers, our measures is total female employers divided by all employers in the municipality. This also helps explain why the observation numbers drop for these separate categories: some municipalities are very small and have *no* employers and/or positions of leadership. On average, around 33% of all leadership positions are held by women in our sample (see Table 2). This number is lower when looking at shares of employers (26.3%), and higher when looking at employees in management roles (35%)."

We have also added some discussion to Section 3.4 in the paper to address some concerns with the control variables mentioned in these specific comments. First, we have added two citations motivating the selection of our control variables. These are both studies focused on corruption in Brazil and the latter, Bologna (2016), emphasizes the role of the informal sector relative to corruption in development. (**R5(3)** expressed concerns regarding this control specifically. Given the word count and the number of reviewer comments that we need to address in this revision, we cannot expand beyond the brief discussion we present in this revised section.) Second, we added some discussion regarding the diversity of the municipalities as the maximum (97%) number for informal sector size was mentioned by **R1(3)**. Third, we note that we have added an online appendix where we also show descriptive statistics for all baseline controls in the non-audited, but eligible municipalities (**R3(5)**).

The revised portion of Section 3.4 is copied below.

"Following Bologna and Ross (2015) and Bologna (2016), our control variables aim to capture the general economic environment of the municipality, with an emphasis on the level of overall development. These variables include (logged) GDP per-capita, (logged) population density, the size of the informal sector (measured as a share of employment), the percent of the adult (25 years or older) population with a college degree, the percent of the population that is male, the percent of the population that is of a standard working age (18–65), and the share of the population that lives in an urban area. These are referred to as our baseline controls.

Most of these controls are relatively standard in development, but the size of the informal economy perhaps warrants further discussion. Bologna (2016) emphasizes the importance of controlling for the size of the informal sector in any analysis of corruption and income, as corruption and informality are likely correlated.[1] Indeed, Bologna (2016) finds that informality is a stronger predictor of income than corruption in their sample of Brazilian municipalities. Moreover, given the independent importance of informality in explaining gender gaps (Malta et al., 2019), this is an important control in our analysis.

---

[1]Though, as noted in Dreher and Schneider (2010) the nature of the corruption–informality relationship is likely context specific.

Summary statistics are presented for these baseline controls in Table 3 using the largest sample of municipalities (935) in our study.[2] Note that these municipalities are diverse, and some municipalities are very small (as low as 1,409 in population). This striking diversity is also reflected in the informality and urban measures. For example, while informality is common in Brazil (approximately 60% of employment is informal on average), some municipalities rely almost exclusively on informal employment (97%). The share of the population that lives in an urban area ranges from only 5% to nearly 100% across included municipalities."

Lastly, in regard to **R3(2)** and **R6(3)**, we have clarified that the cross-sectional data comes from the 2010 Census and have further clarified why using audits that occurred after 2010 is reasonable. We explain why this is reasonable in the following section.

Clarification of the Audit Program

A portion of the **R3(5)** comment was addressed above, but a portion will be addressed here. Similarly, we expand on the use of the 2010 and beyond audits questioned in **R3(2)**. These edits also aim to address **R6(4)**. **R3(6)** also had a question concerning the program; and an econometric suggestion in **R3(8)**.

We copy all of this below but want to clarify a simple point here regarding **R3(5)**. The Avis et al. (2018) paper focuses on two electoral cycles (2004-2008; 2008-2012) and thus only uses audits occurring from 2006-2013 because the audits involve all expenditures over the previous 3–4 years. Doing so excludes the first 21 lotteries as the program began in 2003, so it is unsurprising the sample size drops so much. We have clarified this in the text. We also hope that this discussion has clarified why the lottery subscript is on the corruption variable and why the lottery fixed effect is included (**R3(6)**). We also added a footnote describing a robustness check with a dummy variable included to capture multiple audits (**R3(8)**).

Our main edits to address these three comments (**R3(2)**, **R3(5)**, **R3(6)**, **R6(4)**) are to the 3.1 and 3.2. We have these copied below.

3.1 Empirical Strategy

We use cross-sectional, municipal level data to test whether corruption has an effect on female presence in leadership positions in Brazil. All outcome and control variables come from the 2010 Census, except municipal GDP per capita and population density, collected from IPEAData. The corruption data comes from Avis et al. (2018), who construct a municipal corruption measure stemming from a random audit program. While this measure considers any

---

[2]See Online Appendix C for additional summary statistics using alternative (smaller) samples.

corrupt activity occurring between 2003 – 2013, we treat it as cross-sectional as (1) we do not know the precise timing of the corrupt activity, and (2) corruption is a highly persistent activity (Mishra, 2006). However, we note that there could be some variation in corruption levels depending on when the municipality was audited and we therefore always include audit number fixed effects. (These audits occur according to national lotteries where a subset is selected at a given time.) Our cross-sectional sample is limited to the 935 municipalities that were audited and have corruption data available. We give more details pertaining to the audit program, the construction of the corruption measurement, and the sample in Section 3.2.

We aim to estimate the effect of corruption on two sets of female leadership variables, described in detail below, after controlling for a number of municipal level characteristics. In other words, we estimate the following equation:

$$y_{ms} = \beta \times Corruption_{msl} + \theta \times X_{ms} + \alpha_s + \delta_l + \epsilon_s \qquad (2)$$

where $m, s, and l$ index the municipality, state, and audit number (lottery) of the program; y is one of our several outcomes of interest, detailed in Section 3.3; $Corruption$ is a measure of corruption per capita; X is a matrix of municipal level controls; δ is an audit/lottery fixed effect; α is a state fixed effect; ϵ is the error term. Standard errors are clustered by state. Though the corruption data is cross-sectional, we use the $m, s, and l$ subscripts to emphasize that these corruption scores come from different audits.

We estimate Equation (1) using Ordinary Least Squares (OLS). This is our baseline and focus. We also test the robustness of our results using Two-Stage Least Squares (2SLS) and an instrumental variable analysis, where we instrument for corruption with measures of political participation, management capacity, and judicial representation. We motivate and discuss the plausibility of these instruments in Section **??**. However, because our sample size is dictated by the corruption data and understanding the details of this measure is necessary to describe the nature of the IVs, we detail the corruption measure first.

