INSTITUTO DE ECONOMÍA
FACULTAD DE CIENCIAS ECONÓMICAS
Y ADMINISTRATIVAS

Lecture : Econometric Theory II
Professor : Tomás Rau TA : Valle and Barrueto
Date : June 28, 2023
Student : Valentina Andrade
email : vandrade@uc.cl

Report 4

Causal inference

Question 1

a. Summary "Engaging Teachers with Technology Increased Achievement, Bypassing Teachers Did Not"

Summary:

The paper titled "Engaging Teachers with Technology Increased Achievement, Bypassing Teachers Did Not" examines the impact of technology integration in government middle schools on student achievement. The study focuses on the implementation of eLearn, a government program aimed at enhancing student learning in math and science through brief videos of expert content. Two implementation models have been tested: eLearn Classroom, which encouraged teachers to become more effective and provided students with a more engaging learning environment, and eLearn Tablets, which bypassed teachers and promoted independent learning.

Interventions:

The interventions revolved around introducing technology into the training process through video lectures. In the eLearn Classroom intervention, teachers and students in the 8th grade were exposed to video lectures that could be used to prepare lessons or projected in the classroom. On the other hand, the eLearn Tablets intervention provided 6th-grade students with tablets containing lecture content, but teachers were not actively involved in incorporating the technology.

Results:

The two interventions yielded opposite effects. The eLearn Classroom intervention resulted in a significant increase in student achievement, with scores improving by 0.30 standard deviations (SD). In contrast, the eLearn Tablets intervention led to a decline in student achievement, with scores decreasing by approximately 0.43 SD on combined math and science exams. The gain observed in the Classroom intervention accounted for 60% more improvement compared to the control group, while the decrease in the Tablets intervention represented 95% of the learning that occurred in the control group. Furthermore, the Classroom intervention increased the likelihood of students passing the standardized grade 8 test by 5 percentage points. Students and teachers in Classroom treatment schools also exhibited higher attendance rates and increased effort, unlike those in the Tablets intervention. The study also found that interventions providing materials or infrastructure without training or monitoring support, such as eLearn Tablets, did not yield improvements in student test scores.

Contributions:

- The study challenges the implicit assumption found in many previous studies, as well as the one associated with the Tablets intervention, that existing teachers must be bypassed or undergo extensive re-training and monitoring in order to enhance student learning. By demonstrating the effectiveness of combining external expertise with existing teachers, as seen in the Classroom intervention, the paper emphasizes the importance of integrating changes with teacher monitoring. This finding has implications for education technology initiatives, highlighting the potential benefits of leveraging the existing expertise of teachers while incorporating external resources.
- The paper provides additional evidence on the complementarity of different elements within educational interventions and sheds light on the marginal rate of technical substitution for technology-based interventions. The contrasting effects of the Classroom and Tablets interventions offer insights into how different approaches to technology integration can impact student achievement. The Tablets intervention, which lacked teacher involvement and guidance, resulted in decreased scores, indicating that simply increasing technology exposure without teacher support may displace other valuable learning activities. This finding emphasizes the need to carefully consider the role of teachers and the instructional design when implementing technology-based interventions.

- The interventions examined in the study were designed and implemented by the provincial government of Punjab, highlighting the importance of directly addressing the specific educational issues identified by the government. By involving the government in the program design and implementation, the paper demonstrates a collaborative approach that increases the likelihood of developing effective interventions tailored to the local context. This contributes to the potential for scalability, as successful programs designed in collaboration with government entities can be more easily adopted and implemented on a larger scale.
- The study focuses on middle school education, specifically grade-level content. By demonstrating the efficacy of grade-level solutions that engage teachers, even when dealing with more complex subject matter such as middle school, the paper provides valuable insights into improving grade-level competencies. This finding suggests that targeted interventions that actively involve teachers can effectively enhance student learning outcomes at specific grade levels. Understanding the importance of context-specific interventions can inform future educational initiatives aimed at addressing the unique challenges associated with middle school education.

In summary, the four contributions highlighted in the paper challenge prevailing assumptions, provide evidence on complementarity in educational interventions, emphasize the role of government collaboration for scalability, and highlight the potential for grade-level solutions to improve competencies. These contributions contribute to the broader understanding of effective strategies for enhancing student achievement through technology integration and teacher engagement in the educational process.

Empirical Strategy:

To address the issue of biased estimates stemming from non-random allocation of resources, the authors employed parallel randomized controlled trials for their interventions. By conducting these trials, the study was able to estimate the causal effects of the interventions by comparing outcomes between treatment and control groups. This approach helped mitigate potential confounding factors and establish a clearer causal relationship between the interventions and student achievement, providing more robust empirical evidence. Formally,

$$y_{is} = \alpha + \beta \text{ itreatment }_s + X'_{is}\Gamma + \varepsilon_{is}$$

and we know that $E(\varepsilon_{is}|treatment_s) = 0$

b. Table 1

Table 1 is presented with the purpose of evaluating whether there is any imbalance in the characteristics of students (Panel A), teachers (Panel B), and schools (Panel C). The examined characteristics are potential confounding factors that could bias the effect of the treatment (introduction of technology in learning). For example, the study investigates whether students in the treatment and control groups have parents with no formal education. The authors argue that this could result in less control over the tablets given to the children, potentially leading to a more negative effect.

Overall, as shown in both interventions, there are no significant differences between the treatment and control groups regarding the presented characteristics.

However, for student characteristics, significant differences can be observed in the number of days absent from class, with the treatment group in the eLearn Classroom intervention having a higher average absenteeism. As the table demonstrates, on average, the treated group has 0.3 more days of absenteeism per month.

Regarding teachers' characteristics, no significant differences are found between the treatment and control groups in both interventions. The considered characteristics include having an advanced degree, years of experience, minutes per day spent planning class sessions, use of technology for lesson preparation, and use of technology in the classroom.

In terms of school characteristics, significant differences are observed only in the eLearn Classroom intervention. Specifically, the control group has, on average, more computer labs than the treatment group. This is because out of the 30 schools in the intervention, three do not have computer labs. However, when examining other characteristics such as the number of sections per course and enrollment, no significant differences are found.

c. Internal Validity threat: Attrition

The internal validity threats of attrition refer to potential biases and threats to the internal validity of a study that may arise due to the loss of participants or missing data over the course of the research. Attrition occurs when participants drop out or are lost to follow-up, leading to an incomplete dataset.

There are several internal validity threats associated with attrition:

1. Selection Bias: Attrition can introduce selection bias if the participants who drop out of the study are systematically different from those who remain. If the characteristics or behaviors of the participants who drop out are related to the outcome being measured, it can lead to biased estimates of treatment effects.

2. Nonresponse Bias: When a subset of participants fails to respond to certain measurements or follow-up assessments, nonresponse bias can occur. If nonresponse is related to the variables being measured or the treatment group assignment, it can distort the results and introduce bias.

3. Differential Attrition: Differential attrition occurs when the rate of dropout differs between the treatment and control groups. If the attrition rates are significantly different and related to the outcome or treatment, it can lead to biased estimates of treatment effects.

4. Missing Data: Missing data resulting from attrition can create bias if the missingness is related to the outcome variable or other relevant variables.

The authors address the issue of attrition in their study. On page 23, they mention that they cannot access 15% of their baseline sample at the end of the study, leading them to conduct various tests to assess whether this affects the treatment effect. The authors perform two exercises to account for different types of attrition:

1. Differential Attrition:

They estimate a linear probability model to examine whether the probability of response differs based on treatment status. As shown, the treatment group has a 4-percentage point higher likelihood of providing follow-up data. However, they demonstrate that while these attrition differences exist, they are not related to student outcomes, thereby mitigating potential selection biases among only those who respond (Table 3, Column 2).

2. Nonresponse bias:

A second potential source of attrition is the number of students taking the PEC admission exam. As indicated in Column 3 and 4 of Table 3, the attrition differentials do not yield distinct effects.

It is important to note that this analysis is conducted for both the eLearn Classroom and eLearn Tablet interventions, although it is primarily discussed in relation to the eLearn Classroom intervention. By addressing attrition and evaluating its potential impact on the treatment effect, the authors strengthen the internal validity of their study. Additionally, it is important to highlight that the authors employ the methodology of Lee bounds (2009) to calculate upper and lower bounds of the treatment effects, ensuring that their results are robust to attrition. By using this approach, the authors account for the potential bias introduced by attrition and provide a range of treatment effects that account for this uncertainty. This strengthens the validity of their findings and allows for a more comprehensive understanding of the impact of the interventions despite the attrition observed in the study.

d. Other internal validity threats

Internal validity threats related to ITT (Intention-to-Treat) and SUTVA (Stable Unit Treatment Value Assumption) are two considerations in this research to evaluate treatment effects. First we will review some points of ITT and SUTVA.

Internal Validity Threats of ITT:

The ITT principle refers to the analysis of participants in the treatment groups based on their original assigned treatment, regardless of their adherence or compliance with the assigned treatment. The primary goal of ITT is to maintain the integrity of randomization and preserve the balance between treatment and control groups.

Internal validity threats associated with ITT include:

- Non-compliance: Some participants may not fully adhere to the assigned treatment or may switch treatments during the study. This can introduce bias and impact the estimated treatment effect if there are significant differences in the treatment received.
  - This is especially possible in interventions where there is no control over what students do or don't do, such as the intervention that solely involves the tablet. In this case, we do not have a report on the rates of use specifically for academic purposes in the use of tablets, and instead, we only know that students used them until the end of the period, but we don't know for what purpose. In that sense, the intervention is not strictly applied rigorously because the intervention is to participate in eLearn Tablet, not just "receive a tablet."
- Crossover: Crossover occurs when participants in the control group receive the treatment, or vice versa, during the study. This can complicate the interpretation of treatment effects and may introduce biases if there are differential effects between early and late treatment.
  - The paper does not discuss the rate of school turnover, so it is possible that students switch from one school to another and end up being part of the treatment or control group, even if they were not initially assigned to it. In that sense, the change of schools can be a potential issue that should not be overlooked.

Internal Validity Threats of SUTVA:

SUTVA assumes that the treatment assigned to one participant does not affect the outcomes of other participants. It also assumes that there is no interference or spillover effects between different units receiving the treatment.

Internal validity threats associated with SUTVA includes:

- Contamination: Contamination occurs when participants in the control group are inadvertently exposed to the treatment, or vice versa. This can arise in situations where participants interact or share information, leading to potential spillover effects that violate the assumption of SUTVA.
  - It is difficult to think of externalities between schools due to the introduction of the intervention. One possibility that comes to mind is that, since the intervention is known in the selected schools, families may choose to transfer their children to other schools to access the interventions. However, this issue is not addressed in the article, nor is there information about school turnover.
- Diffusion of treatment effects: If the treatment has indirect effects on non-treated individuals through social or environmental channels, it can undermine the assumption of SUTVA. This can occur when individuals in the control group benefit from the treatment received by others, leading to biased estimates of the treatment effect.

– While it is difficult for results to spread between schools, it is important not to forget that the intervention is implemented at the student level. Therefore, we cannot dismiss the potential effect that the tablet could have on other household members who attend different schools. Although it may seem far-fetched, similar to the argument regarding school turnover, the authors should report whether students have siblings and if they attend the same schools, considering the possibility of contagion among children.

e. Lasso, why?

We know that while ordinary least squares allows us to estimate unknown population parameters between two or more variables, in practice, we may encounter limitations when we have many variables available to include in our regression. This could result in an overfitted model, which translates into low predictive capacity. On the other hand, including explanatory variables that do not provide relevant information can lead to models with unnecessary complexity.

Therefore, the authors use the LASSO method to specify the optimal number of controls to be included in the analysis. These methods, known as reduction, shrinkage, or regularization methods, involve adjusting the model by including all explanatory variables but applying a penalty that forces the slope parameter estimates to approach zero. Specifically, the coefficients are derived by minimizing the sum of squared residuals plus a contraction penalty, such that:

$$\text{Lasso: } \min \left\{ SRC + \lambda \sum_{j=1}^{k} |b_j| \right\}$$

Where λ is a parameter that controls the relative impact of these two terms. For example, when $\lambda = 0$, the penalty has no impact, and both approaches produce exactly the same estimation as Ordinary Least Squares. In contrast, as the value of λ increases, the penalty becomes stronger, resulting in a reduction of the estimates towards zero.

This reduction in estimates can significantly reduce their variance, and in the case of LASSO, it may even force some estimated coefficients to be exactly equal to zero, allowing for variable selection.

An important technical point is that although the authors show that the coefficients are similar, there is a reason for this. Unlike OLS, these methods produce different estimates for different values of the λ parameter, so the selection of its value is crucial. To determine the optimal value, a grid of values is considered, and as seen in their implementation algorithm, a value is chosen based on some cross-validation method.

f. Replicate panel A and B, column 1 a d 3 and explain results

	Project (1)	Combined project (2)
Treatment school	0.256*	0.269**
	0.135	0.119
Observations	2622	2463
R^2	.1325435	.2022723

Table 1: Achievement Effects- Panel A

	Project (1)	Combined prooject (2)
Treatment school	0.270**	0.269**
	0.136	0.117
Observations	2622	2463
R^2		

Table 2: Achievement Effects- Panel B

Tables 1 and 2 show the replication of Table 2 from the article, which is located on page 54. As we can see, the replicated results are quite similar, except for the coefficient in column three estimated by LASSO.

The table displays the effect of the treatment, in this case, the eLearn Classroom intervention, on student achievement. The first column in our table presents the results for the project, and the second column (analogous to (3) in the paper) shows the measure constructed between the Project and PEC.

As we can see, the treatment has a positive and significant effect on academic achievement, with a magnitude of 0.256 standard deviations, holding all other variables constant (column 1). Similarly, for the combined measure, the intervention has a positive and statistically significant effect, with an effect size of 0.269 standard deviations (column 2).

The results remain consistent in terms of direction and significance for the estimation with optimal control reduction (LASSO), except that the effect size is slightly larger in the case of the Project, with 0.296 standard deviations.

Question 2

a. SummaryPopulation and Conflict

Summary

Population and Conflict by Acemoglu, Fergusson, and Johnson is an important study published in The Review of Economic Studies. It explores the relationship between population dynamics (1940-1980) and conflict, examining how changes in population size can affect the likelihood and intensity of violent conflicts.

The authors present a theoretical framework and empirical evidence to analyze this relationship. They argue that population changes can lead to conflicts through various channels, including resource scarcity.

Conclusions and Contributions

In general, although population has not been a prime focus in the economics of conflict literature, this paper addresses this gap by examining the relationship between population dynamics and conflict. By considering population as a significant factor in conflict analysis, the study expands the understanding of the underlying factors contributing to violent conflicts, contributing to a more comprehensive perspective on the economics of conflict. Also we can enumerate the following contributions

1. The paper incorporates both cross-sectional analysis and within-country analysis in Mexico. By utilizing these different analytical approaches, the study provides a broader understanding of the population-conflict relationship across different contexts. This comparative analysis enhances the robustness of the findings and allows for insights into how population dynamics impact conflict outcomes at both the macro and micro levels.

2. The study contributes to the literature on malaria and the effectiveness of anti-malaria campaigns from a health perspective. While previous research in this field has mainly focused on the health implications of malaria and the effectiveness of interventions, this study expands the analysis by considering the potential impact of population dynamics and health on social conflict.