3.2 Corruption

We use Avis et al. (2018)'s measure of corruption, defined as the (log) number of corruption instances uncovered in a random audit of the municipality. We scale this number by population to account for the large variation in municipal size. We explain the details of the random audit program, the corruption measure, and sample in this section.

In Brazil, a significant portion of municipal funding comes in the form of constitutionally-mandated transfers from the federal government. Discretion

concerning how to use these funds is largely left to the municipality, which invites misuse and corruption at the local level. To combat this corruption, Brazil implemented a random audit program (*Programa de Fiscalização por Sorteios Públicos*) commencing in 2003 where municipalities were selected via national lottery and subsequently audited for corruption.[3] Each lottery includes a subset of municipalities, ranging anywhere from 5 to 60, and the frequency of the lotteries has varied throughout the program. The goal of the audits is to uncover malfeasance in the use of public funds. Teams of auditors are sent to investigate any irregularities in the application of the federal transfers, and the results of the audits are made public. Avis et al. (2018) use these reports to construct their measure of corruption. They define any irregularity that the *Corregedoria Geral da União* (CGU) defines as moderate to severe as an act of corruption; the sum of these acts is their measure of municipal level of corruption.[4] Similar measures derived from these audits are now commonly employed in the literature (e.g., Ferraz and Finan, 2008, 2011; Bologna and Ross, 2015; Brollo and Troiano, 2016; Bologna Pavlik, 2018; Colonnelli and Prem, 2022).

Brazil has a total of 5,570 municipalities. From 2003 through 2015, 1,949 of these municipalities experienced at least one audit with 2,241 audits in total (Avis et al., 2018). Because Avis et al. (2018) focus on two electoral terms (namely, 2004-2008 and 2008-2012) and because audits ceased the random component in 2015, their dataset focuses on corruption uncovered in audits that occurred anytime between July of 2006 to March of 2013; i.e., lotteries 22 through 38. Since these audits uncover corrupt activity occurring anytime in the 3–4 years prior to the audit, they expose corruption *ex post*. In principle, then, this corruption measure contains corruption occurring anytime in the 2003 through 2013 period. We cannot obtain information on the exact timing of the corrupt activity.

The fact that these audits expose corruption *ex post* is particularly important for our analysis. First, even audits occurring in 2013 are likely to contain corrupt activities from 2010 or earlier. Thus, given our focus on 2010 outcomes, these later audits are relevant for our analysis.[5] Second, we are interested in the cross-sectional variation of corruption. Because corruption is

---

[3]Extremely large municipalities, specifically those greater than 500,000 in population, and state capitals were excluded from nearly all audits. Only one state capital (Aracaju) was audited once in lottery 22. Our analysis includes only those that were audited and therefore excludes these large municipalities.

[4]The CGU is the agency responsible for conducting audits and disseminating reports.

[5]In Online Appendix A we test the robustness of our results using two different measures of corruption. First, we use the Avis et al. (2018) discussed here but exclude all audit data that comes from 2010 or beyond. Our outcomes come from the 2010 Census and including post-2010 corruption could introduce reverse causation concerns. Second, we use a much more limited sample and measure from Ferraz and Finan (2011). This measure considers only the first 11 lotteries and is constructed using the authors' own definition of corruption. Lastly, though our focus is on the municipal level, we also weight regressions based on population for robustness. We construct our outcomes using averages of individual level data and thus smaller municipalities could have measures that are inaccurate. Our results are robust to these changes. We only lose significance as sample size falls.

extremely persistent, corrupt activity occurring at any point in this ten year period is likely reflective of the corrupt environment that exists within the municipality. For example, taking the sub-sample of municipalities in Avis et al. (2018) that were audited multiple times *and* have corruption measurements available in each audit (48),[6] we find that the average change in their corruption scores (log number of corruption instances) is approximately zero.[7] More importantly for this paper, the relative ranking of corruption scores moves little over time for these municipalities. Though, as cautioned in Avis et al. (2018), audit protocol differed across lotteries making corruption levels difficult to compare through time. This is why all regressions include lottery fixed effects and our focus is on cross-sectional, within lottery effects.

Avis et al. (2018)'s sample contains 1,020 audits across 967 municipalities. For the 53 municipalities that were audited multiple times in this period, we average their corruption scores across the multiple audits (when possible).[8] We also drop municipalities where corruption scores were missing in the Avis et al. (2018) dataset. There were only 37 instances of this, 5 of which were municipalities that were audited twice and thus we have a corruption score from the other audit.[9] This leaves with us 935 municipalities for which we can study the corruption–female leadership relationship. This sample is limited to only those eligible and selected for audit; large municipalities are excluded.[10]"

Concerns with Industry Share Controls

R3(4) suggests that seeing the coefficients on industry share controls would be beneficial. While we agree that they could be useful, we are reluctant to include them in the main analysis for a few reasons. First, we are not aiming to estimate the effect of industry shares on leadership outcomes and thus their interpretation is not clear. This is a common problem when interpreting control variables in regression (see, e.g., Hüner-

---

[6]The Avis et al. (2018) identification strategy is not a standard difference-in-difference and/or event study and thus does not rely on time variation. Rather, they compare corruption scores within each lottery of those that were audited for the first time versus those that were audited in the past and rely on the randomization of the program to difference out unobservable effects.

[7]This is not contradicting the Avis et al. (2018) result that the audits reduce corruption as there is no counterfactual here. The mean change in logged corruption irregularities is -0.067. To understand the impact of the audits and whether this change is significant, we would need a counterfactual over the same time period where corruption is observed but no audit occurred. However, measurement of corruption depends on being audited.

[8]There are 224 that experienced a "treatment" in their sample. Given the results of Avis et al. (2018), in Online Appendix A we include a dummy variable to control for whether the municipality was audited twice. Our results are largely unchanged.

[9]It is unclear why the data is missing in these cases, but we note that these observations are missing all audit data. Since Avis et al. (2018) rely on the CGU's coding of corruption rather than their own, this likely indicates that the CGU did not code the results of these specific audits.