3. In addition, the study addresses concerns raised in the literature regarding the potential bias and reverse causality in previous studies that have examined the correlates of civil war outbreaks and persistence. By employing credible econometric methods and utilizing instrumental variables, the paper tackles these issues and enhances the credibility of the findings. This contributes to the methodological rigor of the literature on conflict and establishes a more robust causal relationship between population dynamics and conflict outcomes.

Overall, the paper's contributions lie in its examination of population dynamics within the economics of conflict literature, its utilization of diverse analytical approaches, its expansion into the health dimension of conflict research, and its methodological advancements to address potential biases.

Endogeneity, assumptions and strategy

In the paper, the authors address the endogeneity problem that arises when estimating the effect of population growth on social conflict. They acknowledge that ordinary least squares (OLS) estimates can be biased downwards due to the potential correlation between population changes and other economic and institutional factors that also affect conflict levels. To overcome this endogeneity issue, the authors employ an identification strategy that exploits exogenous variation in population growth.

The identification strategy relies on utilizing a predicted mortality instrument, which captures the variation in population growth arising from the initial distribution of mortality due to infectious diseases. The authors focus on the period between 1940 and 1980 when global health technology and conditions experienced significant improvements. They argue that most medical and public health breakthroughs during this period were primarily driven by a few industrialized countries and can be considered exogenous to the development prospects and likelihood of internal conflict in the rest of the world.

By using the predicted mortality instrument, the authors can establish a causal link between population growth and social conflict. They emphasize that their identification strategy does not depend on the timing of when a particular country adopted better public health measures or how effectively these measures were implemented. Additionally, they find no differential trends in life expectancy, population, conflict, or economic outcomes between countries with initially high and low mortality rates before the onset of the international epidemiological transition.

In addition to the cross-country analysis, the authors apply a similar identification strategy focused on the role of malaria. They utilize the international epidemiological transition as a source of exogenous variation in health conditions and life expectancy across various countries.

Overall, the authors' identification strategy overcomes the endogeneity problem by exploiting exogenous variation in population growth driven by the initial distribution of mortality rates. This approach allows them to establish a more robust causal relationship between population dynamics and social conflict outcomes.

b. Predicted mortality instrument

The instrument

In the study, the authors construct an instrument called the "predicted mortality instrument" using the identification strategy based on the international epidemiological transition. This transition led to significant improvements in health conditions and life expectancy across various countries. Key innovations in drugs (such as penicillin and other antibiotics), chemicals (like DDT), and global public health campaigns implemented by organizations like the World Health Organization played a vital role in this transition. These interventions had a greater impact on countries where infectious diseases were more prevalent initially.

To construct the instrument at the global level, the authors adopt the "predicted mortality instrument" developed by Acemoglu and Johnson (2007). This instrument measures the cumulative mortality rate from infectious diseases for each country until the occurrence of a global intervention (innovation or campaign). After the intervention, the mortality rate of the specific disease declines to the lowest or "frontier" mortality rate observed in the study's sample.

$$M_{it}^{I} = \sum_{d \in D} \left((1 - I_{dt}) \, M_{di40} + I_{dt} M_{dFt} \right)$$

where Mdi40 represents the mortality rate in 1940 (measured as the number of deaths per 100 individuals per year) due to disease d ∈ D for country i. Idt is a dummy variable that takes the value of 1 for all dates following the intervention for disease d. MdFt represents the mortality rate of disease d at the global health frontier at time t (assumed to be zero mortality, as in Acemoglu and Johnson, 2007). D refers to the set of 13 diseases considered in the analysis.

To summarize, the "predicted mortality instrument" is constructed by considering the initial mortality rates, global interventions for specific diseases, and the frontier mortality rate in global health. This instrument captures exogenous variation in health conditions and is used to identify the causal effect of population changes on social conflicts.

IV assumptions

The authors make several key assumptions to ensure that the necessary conditions for instrumental variables (IV) estimation are met:

1. Instrument relevance: The authors justify the relevance of their instrument, the predicted mortality instrument (MI it), by explaining that it captures the exogenous variation in health conditions resulting from global interventions for specific diseases. They argue that countries with higher mortality rates from infectious diseases in 1940 are expected to experience larger increases in population after the interventions. This assumption implies that the instrument is correlated with the endogenous variable (population growth), which is necessary for the IV strategy.

2. Exclusion restriction: The exclusion restriction assumption states that the instrument affects the outcome variable (conflict) only through its impact on the endogenous variable (population), and not through any other pathways. The authors argue that the unique channel through which the predicted mortality instrument affects conflict is changes in population. They also mention conducting falsification exercises to test the correlation between predicted mortality and various pre-existing determinants of economic growth and conflict, suggesting that the instrument is not capturing other factors.

3. No correlation between instrument and error term: The assumption of no correlation between the instrument (MI_{it}) and the error term (ε_{it}) is crucial for unbiased estimates. The authors state the requirement $Cov(MI_{it}, \varepsilon_{it})$ =0, which indicates that the instrument is uncorrelated with the unobserved factors that might influence conflict independently of population changes.

4. Control variables: To address concerns about potential alternative channels through which changes in predicted mortality may affect conflict, the authors include additional control variables (Z_{it}) in their estimation model. These control variables aim to account for factors such as overall health conditions, age composition, and education (**in mexican analysis**)

Overall, the authors carefully consider the assumptions necessary for instrumental variables estimation. They provide justifications and evidence to support these assumptions, such as the instrument's relevance, the exclusion restriction, and the lack of correlation between the instrument and the error term. By addressing these concerns, they strengthen the validity of their IV estimates and enhance the robustness of their findings.

c. Falsification exercises

In the paper, the authors conduct a falsification exercise to strengthen the causal interpretation of their findings regarding the relationship between the international epidemiological transition, population dynamics, and conflict. The purpose of this exercise is to test whether there is any correlation between changes in predicted mortality and pre-1940 changes in population or conflict, which would undermine the causal link established through their identification strategy.

To perform the falsification exercise, the authors estimate a variant of their main equation (equation 4.7) in columns 5-8 of Table 3, where the time period considered is from 1900 to 1940 (t1 = 1940 and t0 = 1900). If the causal interpretation is correct, there should be no significant correlation between changes in predicted mortality and pre-1940 changes in population or conflict.

The authors report that the results of this falsification exercise are reassuring. They find no significant correlation between changes in predicted mortality and pre-1940 changes in population or conflict. The coefficient estimates are both insignificant and very small compared to the reduced form estimates. This lack of correlation indicates that the predicted mortality instrument is not capturing some other underlying economic or political trends that could potentially confound the results.

By conducting the falsification exercise, the authors provide additional confirmation that there were no preexisting trends in population or conflicts related to changes in predicted mortality. This strengthens their confidence in the causal interpretation of their findings and reinforces the validity of the predicted mortality instrument as a robust identifier for the impact of the international epidemiological transition on population dynamics and conflict.

The scatter plots shown in Figure 1 further support the findings of the falsification exercise. These plots illustrate the patterns observed in Table 3, including the first stages, reduced form estimates, and the falsification exercise. By demonstrating that the patterns are not driven by outliers, the scatter plots provide visual evidence that reinforces the robustness of the results obtained in the falsification exercise and the main analysis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Change in predicted mortality, 1940-1980	-0.782	-0.764	-0.491	-0.660	0.085	0.197	-0.189	-0.198
	(0.141)	(0.191)	(0.179)	(0.236)	(0.055)	(0.126)	(0.138)	(0.196)
Observations	51	40	52	41	36	28	52	41
R-squared	0.306	0.238	0.166	0.200	0.012	0.039	0.033	0.029
Robust standard errors in parentheses								

Table 3: First stages, reduced forms and falsification exercises

d. Table 4

	(1)	(2)	(3)	(4)
	propconflictCOW2	propconflictU	propconflictFL	logdeathpop40U
log of population	0.617	0.576	0.879	1.347
	(0.213)	(0.238)	(0.303)	(0.598)
Observations	102	104	104	104
Number of clusters	50	51	51	51
Weak exogeneity (Hausman p-value test) =	0.0128/0.0632	0.1258/0.1054	0.0290/0.0209	0.0733/0.2003
Weak identification test	(Cragg-Donald Wald F statistic):	21.582/22.533		
	(Kleibergen-Paap rk Wald F statistic)	31.217/32.506		
Stock-Yogo weak ID test critical values:	10% maximal IV size	16.38		
	15% maximal IV size	8.96		
	20% maximal IV size	6.66		
	25% maximal IV size	5.53		

Table 4: The effect of population on conflict: 2SLS estimates. Panel A: Long differences, just 1940s and 1980s. Robust standard errors in parentheses

	(1)	(2)	(3)	(4)
	propconflictCOW2	propconflictU	propconflictFL	logdeathpop40U
log of population	0.609	0.304	0.873	1.106
	(0.205)	(0.250)	(0.461)	(0.454)
Observations	307	308	308	273
Number of clusters	63	63	63	54
Weak exogeneity (Hausman p value test) =	0.5387/0.0506	0.9977/ 0.9713	0.1360/0.1595	-5.87/ 0.3255
Weak identification test (Cragg-Donald Wald F statistic)	32.756	34.388	34.388	31.823
Stock-Yogo weak ID test critical values:	10% maximal IV size	16.38		
	15% maximal IV size	8.96		
	20% maximal IV size	6.66		
	25% maximal IV size	5.53		

Table 5: The effect of population on conflict: 2SLS estimates. Panel B: Panel regressions, 1940s–1980s. Robust standard errors in parentheses

Test

In the case of these models we cannot perform an over-identification test because we only have one instrument for one endogenous variable, i.e. we are equally identified.

Now let us look at the Durbin-Wu-Hausman test for weak exogeneity, the null hypothesis (H_0) establishes that there is no endogeneity in the model, that is, that there is no correlation between the model errors and the explanatory variables. More specifically, the null hypothesis $H_0 : E(x'u) = 0$ implies that the covariance between the explanatory variables (x) and the model errors (u) is zero, indicating that the explanatory variables are truly exogenous and uncorrelated with the errors. If the p-value associated with the Durbin-Wu-Hausman test statistic is less than a predefined significance level (e.g., 0.05), the null hypothesis is rejected. This indicates that there is sufficient evidence to conclude that there is endogeneity in the model, i.e., the explanatory variables are correlated with the errors and weak exogeneity is not satisfied.

As we can notice in both panel A and panel B two values appear. The first one corresponds to the test when we have non-robust standard errors (do file problem3-hausman.do) while the second value corresponds to the test when we have robust standard errors (endog() specification in do file problem3.do). We report these two results even though they are similar because of their differences in rejecting the null hypothesis in model 4 of panel A and model 1 of panel B. We will now analyze these results considering the robust result.

As we can see the text gives us statistically significant evidence that there is no endogeneity of age in models 2 and 4 in panel A (recall that the null hypothesis tells us that there is evidence that OLS and IV are not different, so we have no evidence of endogeneity at least for this test). Similarly, the null hypothesis cannot be rejected for all models in panel B.

Let us now review the Cragg-Donald statistic. The null hypothesis in the Cragg and Donald (1993) test for detecting weak instruments is that there are no weak instruments in the instrumental variables model used. In other words, the null hypothesis assumes that the instruments used in the model are valid and have sufficient capacity to explain the variation in the endogenous variable, without introducing biases in the model results. This test should be interpreted jointly with the critical values provided by Stock and Yogo (2005), in particular for the table on significance and not on bias, since we are in the case of being exactly identified.

Then, Stock and Yogo will indicate different tolerance levels for significance in the IV estimates. If for example the Cragg Donald statistic is above 10% of the maximal IV size then 10% is significant at most. The stronger the instrument the closer to 10%. Whereas if the instrument is weaker our significance level will be biased upwards. That is why 10% is like a threshold to be able to say that the instrument is strong.

As we can see in the results of both panels, the instrument is strong because the Cragg-Donald values are above the 10% of Stock and Yogo (2005).

Importance

The results presented in Table 4 are crucial for the authors as they provide evidence on the relationship between population growth and conflict, addressing the endogeneity issue through their identification strategy. The main objective of the paper is to establish a causal link between population and conflict while addressing potential biases arising from omitted variables and reverse causality.

By employing 2SLS estimation, the authors address the endogeneity problem that arises from the potential correlation between population changes and other economic and institutional factors affecting conflict. The predicted mortality instrument captures the exogenous variation in population growth resulting from the international epidemiological transition, which led to improvements in health conditions and life expectancy across countries.

The IV estimates presented in Table 4 provide the causal effect of population on conflict by addressing potential biases and endogeneity concerns. These estimates are compared with the OLS estimates in Table 2, highlighting the advantages of the IV approach in mitigating omitted variable biases and capturing local average treatment effects.

Therefore, the results presented in Table 4, based on the IV/2SLS estimation, are crucial in establishing the causal relationship between population growth and conflict while addressing endogeneity concerns.

Conclusions The authors conclude that there is a positive and highly significant effect of population on conflict in almost all specifications. They find that a 10% increase in population leads to approximately 0.62 more years of conflict in the 1980s compared to the 1940s. The IV estimates, obtained through their instrumental variable approach, are larger than the OLS estimates reported in Table 2. This difference can be attributed to measurement error, capturing different local average treatment effects, and omitted variable biases affecting the OLS estimates.

Furthermore, the authors conduct robustness checks in columns 2-4 of Table 4. These checks involve using alternative datasets and different dependent variables, such as years in internal conflict as a fraction of total years and log battle deaths divided by population. In all cases, the estimated coefficients remain positive and typically significant, supporting the robustness of the IV estimates.

Overall, the results in Table 4 provide strong evidence supporting the authors' hypothesis that population growth has a significant impact on conflict. By employing their identification strategy and comparing the IV estimates with OLS estimates, the authors address endogeneity concerns and demonstrate the importance of considering the causal relationship between population and conflict.

e. Table 11

Identification Strategy

In the empirical analysis for Mexico, the authors adopt a similar identification strategy to the cross-country analysis, but with a focus on the role of malaria. Due to the unavailability of data on mortality from various diseases at the municipality level in Mexico, the authors concentrate on malaria, which was a significant factor influencing differential changes in mortality and population during the time period under study.

To construct the predicted mortality instrument at the municipality level in Mexico, the authors use malaria suitability as the basis. In the base period of 1940, the predicted mortality instrument, denoted as Mi, t, is equal to the measure of malaria suitability. Following the logic of the cross-country instrument, after the implementation of anti-malarial campaigns, particularly in 1960, the predicted mortality instrument is set to 0 in all municipalities.

The dependent variable, c_{it} focuses on protests against the PRI (Institutional Revolutionary Party) in Mexico. As there were essentially no protests against the PRI in the 1940s but reliable data on protests in the 1960s, the authors utilize protests in the 1960s normalized by population, similar to their cross-country work.

The key difference from the cross-country equations is that the predicted mortality variable, M_{it}, is based on a single disease, malaria, and uses the index of malaria suitability instead of baseline mortality rates. Additionally, x_{it} represents the log population for a given municipality and incorporates municipality fixed effects and time fixed effects as in the cross-country analysis. The vector Z_{it} includes covariates and incorporates interactions between state dummies in Mexico and a dummy variable denoting the post anti-malarial campaign period. These interactions help account for the substantial differences across states within Mexico.