[10]We compare summary statistics for control variables in this sample of 935 with other eligible but non-audited municipalities in Appendix C, Table C2. Given the random nature of the audit program, there is very little (meaningful) difference between these groups.

mund and Louw, 2025). Second, even if the coefficients on the "corrupted" sectors are insignificant, there could still be significant correlation between corruption and industry shares where the individual significance is lacking due to multicollinearity. Lastly, while the coefficients could be beneficial, we chose to relegate them to an Online Appendix due to the aforementioned concerns and journal length constraints set by the *Journal of Institutional Economics.* Please see the Online Appendix F for these coefficients.

Concerning the latter half of the comment, we have added results on the non-corrupted sector (this is simply the remaining industries that we have not labeled as corrupt). These results are now in Online Appendix D. Within this appendix, we have also added additional summary statistics where we test the difference in means of our outcome variables between the corrupted sectors and the non-corrupted sectors. As expected, the means are statistically different. We believe this greatly benefits the discussion and also addresses **R5**'s comments **(5)** and **(6)** as well. Lastly, we have tried to clarify our terminology throughout. We hope this clarification also addresses **R5**'s comments **1** and **2**.

R5(4) notes that we need to clarify why we included sectoral shares only as a robustness test yet include an analysis on a sub-sample of corrupt sectors. This is an important point to clarify. Controls are important in that they address omitted variable concerns that something other than corruption could be the underlying cause. However, we are still estimating an overall, average effect of corruption on the outcome. The sub-sample regressions have a different purpose. These regressions allow the effect of corruption to change in different sub-samples. This would be analogous to including an interaction term in a regression. (Note that we cannot include an interaction term here as we are calculating these measures at the municipality (not individual) level.) In this case, we are no longer interested in an average effect on the whole sample but the average effect in the corrupt sample only. These effects can be different. We have clarified why including sectoral shares could be important in our analysis (copied below). We have also better motivated the need for sub-sample testing (also copied below).

3.4 Control Variables

Following Bologna and Ross (2015) and Bologna (2016), our control variables aim to capture the general economic environment of the municipality, with an emphasis on the level of overall development. These variables include (logged) GDP per-capita, (logged) population density, the size of the informal sector (measured as a share of employment), the percent of the adult (25 years or older) population with a college degree, the percent of the population that is male, the percent of the population that is of a standard working age (18–65), and the share of the population that lives in an urban area. These are referred to as our baseline controls.

Most of these controls are relatively standard in development, but the size of the informal economy perhaps warrants further discussion. Bologna (2016) emphasizes the importance of controlling for the size of the informal sector

in any analysis of corruption and income, as corruption and informality are likely correlated.[11] Indeed, Bologna (2016) finds that informality is a stronger predictor of income than corruption in their sample of Brazilian municipalities. Moreover, given the independent importance of informality in explaining gender gaps (Malta et al., 2019), this is an important control in our analysis.

Summary statistics are presented for these baseline controls in Table 3 using the largest sample of municipalities (935) in our study.[12] Note that these municipalities are diverse, and some municipalities are very small (as low as 1,409 in population). This striking diversity is also reflected in the informality and urban measures. For example, while informality is common in Brazil (approximately 60% of employment is informal on average), some municipalities rely almost exclusively on informal employment (97%). The share of the population that lives in an urban area ranges from only 5% to nearly 100% across included municipalities.

In Online Appendix B, we also include industry employment shares as controls. Regressions are at the municipality (as opposed to individual) level and municipalities have different sectoral compositions. However, our focus is on the estimates without these measures, as they are likely "bad" controls à la Angrist and Pischke (2009). The effect of corruption on female leadership is likely to work (in part) through industry choice and, therefore, by including sector composition as a control we could be eliminating one of the main channels through which corruption can impact female leadership. Our results in Online Appendix B confirm this suspicion. Once we control for for sectoral composition, much of the effect of corruption disappears. This does not suggest that corruption is irrelevant, however, it simply implies that much of corruption's effect works through sectoral composition and that the two are highly correlated. We believe studying the effect of corruption on sector choice for women would be beneficial for future research, but it is beyond the scope of this paper to do so.

3.5 "Corrupt" Sector Sub-Sample

Our main set of results uses all individual level Census data to create female leadership presence and labor force measures for each municipality. As such, we get an estimate of how municipal corruption impacts female leadership overall. However, this average effect includes corruption's impact on female leadership across all sectors and may not be reflective of corruption's impact on leadership in sectors that are more or less prone to corruption. Traditionally corrupt sectors could have lower female presence in general, making it particularly difficult for women to advance their careers. On the other hand,

---

[11]Though, as noted in Dreher and Schneider (2010) the nature of the corruption–informality relationship is likely context specific.

[12]See Online Appendix C for additional summary statistics using alternative (smaller) samples.

the women that are willing to work in these sectors within a relatively corrupt country could be more risk-tolerant and willing to engage in corruption which could increase female representation among leaders in this group. A highly corrupt municipality could signal a stronger corruption "treatment" in these sectors, relative to others, and thus may have a different effect on leadership in corruption prone sectors than the average.

We follow Bologna and Ross (2015) and consider the following four sectors as corruption-prone: (1) extractive (e.g., mining), (2) manufacturing, (3) construction, and (4) transportation/communication sectors. Bologna and Ross (2015) identify these four sectors as "corrupt" using the the Organization for Economic Cooperation and Development (OECD) Foreign Bribery Report (2014). Within this report, these four sectors accounted for nearly 70% of all corruption cases studied. While corruption can and likely does exist in any industry,[13] we define these four sectors as "corrupt" sectors and focus our analysis on this sub-sample. However, in Online Appendix D, we also present data and results for the "non-corrupt" sub-sample, defined as all remaining industries as a comparison.

Table 2 presents summary statistics for this sub-sample along with the full sample discussed above. Female labor force participation in these "corrupt" sectors is much lower than the overall average. Only 2.1% of working age women are employed in these sectors, though 33% of women in this group are working in general (across any sector). Thus, these industries comprise only 6.36% of total female workers in the economy; this is compared to the 19.72% of total employment that these industries constitute overall. Perhaps more importantly, women are less likely to be leaders within these industries. Female leaders as a share of total leadership positions across all industries averages 33.1%, whereas female leaders within these four sectors is less than 20%. Moreover, female leadership shares in the other (non-corrupt) sectors is much higher at 35.2% (Online Appendix, Table D1). These statistics highlight potential differences in female representation in corruption-prone industries and thus motivate the idea that the effect of corruption could be different in this more susceptible group.