Considering that spatial spillovers and correlation are of greater concern at smaller geographic units, the analysis for Mexico incorporates spatial correlation between municipalities to address this issue.

In summary, the empirical strategy for Mexico parallels the cross-country analysis but focuses on the specific context of malaria as a major determinant of changes in mortality and population within the country. The within-Mexico models employ a predicted mortality instrument based on malaria suitability, utilize protests against the PRI as the dependent variable, and incorporate municipality fixed effects, time fixed effects, and spatial correlation considerations

Interpretation

Table 11 presents the results of the analysis for Mexico, focusing on the relationship between population and violent protests. The table includes 2SLS (Two-Stage Least Squares), first stage, and OLS (Ordinary Least Squares) estimates.

In Panel A, the coefficient on log population is statistically significant in all specifications. The magnitude of the coefficient ranges from 10.9 to 14.3 in the 2SLS specifications. This suggests that an increase in population is associated with a higher incidence of violent protests per capita. For example, the 2SLS estimate in column 9 indicates that a 10% increase in population is associated with 1.3 more protests per hundred thousand inhabitants. Alternatively, the average change in log population in the sample (0.41) is associated with 5.35 more protests.

Panel B presents the first-stage estimates, focusing on the predicted mortality variable. The coefficients of the predicted mortality variable range from -0.28 to -0.21 and are consistently significant at less than 1%. These results indicate that the predicted mortality instrument is robust across specifications and supports the exogeneity of population changes.

In Panel C, the OLS estimates are compared to the 2SLS estimates. The OLS coefficients for log population are smaller than the IV estimates. This suggests that the OLS estimates are downwardly biased due to potential correlations between population changes and other economic and institutional factors that can impact conflict.

Here is the interpretation of the coefficients:

Panel A (2SLS estimates - Violent Protests):

- The coefficient on log population ranges from 10.9 to 14.3 in different specifications. These coefficients indicate that an increase in population is associated with a higher incidence of violent protests per capita.
- For example, a 10% increase in population is estimated to lead to approximately 1.3 more protests per hundred thousand inhabitants

Panel B (First-stage estimates - Predicted Mortality):

- The coefficients on the predicted mortality variable range from -0.284 to -0.211. These coefficients indicate the effect of predicted mortality on log population.
- A one-unit increase in the predicted mortality variable leads to a decrease in log population by the magnitude of the coefficient.

Panel C (OLS estimates - Log Population):

- The coefficients on log population range from 4.469 to 6.152. These coefficients indicate the estimated effect of log population on the occurrence of violent protests.
- The OLS coefficients are smaller than the 2SLS estimates, suggesting that the OLS estimates are downwardly biased due to potential correlations between population changes and other factors impacting conflict.

Overall, the coefficients in Table 11 provide evidence that higher population is associated with an increased likelihood of violent protests, supporting the hypothesis that population growth can contribute to social unrest. The instrumental variable approach helps address the endogeneity issue and strengthens the validity of the population's effect on violent protests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: 2SLS									
log population	12.372	13.333	12.813	13.102	12.389	12.996	14.263	10.948	13.049
	(3.218)	(3.040)	(3.069)	(3.047)	(3.166)	(3.056)	(3.525)	(3.113)	(4.130)
yr1960xcorr_logpop40			0.621						-0.108
			(0.220)						(0.283)
yr1960xprimary_schooling40				22.070					22.467
				(6.101)					(7.709)
yr1960xuniversity40					161.524				3.572
					(90.083)				(43.801)
yr1960xbattles_centroid40						3.021			1.746
						(0.716)			(0.596)
yr1960xshare_basin40							−0.617		−0.558
							(0.567)		(0.830)
yr1960xshare_ind40								−1.818	−1.029
								(0.752)	(0.913)
Observations	4,756	4,756	4,756	4,128	4,176	4,756	4,756	3,760	3,572
Panel B: First Stage									
predicted mortality (malaria)	-0.284	-0.262	-0.259	-0.247	-0.236	-0.259	-0.240	-0.246	-0.211
	(0.030)	(0.030)	(0.030)	(0.029)	(0.030)	(0.030)	(0.032)	(0.028)	(0.031)
Panel C: OLS									
log population	4.700	4.780	4.700	4.469	5.163	4.577	4.748	6.152	5.132
	(0.841)	(0.853)	(0.843)	(0.880)	(0.914)	(0.850)	(0.858)	(1.033)	(1.067)

Table 6: Mexico: population and violent protests per capita. 2SLS, first stage and OLS estimates. Robust standard errors in parentheses

Question 3

a. Summary

1. Summary of the paper: The paper examines the impact of mayors' tenure in office on public procurement outcomes in Italy. It analyzes procurement auctions for public works conducted by Italian mayors between 2000 and 2005. The study investigates the relationship between mayoral tenure and various procurement outcomes, including the number of bidders, winning rebates, the probability of awarding contracts to local firms, and repeated contract awards to the same firm.

2. Contribution to the literature: The paper contributes to the existing literature by focusing on the effect of political tenure on public procurement. It adds to our understanding of how the duration of mayoral office can influence procurement outcomes and sheds light on potential mechanisms driving these effects

3. Most important results: The study finds that longer mayoral tenures are associated with a decrease in the number of bidders participating in procurement auctions. Additionally, longer tenures lead to lower winning rebates, indicating higher costs for public works. The research estimates that, on average, public works in municipalities with second-term mayors cost about €3,426 more compared to municipalities with first-term mayors. Furthermore, longer tenures increase the probability of awarding contracts to local firms and repeated contract awards to the same firm.

4. Identification strategies: The paper employs two identification strategies to establish causal relationships. Firstly, a Regression Discontinuity (RD) design is used by comparing elections where the incumbent mayor narrowly won another term with elections where the incumbent mayor narrowly lost and a new mayor took over. This approach helps isolate the effect of tenure on procurement outcomes. Secondly, the study cross-validates the RD estimates by leveraging a quasi-experiment created by the introduction of a two-term limit on mayoral office. This allows for comparisons between mayors appointed before and after the reform, using the date of the first election as an instrument for measuring tenure.

b. Figure 1

Based on Figure 1 of the article "Tenure in Office and Public Procurement," if we use the number of bidders as the outcome variable and relate it to the margin of victory of mayors in the regression discontinuity, we can say the following:

For "numbers of bidders" if the margin of victory is low and the mayor is not elected: If the mayor's margin of victory is low and they fail to be elected, this is likely to have an impact on the number of bidders. Mayors with less solid electoral support and less stability in office may have less influence on procurement processes. This can create a more competitive environment and incentivize participation in the bidding process. On the other hand, once the election is won and there is more tenure in office, following the same argument as before, companies seek to leverage their connection and perceived favoritism with the elected mayor, which could result in a lower number of bidders compared to a more competitive scenario. This is why in the graph in panel A, the number of competitors is higher before the margin of victory. Among all the outcomes, this is the one that shows the most discontinuity in response after the cutoff.

If the number of bidders is higher below the margin, in summary, this could be due to several reasons related to auction theory and competition in the public procurement process.

1. Intense competition: When the mayor's margin of victory is low, it indicates a tighter competition among candidates. In this scenario, there is likely to be a greater number of companies interested in participating in the bids, which increases competition and, in turn, the number of bidders. Companies may perceive that they have a greater chance of winning contracts, which increases their participation in the bidding process.

2. Opportunities-seeking: A low margin of victory can generate uncertainty about political stability and the continuity of policies under the outgoing mayor. This may motivate more companies to participate in the bids as they see an opportunity to secure contracts before potential changes in the local administration occur. Companies may consider their bids to be more competitive in a lower margin of victory environment.

3. Diverse interests: A low margin of victory may reflect a greater diversity of interests and opinions among voters. This could translate into a greater variety of local projects and needs that require public services or projects, which in turn attract a higher number of bidders interested in addressing those specific demands. The more companies perceive profitable opportunities in the public procurement market, the higher the number of bidders willing to participate.

Additionally, it can be observed that the relationship is cubic below the margin. This implies that the number of bidders may increase more pronouncedly within a specific range of the margin of victory and then stabilize or decrease after that optimal point. From this, we can say that at the margin of victory, the number of bidders will resemble the context where the victory has just been achieved, which is why it will be lower.

Regarding panel B, we can notice that before the margin of victory, there is a lot of dynamics in the relationship between election margin and winning rebate. Initially, the rebates are low and then increase, and finally, they decrease again when close to the margin of victory. On the other hand, above the margin of victory, the rebate is always lower, and the relationship is linear, similar to a constant.

Here is a possible explanation for this relationship based on auction theory and economic incentives:

1. Initial effect of uncertainty: Initially, when the mayor's margin of victory is low, there may be greater uncertainty and risk for the winning companies in terms of political stability and continuity of the outgoing mayor's policies. In this scenario, companies may offer lower winning rebates due to the uncertainty and lack of information about the new mayor. Companies may be cautious and reluctant to

offer higher financial incentives at this point.

2. Increase in the winning rebate: As the margin of victory increases, there may be greater political stability and a greater influence of the mayor in procurement decisions. This could lead to an increase in the winning rebate as winning companies feel more secure about the continuity of favorable policies for them. Companies

may be willing to offer higher financial incentives to ensure winning the contracts and maintaining favorable relationships with the elected mayor.

3. Decrease near the margin: Near the margin of victory, where the election is tighter, the winning rebate may start to decrease. This could be due to a combination of factors such as increased competition among companies and reduced ability of the mayor to influence contract allocation. As competition intensifies and the margin narrows, companies may reduce the offered financial incentives as they perceive fewer opportunities to obtain significant economic benefits.

Regarding panels C and D, we can notice that there are no significant differences before and after the margin of victory. Some approximations about this result:

1. Since the "winner local" outcome is constant for mayors both below and above the margin of victory, this could imply that the mayor's margin of victory does not have a significant impact on the preference for local firms in the public procurement process. In this case, regardless of whether the margin of victory is low or high, local companies are expected to have an equal chance of winning public contracts. This could be due to consistently implemented preferences for local companies or a legal framework that promotes the participation of local companies in public procurement, regardless of the mayor's margin of victory.

2. Since the "maximum percentage of wins in the same firm" is equal for mayors both below and above the margin of victory, this suggests a constant linear relationship between the mayor's margin of victory and the mentioned outcome. One possible interpretation of this constant relationship could be that the mayor's margin of victory does not have a significant impact on the preference for the same firm in the public procurement process. This could indicate that, regardless of the margin of victory, the same company or group of companies has a constant probability of obtaining contracts in the region or municipality in question.

There are different possible explanations for this relationship:

- Long-term associations: There may be long-term contracts or agreements between the local government and a specific company or group of companies, regardless of the electoral outcome. These long-term associations may lead to a constant percentage of contracts won by the same firm, whether the margin of victory is high or low.

- Capabilities and expertise: It is possible that a company or group of companies has consistently demonstrated their expertise and capacity in the public procurement sector in that region or municipality. This can lead to them being selected for the majority of contracts, regardless of the electoral outcome.

- Inertia in the procurement process: There is a possibility that the public procurement process has inertia or inefficiencies that lead to continuity in the allocation of contracts to the same company or group of companies, regardless of the mayor's margin of victory.

It is important to note that these are potential explanations, and the precise relationship between the mayor's margin of victory and the "maximum percentage of wins in the same firm" will depend on the specific context and public procurement dynamics addressed in the article you are reviewing. Please refer to the original study for detailed information on the results and empirical analysis presented regarding this relationship.

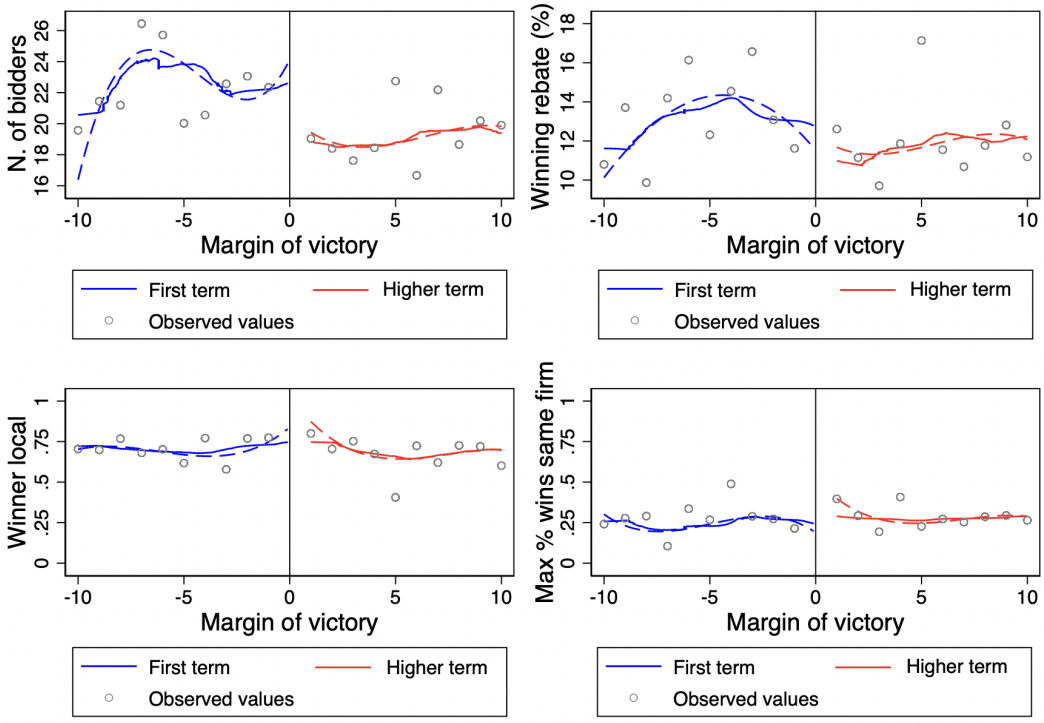

Figure 1: Outcomes RD

c. Continuity in RD

The RD framework offers a valuable tool for examining the validity of the continuity assumption. As discussed by Lee (2008), it enables us to compare pre-intervention characteristics between the treated and control groups. By doing so, we can identify any nonrandom selections that may have occurred around the threshold and observe if these characteristics exhibit systematic differences. To assess the validity of the continuity assumption more thoroughly, we will estimate equation (2) using pre-intervention characteristics as an outcome variable. To evaluate this, it can be presented both in regression tables and in classic RD plots where discontinuities can be seen. As we can notice in Figure 2, in general the continuity is fulfilled for most of the covariates, indicating balance of the characteristics with the exception of "Education in College", "Center" and "Northwest". These variables show discontinuities in the cutoff. Now we will go deeper into this analysis using the regression table.