R6(6) states that they are "reluctant to buy the argument that the sectoral share of females in leadership is driven by corruption". We apologize but we are unsure if the reviewer is referring to our main analysis "share of females in leadership positions" or our control variables which include sectoral shares. We believe they are referring to the latter and want to emphasize that studying specific sectoral shares is beyond the scope of this paper, especially given journal length limits with *Journal of Institutional Economics.* We have rewritten the section discussing sectoral shares to clarify the "bad controls" discussion and that studying this channel is not our focus but would be interesting for future research. This text is copied below. (See text, copied from Section 3.4, above.)

---

[13]See, e.g., Colonnelli and Prem (2022) for a discussion of corruption uncovered in these audits.

Concerns Regarding Instrumental Variables

R1(1), R1(2), R3(3), R5(7), and R6(7) all expressed similar, yet unique concerns with the instrumental variable section. Most of these comments have been addressed in our rewriting of Section 3.6, copied below. But we wanted to give some specific responses here with portions of the comments that were not directly addressed in the paper.

3.6 Instrumental Variables

A concern with our OLS results, and any study of corruption, is endogeneity. There are many potential unobservable causes of corruption that could result in omitted variables biasing the estimates. Culture, for example, has been linked to corruption and is notoriously difficult to measure and control for (Barr and Serra, 2010; Pillay and Kluvers, 2024). An ideal solution would be to utilize an experimental or quasi-experimental design. However, given our cross-sectional data, this is not possible here. We therefore rely on an instrumental variable approach. More specifically, we utilize a two stage least squares (2SLS) estimator where we (1) get an estimate of corruption in the first stage using a set of instrumental variables (along with our controls) and (2) use this predicted corruption value to estimate the causal effect of corruption on our outcomes in the second stage.

For instruments to be valid, they need to satisfy two criteria: relevance and exogeneity. The first is relatively easy to satisfy in that many factors are related to corruption. It is the second, exogeneity, that makes finding a plausible instrument more difficult.

Our instruments include two measures of political competition and participation: the existence of local councils and whether these councils are active.[14] The former counts the number of municipal councils that exist and creates an index from this information (scaled from 1 (least councils) to 6 (most councils)). A council is coded as active if they have individuals appointed in positions. These measures are taken from a 1998 index (*Indicador de Qualidade Institucional Municipal* - IQIM) constructed by the *Instituto Brasileiro de Geografia e Estatística* (IBGE). Municipal councils serve as a check on corruption.

We additionally include a measure of management capacity, also from the IQIM index and defined at the municipal level. In sum, this indicator measures the government's ability to implement zones, codes, and other laws with

---

[14]In the context of Brazil, municipal councils (*conselhos municipais*) are commissions established by law to propose or advise on policy initiatives in a specific area (e.g., health, education) and oversee their implementation. These councils typically include representatives from the local public administration and civil society organizations. Notably, they differ from city councils (*câmaras municipais*), which serve as the local legislative branch.

the purpose of municipal planning (e.g., zoning laws or building codes). We interpret this as a measure of state capacity where state capacity is defined broadly as the ability to govern, enforce the law, and tax (Piano, 2019). While stronger states might engage in more corruption, they also have a stronger ability to limit it. There is an extensive literature connecting state capacity to development (see, e.g., (Johnson and Koyama, 2017)), and while specific the connection between corruption and capacity has received relatively less attention, it is likely that state strength is an important factor in determining corruption levels (Owen and Vu, 2022).[15]

Lastly, we include an indicator for whether the municipality is a judiciary district (*comarca*). It implies that the municipalities has a branch of the state court. Ferraz and Finan (2011) theorize that the presence of a judge increases the likelihood of being prosecuted for wrongdoing and thus likely reduces corruption as a result.

All four instruments are measured before any occurrence of corrupt activity studied in this paper.[16] This is beneficial because it makes reverse causality less of a concern. However, the length of time between instrument measurement (1998 for the IQIM data) and corrupt activity could be concerning (any time between 2003-2013). One might be worried that these instruments are not relevant at the time of the corrupt activity and therefore may not be strong predictors of corruption – in other words, these instruments are predetermined but could be weak. To address this concern, we always report the F-Statistic from the first stage to gauge the strength of the instruments. First stage results are reported in full in Online Appendix E.

Another concern with these instruments is that they are not truly exogenous. We note that all four variables are political instruments with a focus on implementing some sort of check and balance in local government. We argue that these checks and balances only influence our outcomes through corruption. We believe this is a reasonable assumption but cannot rule out other potential channels. We do provide the J-Statistic from a test where the null hypothesis is that the instruments are exogenous. However, this is not a particularly strong test because even if we fail to reject the null at a standard threshold (e.g., 90%) the probability that the null is false can still be reasonably high. We also provide appendix results (Appendix E) where we present "just-identified" 2SLS estimates using each instrument separately to show the consistency of our results. Even so, endogeneity could remain. We therefore view these instrumental variable results as a robustness check only and refrain from making strong causal statements throughout the paper.

---

[15]Defining the causal association between state capacity and corruption is not necessary in determining the relevance of an instrument. All that matters is that the two variables are correlated; and that the instrument is not otherwise associated with the outcome.

[16]Because there were 6 municipalities emancipating after 1998 but before 2010 we have only 929 observations for IV estimates. Summary statistics for instruments are presented in Table 3.

R3(3) questions why the IV strategy would only "work" for the full sample and the leadership category. We want to clarify that the goal of the IV analysis is to separate the exogenous from the endogenous variation of corruption and to exclusively focus on the former. If a result is not robust when using only the exogenous variation, and the instrument is strong, this implies that the initial results were biased. This does not suggest that the IV strategy did not "work". However, as we note in the paper, we caution the reader from interpreting any result in our paper as truly causal. Even if the IV result does not present a statistically significant effect of corruption, it may be that the instrument is not picking up enough exogenous variation to get precises estimates. It is likely that the IV strategy is most effective for the full sample and leadership categories because this sample has the largest sample size and thus the most precision.