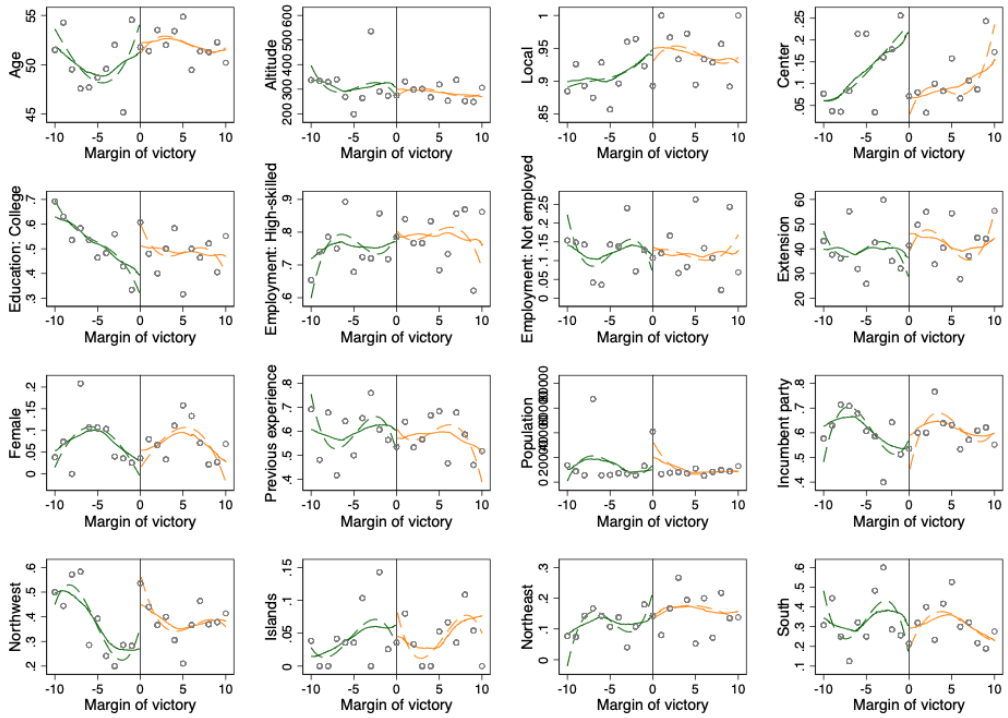

Figure 2: RD with covariates as outcomes

To evaluate this we present the table 7 which evaluates the pre-treatment characteristics in the vicinity of the threshold. Thus, this table estimates a simplified version of equation two by making the pre-treatment variables the dependent variables. As we can see the characteristics at the municipal and individual levels are well balanced which shows that the continuity assumption is likely to be met. In addition, descriptive statistics of the individual characteristics of the mayors have been left in the appendix. Now, unlike the table 7 which is much more rigorous as it evaluates continuity in the threshold, the nexus table contains all mayors.

Additionally, in **question (e)**, we will closely examine the distribution of the margin of victory around the threshold and apply the McCrary (2008) test to gain further insights into the validity of the continuity assumption.

Panel A. City characteristics	
Northwest	0.069
	(0.04)
Northeast	−0.002
	(0.03)
Center	−0.089
	(0.03)
South	0.02
	(0.04)
Islands	0.002
	(0.016)
Population	−1, 520.97
	(5, 284.21)
Altitude	−3.255
	(22.241)
Extension	−3.362
	(4.458)
Panel B. Mayor's characterist	
Female	−0.015
	(0.022)
Age	2.634
	(0.767)
Local	0.009
	(0.021)
Education: College	0.044
	(0.04)
Employment: Not employed	0.010
	(0.026)
Employment: High-skilled	0.021
	(0.035)
Previous experience	−0.050
	(0.041)
Incumbent party	−0.007
	(0.040)
Number of mayors	2,268

Table 7: Balance characteristics. Notes: All cities/mayors in the RD estimation sample. Altitude is the city's altitude above sea level. Extension is the geographical extension of the city administrative territory. Population is the census popula tion as of 1991. Local is a dummy for being born in the same region. High-skilled includes entrepreneurs and self-employed. Previous expe rience is a dummy for whether the mayor was in the council or in the executive committee before. Incumbent party is a dummy for whether the mayor belongs to the incumbent party

d. Table 5 and 6

VARIABLES	(1) N. of bidders RDD	(2) N. of bidders RDD	(3) Winning rebate OLS	(4) Winning rebate RDD
N. terms in office	-1.866	-2.469***	-0.903*	-0.705**
	(1.859)	(0.930)	(0.543)	(0.308)
Term limit binding		3.740**		1.181**
		(1.715)		(0.499)
Population		0.090***		0.031***
		(0.025)		(0.004)
Starting value		0.746***		0.104***
		(0.093)		(0.017)
Female mayor		0.006		0.146
		(0.845)		(0.309)
Age of the mayor		0.035		0.025**
		(0.030)		(0.010)
N. terms in office (party)		-0.885		-0.103
		(0.583)		(0.192)
Observations	12,687	12,687	12,687	12,687
R-squared	0.004	0.255	0.004	0.464
Number of Mayors	2195	2195	2195	2195
Avg. outcome	21.52	21.52	12.26	12.26

Table 8: Tenure in Office and the level, RD- Table 5

In order to examine the impact of mayoral tenure on auction outcomes, the coefficients are calculated by dividing the effect of the variable by its mean.

Table 5 presents the findings regarding the effects of tenure on two important auction outcomes: the number of bidders and the winning rebate. The analysis, conducted with controls for a comprehensive set of observable characteristics, reveals significant results. Specifically, in columns 2 and 4, where the full set of controls is included, the effect of tenure on the number of bidders is -11.48 percent (-2.469 divided by 21.52), while the effect on the winning rebate is -5.7 percent (-0.705 divided by 12.26). It is noteworthy that when controls are excluded (columns 1 and 3), the qualitative patterns of the results remain consistent. This consistency provides initial evidence supporting the validity of the regression discontinuity (RD) strategy employed in the analysis. Other variables are not analyzed in this context as they serve as controls.

Moving to Table 6, the focus shifts to examining the effects of tenure on two additional auction outcomes that characterize the nature of competition. The RD estimates indicate statistically significant positive effects of tenure on whether the winning firm is registered in the same region (5 percent) and the highest percentage of auctions awarded to the same firm within the tenure (25.6 percent). The 5 percent figure is derived from dividing 3.458 by 70.20, while the 25.6 percent figure is obtained by dividing 5.729 by 22.37.

While direct evidence of mayoral misconduct is not available, the last two estimates offer valuable insights into the potential mechanisms underlying the deterioration observed in the procurement process described in Table 5. Specifically, the findings regarding the geographical origin of the winning firm suggest a plausible compatibility with the hypothesis that, as mayors remain in power for longer durations, there is a higher probability of favoritism towards local bidders. These results indicate that extended tenures might lead to an increased likelihood of mayors distributing benefits and favors to local firms during the procurement process.

VARIABLES	(1) Winner local RDD	(2) Winner local RDD	(3) Max % wins same firm OLS	(4) Max % wins same firm RDD
N. terms in office	4.776**	3.458	6.023***	5.729***
	(1.926)	(2.173)	(1.861)	(2.189)
Term limit binding		-1.604		-2.491
		(3.325)		(2.837)
Population		0.003		-0.108**
		(0.015)		(0.044)
Starting value		-0.992***		-0.157***
		(0.105)		(0.036)
Female mayor		3.035*		-4.637**
		(1.730)		(1.841)
Age of the mayor		-0.078		-0.222***
		(0.075)		(0.067)
N. terms in office (party)		0.552		1.367
		(1.335)		(1.146)
Observations	12,687	12,687	11,099	11,099
R-squared	0.002	0.100	0.032	0.309
Number of Mayors	2195	2195	1814	1814
Avg. outcome	70.20	70.20	22.37	22.37

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Tenure in Office and the nature, RD - Table 6

It is important to note that these findings provide suggestive evidence and do not establish causality or definitively prove the existence of any wrongdoing. Further investigations and additional evidence would be necessary to establish the underlying mechanisms with greater certainty. Nevertheless, the observed patterns lend support to the notion that longer tenures could potentially influence procurement outcomes, thereby compromising the fairness and competitiveness of the process.

e. Data manipulation and McCrary test

In the context of the article we are working on, which focuses on the relationship between mayors' tenure and the outcomes of public tenders in Italy between 2000 and 2005, direct manipulation of election results to influence the margin of victory appears to have a low probability of occurrence in democratic contexts.