Other Econometric Concerns

R5(2) this is an excellent point and would be a very interesting analysis but we reserve this for future research. First, our sample includes only a subset of audited municipalities that are (sometimes) not necessarily geographically close. Second, and more importantly, we are very limited in the word count given the journal's requirements. However, we emphasize that this would be an interesting analysis for future work. This is now discussed in the final paragraph of the paper (copied below).

> "We also believe that examining potential spillover effects of corruption on female labor market outcomes would be an interesting future analysis. It could be that women are escaping corrupt environments and moving to nearby, presumably less-corrupt municipalities. Thus, while corruption might reduce female leadership presence in one municipality, it increases it in others. Understanding the spatial dynamics of corruption would help shed light on the full effect of corruption on female leadership."

R3(7) questions why we do not include municipal fixed effects. We are looking at cross-sectional municipal data, and we focus on this cross-sectional variation within each lottery. We cannot include municipal fixed effects as this would eliminate all relevant variation since this is not a panel.

R6(9) We do not provide coefficients for the controls in the main tables given journal length limits. We also do not believe it would be beneficial to do so given concerns outlined in Hünermund and Louw (2025). However, we now provide full estimates for the main estimates in Online Appendix F for reference.

R6(10) It is not clear what benefit the proposed grouping (i.e., corrupt environment versus non-corrupt environment) would have over regression with controls in this context. It is also not clear how one might define a corrupt versus non-corrupt environment. Our corruption is measured on a continuum and this would simply reduce this variation down to a simple dichotomy. We understand that matching and regression each have their own set of benefits (and flaws), but we are unsure of what flaw that matching in this way might

improve our analysis. For example, matching still relies on the assumption that there are no important unobservable characteristics. There is a study that utilizes matching in the corruption context (Pavlik et al., 2023) but this study identifies a specific treatment and focuses on variation through time. We do not have time variation. However, we have added clarification throughout the text as well – much of which is copied throughout this response document – and hopefully this helps address this comment.

Terminology Concerns

R3(4) expressed concerns and confusion over the selection versus quantitative effect terminology and offered a suggestion to use "glass-ceiling" in place of "selection". We appreciate this suggestion and have made this change throughout. Please also note that the "glass-ceiling" effect is present in the non-corrupt sectors (Appendix D) as expected by the reviewer.

Other Concerns Regarding the Outcome Variables

R5(1) and **R5(2)** expressed concerns regarding the interpretation of the outcome variable. We agree that they are not directly comparable and have tried to clarify the discussion as to why all three outcome variables are necessary to understand the full picture. We hope that these revisions, along with the change in terminology mentioned above, helps satisfy these concerns.

Theory - Individual Choice or Firm Choice

In relation to **R6(1)** and **R6(2)**, we have clarified that the mechanism could come from individual choice or firm choice in the sense that firm owners could prefer a leader of a specific gender due to corruption. However, we also now highlight the importance of looking at the employer category to help separate these effects. This is also done at the suggestion of **R4(1)**.

Minor Errors, Typos, and Other Suggestions

R1(a), **R1(b)**, **R3(a)**, **R3(b)**, **R3(c)**, **R3(d)**, **R3(e)**, **R3(f)**, **R3(g)**, and **R5(8)** had minor comments regarding writing. We have addressed all of these.

R3(j) asks if we have data on the public sector. While this sector could theoretically be parsed out, we note that this is a very small share of overall employment in most

municipalities at less than 9 percent of total employment. This sector also wouldn't have much variation in the employer category, which is of primary interest in our analysis.

R2(1) requested that we engage more with recent articles published in the *Journal of Institutional Economics.* We have incorporated several recent articles including Nishioka et al. (2023); Kubbe et al. (2025); Hodgson (2025). We appreciate this suggestion.

References

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Avis, E., Ferraz, C., and Finan, F. (2018). Do government audits reduce corruption? estimating the impacts of exposing corrupt politicians. *Journal of Political Economy*, 126(5):1912–1964.

Barr, A. and Serra, D. (2010). Corruption and culture: An experimental analysis. *Journal of Public Economics*, 94(11-12):862–869. Conducts bribery experiments showing that undergraduate students from more corrupt countries are more likely to engage in bribery, suggesting corruption as a cultural phenomenon influenced by social norms.

Bologna, J. (2016). The effect of informal employment and corruption on income levels in brazil. *Journal of Comparative Economics*, 44(3):657–695.

Bologna, J. and Ross, A. (2015). Corruption and entrepreneurship: evidence from brazilian municipalities. *Public Choice*, 165:59–77.

Bologna Pavlik, J. (2018). Corruption: The good, the bad, and the uncertain. *Review of Development Economics*, 22(1):311–332.

Brollo, F. and Troiano, U. (2016). What happens when a woman wins an election? evidence from close races in brazil. *Journal of Development Economics*, 122:28–45.

Colonnelli, E. and Prem, M. (2022). Corruption and firms. *The Review of Economic Studies*, 89(2):695–732.

Dreher, A. and Schneider, F. (2010). Corruption and the shadow economy: an empirical analysis. *Public Choice*, 144:215–238.

Ferraz, C. and Finan, F. (2008). Exposing corrupt politicians: the effects of brazil's publicly released audits on electoral outcomes. *The Quarterly journal of economics*, 123(2):703–745.

Ferraz, C. and Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, 101(4):1274–1311.

Hodgson, G. M. (2025). Formal and informal institutions: some problems of meaning, impact, and interaction. *Journal of Institutional Economics*, 21:e1.

Hünermund, P. and Louw, B. (2025). On the nuisance of control variables in causal regression analysis. *Organizational Research Methods*, 28(1):138–151.

Johnson, N. D. and Koyama, M. (2017). States and economic growth: Capacity and constraints. *Explorations in Economic History*, 64:1–20.

Kubbe, I., Kırşanlı, F., and Nugroho, W. S. (2025). Corruption and informal practices in the middle east and north africa: a pooled cross-sectional analysis. *Journal of Institutional Economics*, 21:e24.

Malta, V., Kolovich, M. L. L., Martinez, A., and Tavares, M. M. M. (2019). *Informality and gender gaps going hand in hand*. International Monetary Fund.

Mishra, A. (2006). Persistence of corruption: some theoretical perspectives. *World Development*, 34(2):349–358.

Nishioka, S., Sharma, S., and Le, T. V. (2023). Political regimes and firms' decisions to pay bribes: theory and evidence from firm-level surveys. *Journal of Institutional Economics*, 19(6):764–786.

Owen, P. D. and Vu, T. V. (2022). State history and corruption. *Economics Letters*, 218:110774.

Pavlik, J. B., Grier, R. M., and Grier, K. B. (2023). Two birds with one stone: Reducing corruption raises national income. *Social Science Quarterly*, 104(4):406–422.

Piano, E. E. (2019). State capacity and public choice: a critical survey. *Public Choice*, 178(1):289–309.

Pillay, S. and Kluvers, R. (2024). The relationship between culture and corruption: A cross-national study. *Journal of Public Affairs*, 24(3):e2941.

1 Reviewer 1

Comments to the Author The authors rectify the existing literature by focusing on harmful effects of corruption on individuals (existing literature looks at effects of corruption on firms). Fighting corruption is one avenue for rectifying gender inequalities. Less corruption ÔÉ† more females in leadership positions (for a multitude of reasons). The authors use Brazilian municipalities to test whether variations in corruption have any relation to the share of female leadership in firms. I particularly like the distinction between the quantitative effect and the selection effect. I also really like the instrumental variables approach, using an older publication to rule out reverse causality. I have two concerns about that approach though: (1) since those variables are measured years before, is there any concern that they might not reflect the reality at the time your corruption variables were observed? Judiciary district and management capacity are probably relatively stable over time, but the existence of local councils and whether they are active could be more subject to change, and (2) might management capacity be endogenous to corruption? This is along the lines of the grease the wheels hypothesis ,Äì less functional areas might see more corruption just in the everyday course of business. There is literature that suggests an inverse relationship between corruption and state capacity, so maybe just citing that will overcome this particular qualm.

In looking at your summary statistics, I can't figure out why your share of female employers, managers/directors/executives, and women in leadership would have a min and max of 0 and 1. A better explanation of this would be warranted in the text, otherwise this stands out like a weird measurement error. Also, there are some municipalities have 97

Small points: Bottom of page 4, into page 5 ,Äì Those first two sentences are confusing. They seem to be saying the same thing, but suggesting there is a difference. I think you're trying to get at the selection vs. quantitative effects, but this is worded confusingly. Middle of page 5 ,Äì "Corruption has no effect on female leadership presence, but female leadership representation is much lower across the board in these corruption prone sectors." This is another case of weird wording. It sounds like the two parts of the sentence are contradicting each other.

2   Reviewer 2

Comments to the Author This manuscript makes a significant contribution to the Journal of Institutional Economics by skillfully examining how corruption–a key informal institution–shapes economic opportunities and organizational leadership across gender lines in Brazil. The paper aligns excellently with the journal's focus on understanding the role of institutions in economic growth by analyzing how corruption creates institutional barriers that can systematically exclude women from leadership positions. The authors bring a valuable institutional economics perspective to gender inequality research, demonstrating how corruption operates as an informal institution that influences firm behavior and leadership composition. This approach is particularly relevant to the journal's mission of studying the nature, role and evolution of institutions in the economy, including firms, as it reveals how corruption creates parallel institutional structures that may disadvantage certain groups. The paper's cross-disciplinary approach–bridging economics, gender studies, organizational behavior, and development economics–fits perfectly with the journal's multidisciplinary ethos. The authors have developed a rigorous empirical framework using municipal-level corruption data and detailed occupational information from census data, which allows them to make causal inferences about the relationship between corruption and female leadership presence. I particularly appreciate how the research distinguishes between quantitative effects (reducing female labor force participation) and selection effects (limiting female advancement to leadership positions). This nuanced analysis of institutional mechanisms helps us understand not just that corruption affects gender equality, but precisely how these effects manifest in different sectors and contexts. The focus on Brazil provides valuable insights into institutional dynamics in developing economies where corruption may function differently than in more developed contexts. The finding that corruption has different effects in corruption-prone sectors versus the broader economy adds important complexity to our understanding of how institutional barriers operate. This manuscript would be of great interest to the journal's broad readership interested in real world economic institutions and organizations. It advances our understanding of how informal institutions can perpetuate gender inequality in economic leadership, with implications for researchers studying institutional economics, gender in organizations, development economics, and labor markets. Despite the excellent fit with the journal's focus, one area for improvement is the limited engagement with recent related literature published in the Journal of Institutional Economics. The authors could strengthen their contribution by connecting their findings to the journal's ongoing scholarly conversation about informal institutions, governance, and economic outcomes. I recommend the authors incorporate several relevant recent publications from the journal:

Nishioka, Sharma, and Le's (2023) "Political Regimes and Firms' Decisions to Pay Bribes" would provide an excellent comparative perspective on how political systems influence corruption at the firm level, complementing this paper's municipal-level analysis. Belloc, Burdin, and Landini's (2023) "Corporate Hierarchies and Workplace Voice" examines institutional structures within firms that affect power dynamics–directly relevant to understanding how corruption might reinforce or alter these hierarchies along gender lines. Mahn, Poblete, Wang, and Heaton's (2024) "The Role of Culture as an Informal Institution in Cross-Border Venture Capital Investments" offers insights on how infor-

mal institutions like culture shape economic decisions, which could enrich the discussion of how corruption functions as an informal institution affecting gender outcomes. Arif and Dutta's (2024) "Legitimacy of Government and Governance" explores how government legitimacy affects governance quality, providing a valuable lens for understanding the institutional context of corruption in Brazilian municipalities. Aimar's (2023) work on "Destructive Entrepreneurship" could inform the analysis of how corruption diverts entrepreneurial activity, potentially explaining some of the observed effects on female employers in corruption-prone sectors.

By connecting their findings to this recent literature, the authors would not only strengthen their theoretical framework but also more explicitly position their contribution within the journal's ongoing scholarly conversation about institutional economics, enhancing the paper's impact and relevance for the journal's readership.

3 Reviewer 3

Comments to the Author Summary: This paper empirically estimates the effect of municipal corruption on the share of women in leadership positions in Brazil, using the 2010 Census and a corruption measure built from random audits of municipal expenditures. The paper finds that corruption does indeed reduce female representation in leadership positions but through a selection effect in corruption-prone sectors, by deterring female entrance in these sectors. The paper focuses on an interesting topic with good data and empirical strategy. However, I am not entirely convinced by the execution of the analysis (and by the IV results) so I further detail some suggestions below. Major comments: 1. Why do sample sizes differ? In Table 1, the number of municipalities for the share of females among employers is 878 but the number of municipalities for the share of females among managers, directors, or executives is 934 (and it is 930 in Table 2). The descriptive statistics and analysis have to be based on the same sample. 2. It would be good to know from the first sentence in Section 3.1 that the cross-sectional data come from the 2010 Census because otherwise the discussion on the years over which the corruption measure is calculated becomes difficult to follow. Additionally, the robustness done in Appendix A, which only considers corruption data before 2010 should be the main analysis (and not the robustness check). How can corruption cases that occur after 2010 influence female leadership representation in 2010? Otherwise, I would try to make a stronger argument for considering corruption cases after 2010 in the corruption measure (for instance, that it is quite persistent over time by showing statistics over time for municipalities that were audited several times and arguing that because lotteries only select some municipalities each year you have to use the full sample of lotteries, even those after 2010 ‚Äì if that's the case). 3. On the instrumental variable strategy: can you provide concrete examples of "unobservable causes of corruption that could result in omitted variables biasing the estimates"? Given the outcome variable is measured for year 2010, would it be possible to build instruments from later data (1998 is more than 10 years before and corruption has likely changed over a 10 year period)? The measure of management capacity is at State level and not at municipal level, correct? How are the instruments used? Do they form a unique indicator? It would be good to include them in the descriptive statistics and to see first stage results. Why would the IV strategy only work for the full sample and the leadership category? 4. Industry share controls: it would be good to see coefficients on the industry share controls, in particular to see whether they correspond to the "corrupted" sectors. Relatedly, it would be good to see results on the "non-corrupted" sample; if the "selection" and "quantitative" mechanisms are taking place as suggested, we should see a "selection" effect in the "non-corrupted" sample. [I am slightly confused with the use of the term "selection effect" here; I would rather call the "selection effect" the fact that women select out from corrupted sectors. Maybe the other effect can then be called a "glass-ceiling effect", given that corruption prevents women from reaching the higher ranks of the hierarchy?] Some z-tests on the proportion of females among employers and in management in "corrupted" versus "non-corrupted" samples might help the discussion. 5. Is it possible to see descriptive statistics on municipalities that were excluded from the analysis as compared to those that were included (in particular in terms of female leadership representation)? Are the large municipalities included in the anal-

ysis (those that were excluded from the lottery)? How do these municipalities differ, as compared to those included in the analysis? Why are there corruption scores missing for some municipalities? How do these municipalities differ, as compared to those included in the analysis? These descriptive statistics could give a sense of whether results on all municipalities are likely to be weaker or stronger. Additionally, it would be interesting to know why Avis et al. (2018) only use 967 municipalities, while 1949 municipalities experienced at least one audit (excluding the last two years of lotteries would reduce the number of municipalities by more than half?). 6. Is it the case that one lottery draws several municipalities to be audited, or only one? This would clarifies what the lottery fixed effects capture exactly. Related to this question, it is not clear to me why there is the lottery subscript to the corruption variable. 7. Why not using municipality fixed effects? Instead or on top of State fixed effects. Along this line, why not clustering standard errors by municipality? 8. Do you include a dummy variable that captures the fact that some municipalities were audited several times?

Minor comments: - Last paragraph of page 3 and first paragraph of page 4 are too detailed for the introduction. I would shorten them. However, I would extend the sentence on the corruption measure (using a first sentence along the lines of the first sentence of Section 3.2). - In the first paragraph of page 10 "we scale this number by population" and in footnote 8 "we additionally weight our data by population": what is the difference between these two? Is the population weighting done for the main analysis or for the robustness analysis? It is not clear to me what is the difference between results in Table 2, nPanel A and results in Table A1, Panel A. I would use the population weighting in the main analysis and maybe, you can present a separate analysis for large and small municipalities in the robustness checks? - It would be good to have more details on the corruption measure from Ferraz and Finan (2011) and how it differs with the main corruption measure used in the analysis. - Could you provide a discussion on how likely the lottery is itself likely to be corrupted? Descriptive statistics on the municipalities that were audited (and those that were audited several times), as compared to the municipalities that were never audited could help the discussion. - It would be good to have the non-logged (per capita and absolute level) figures for the corruption measure in the descriptive statistics. - It would be good to add the significance levels in tables' notes. - Footnote 10 should appear earlier on in the paper. - Please explain what you mean by "bad controls à la Angrist and Pischke (2009)". - It would be good to see robustness results on the "corrupted" sample as well. - Do you have data on the public sector? To look at the female share in leadership positions in municipalities.

4 Reviewer 4

1. The first outcome variable is the ratio of women in leadership roles out of all leadership positions, while the second is the ratio of women in the labor force out of all working-age women. These two outcomes are not directly comparable because the denominators are inconsistent. The first one shows women's leadership compared to both men and women in leadership, but the second one shows labor force participation relative only to women. So, when interpreting results, saying that corruption reduces both women's leadership and labor force participation can be misleading, since one is relative to men and the other is not. To make them comparable, you could redefine the second outcome as women's labor force participation rate divided by the male labor force participation rate ((women in labor force/working age women)/(men in labor force/working age men)). That way, if you argue that corruption reduces both participation and leadership, it would imply that men are taking more of the jobs and that in corrupt sectors, men may find it easier to navigate and climb to leadership roles. To make this claim strong, you'll need to control for working-age population in the analysis. 2. The third outcome variable actually strengthens the interpretation. Even if women's overall labor force participation doesn't change–or even increases–corruption can still be a barrier specifically to attaining leadership. This helps justify that in corrupt environments, even when women manage to get jobs, it's much harder for them to reach top positions. 3. The paper could better explain the rationale behind the choice of control variables. Please cite some relevant literature to support the selection. For instance, why is the informal sector included as a control? A possible logic is that informal sectors tend to be more prone to corruption, and if you don't control for this, you might miss capturing the true effect of corruption on the outcomes. 4. Sectoral shares are not included in the baseline regressions, but a sub-sample is used for specific "corrupt" sectors later. Please justify why a separate sub-sample makes sense here. It's likely because the regressions are at the municipality level rather than at the individual level, and sectoral composition differs across municipalities. 5. How did you conclude that the sectors are corrupt? Explain the methodology and include some summary statistics comparing corruption levels in these selected sectors (extractive, manufacturing, construction, transportation/communication). 6. In the sentence: "It is quite telling that while the female labor force participation in these sectors is almost one-tenth of the overall economy (3.8of leadership positions held by women is ‚Äòonly' three-fifths of that for the full sample (19.7interesting but could use more explanation. What exactly is "telling"? Is it that leadership representation is relatively higher than participation, or that both are disproportionately low? 7. There should be more explanation for why each of the IVs satisfies the exogeneity condition. Right now, it's mentioned that they are measured before corruption, which helps with reverse causality. Please explain the theoretical link between these IVs and the outcomes, and cite supporting literature to justify why they affect corruption but not female leadership directly. 8. This sentence is confusing: "i.e., employees that are in a manager, general director, or executive role and employers - that are occupied by women." Reword for clarity.

5   Reviewer 5

This paper explores the possibility of corruption creating a barrier to entry in female leadership roles. The authors adopt a measure of corruption per capita in Brazilian municipalities from Avis et al. (2018) and regress against several measures of female leadership - the portion of women leaders (employers and employees) among all leadership roles. Results reveal that corruption per capita and female leadership generally relate negatively (and statistically significantly so in several cases). I think the study poses interesting questions, and the results tend to support the authors priors. I currently have only a couple comments the authors might consider. 1. I think the authors undersell the potential importance of the employers variable. In the women-as-employee situation, it may be that men navigate the red tape to start the business that employs leaders. In this case, it seems women's exposure to corruption affecting their placement in leadership roles as an employee is somewhat limited to whoever they might interact with in that firm to acquire a leadership role. However, employers potentially face the municipality level bureaucratic process of starting a business - which plausibly opens them up to greater corruption exposure. It would be interesting to explore that, separate from the rest. That is, inspect the corruption effect on women employers as a percent of total employers. Currently, it appears the authors are looking at women employers as a percent of all leadership roles or aggregate leadership roles held by women. (If the authors are currently regressing on women employers as a percent of all employer, then that should probably be stated clearly). 2. To address other possibilities, spatial regression analysis might be valuable. Women pushed out of employer and other leadership roles because of corruption might search nearby for less corrupt environments to apply their efforts. This might shed light on the full effect of corruption - does it squash these leadership (and entrepreneurial) opportunities for women or reallocate them elsewhere? Otherwise, I think the study is a nice fit for JOIE, and recommend it for publication, with only minor revision

6   Reviewer 6

The authors aim to analyze the effect of corruption on female leadership in Brazil. Based on crosssectional municipal data, they conclude that corruption is the primary factor deterring women from pursuing managerial careers. I found the paper interesting. Nevertheless, I have several concerns and suggestions. Main Comments: 1) The authors should carefully discuss whether the low presence of females in the leadership of firms (or regions) associated with higher corruption is a result of their own choice or a firm-level choice. It has been noted that corruption is often associated with men's activities; therefore, owners may prefer a corrupt CEO in a corrupt environment. 2) The authors cite the literature studying the effect of corruption on female-led firms, criticizing that the focus of this literature is on the firm rather than on the individuals, and highlight additional firm-level effects, such as women being more likely to hire other women. Nevertheless, while these studies employed sizeable firm-level data with existing cross-sectional and time variation, the causal link between female participation and corruption is hard to identify. Existing corruption literature tends to agree that it is driven by firm owners who prefer "somebody" able to act in corrupt circles. In terms of political representation, the situation could be different. 3) The model, methodology, and data structure are not adequately described. Initially, I was surprised that the primary model does not include the time dimension, because the authors later describe the differences that other authors used with similar data across different periods. Given the FE, I presumed there was some cross-sectional time variation that the authors could use. I realized later that the primary data sample is the cross-sectional data coming from the 2010 census. It must be described clearly; maybe first, the authors should explain the data, the sampling frame they used, any aggregation, etc. Then their methodology would not be misleading. 4) What is the role of audits? Are all of them related to tax collection or the use of public funds? It is unclear. 5) Is there any information about the size and ownership type of the firms used to comprise the sample? It also affects female participation. 6) I am reluctant to buy the argument that the sectoral share of females in leadership is driven by corruption. The authors should present and, if applicable, utilize the share of females employed in the sector and participating in leadership roles. 7) Some measures used as instruments (e.g., an indicator of a juridical district) may be associated with municipality size, as well as with the number of economically active individuals. 8) Table 1, which describes the data and ratios, should be more clearly defined. What is the basis for getting different percentages for leadership positions? What are the bases in each row? Starting number of observations and the aggregation method. Additionally, when one considers the standard deviations of each mean, it suggests that there is no difference between the "corrupted" and non-corrupted sectors. Also, which ratio is eventually used for the regression analysis? The share of labor force participation is way too low if they use the somewhat recalculated ratio; it should be clearly defined and explained with a different number of observations. 9) The regression tables should include a list of all controls, with details provided in the Appendix. 10) To control for the size and possible endogeneity concerns, the authors should employ some variant of differences, using those in a "corrupt" environment at the treatment group and those in a non-corrupt environment as the control group. Perform matching based on the size of the municipality, population size,

age, and education structure, and compute the ATET (average treatment effect on the treated) while evaluating the quality of the matching. Given an inadequately described methodology and data construction, one may be curious whether the observed effect is not driven by selection and sampling issues. Therefore, the authors should be careful when describing the data, sample construction, and variable definitions.