In Italy, during the years under study, the electoral system is considered relatively stable and governed by democratic principles. Although no electoral system is perfect and there is always a theoretical possibility of manipulation, in consolidated democracies like Italy, electoral institutions and oversight mechanisms are usually robust and designed to ensure the integrity of electoral processes.

Direct manipulation of electoral results, such as bribery or fraud, is not only an undemocratic practice that is widely condemned and penalized by law but also carries costs, usually high ones. Political parties, candidates, and citizens tend to value transparency and the integrity of electoral processes. Therefore, in the specific case of the study on the outcomes of public tenders in Italy between 2000 and 2005, direct manipulation of electoral results to be above or below the cutoff threshold seems unlikely. The election results would mostly reflect the will of the voters and the political preferences of the population, rather than intentional manipulation to influence the tender outcomes.

However, it is important to note that while direct manipulation of electoral results may be improbable, there could be other factors or biases in the study that could influence the tender outcomes. These could include political or economic influences, relationships among the actors involved in the tender process, or even biases inherent in the data used in the analysis.

To assess the possibility of manipulation in the outcomes, the McCrary test can be used. This test is a statistical tool designed to detect manipulation of variables around a cutoff or threshold point. In the context of the article, the McCrary test can be applied to examine if there are significant differences in the margin of victory around the cutoff used in the study.

The McCrary test allows determining if the observed differences in the margin of victory near the cutoff point are statistically significant and consistent with the existence of manipulation. It is based on the idea that if there is no manipulation, the margin of victory should exhibit a continuous distribution around the cutoff. On the contrary, if there is manipulation, discontinuities would be expected in the distribution of the margin of victory around the cutoff point.

When conducting the McCrary test, the distributions of the margin of victory before and after the cutoff can be graphically examined, and statistical tests can be performed to evaluate the presence of significant discontinuities. If there is evidence of discontinuities in the margin of victory around the cutoff, this would suggest the possible existence of manipulation in the outcomes.

More specifically, in the McCrary test, the null and alternative hypotheses are formulated regarding the presence of a discontinuous regression in the data around a specific cutoff point.

Null hypothesis (H_0): There is no discontinuous regression in the data around the cutoff point.

Alternative hypothesis (H_1): There is a discontinuous regression in the data around the cutoff point. More precisely, the null hypothesis states that there is no sharp or discontinuous change in the outcomes around the cutoff point, while the alternative hypothesis argues that such an abrupt change does exist.

The McCrary test is based on a regression design approach called "local polynomial regression" or "local regression." The key steps of the test are described below in a general manner:

1. Model specification: A regression model is specified that relates the outcome variable (e.g., the effect of a housing subsidy program) to the threshold variable (e.g., household income). The model includes polynomial power terms to capture possible nonlinear changes around the cutoff point.

2. Parameter estimation: The regression model is estimated using the available data, employing weighted least squares methods or similar approaches to fit the model to the observed data.

3. Testing for regression discontinuity: The test is conducted to assess whether there is evidence of a discontinuous regression in the data. This involves comparing the density of observations just on both sides

of the threshold, in a narrow window around the cutoff point. A statistical test, such as the McCrary test, is used to evaluate whether the density of observations is significantly different from what would be expected under the null hypothesis of no discontinuous regression.

4. Estimating the jump in outcomes: If evidence of a discontinuous regression is found, the jump in outcomes (the sharp change in the values of the outcome variable) around the cutoff point can be estimated. This provides a quantitative measure of the impact of the threshold on the outcome variable.

Now, the estimates depicted in Figure 2 indicate that the density of margin of victory (MV) is relatively smooth and exhibits expected behavior around the threshold, aside from minor fluctuations due to sample noise. Conducting a formal density test (McCrary, 2008) yields evidence against the existence of a statistically significant jump. The estimated log-difference is -0.18, with a standard error of 0.13.

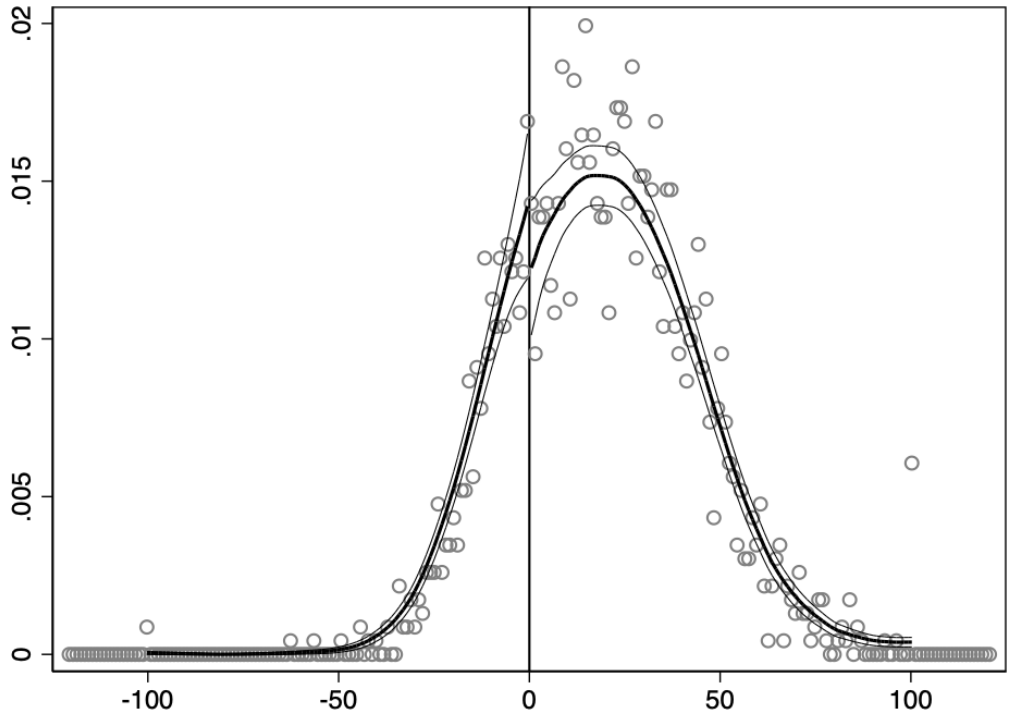

Figure 3: McCrary test

f. Table 5 and 6 - polynomial of degrees 2 and 4

Replicate

VARIABLES	(1) N. of bidders RDD	(2) N. of bidders RDD	(3) Winning rebate OLS	(4) Winning rebate RDD
N. terms in office	-1.840	-2.414***	-0.844	-0.714**
	(1.724)	(0.928)	(0.539)	(0.307)
Term limit binding		3.422**		1.231**
		(1.592)		(0.486)
Population		0.090***		0.031***
		(0.025)		(0.004)
Starting value		0.745***		0.104***
		(0.093)		(0.017)
Female mayor		-0.004		0.148
		(0.843)		(0.309)
Age of the mayor		0.036		0.024**
		(0.031)		(0.010)
N. terms in office (party)		-0.905		-0.100
		(0.585)		(0.193)
Observations	12,687	12,687	12,687	12,687
R-squared	0.004	0.255	0.004	0.464
Number of Mayors	2195	2195	2195	2195
Avg. outcome	21.52	21.52	12.26	12.26

Table 10: Tenure in Office and the level, RD - Table 5 with polynomial degree 2

VARIABLES	(1) N. of bidders RDD	(2) N. of bidders RDD	(3) Winning rebate OLS	(4) Winning rebate RDD
N. terms in office	-2.515	-2.448***	-1.172*	-0.761**
	(2.082)	(0.937)	(0.621)	(0.309)
Term limit binding		3.790**		1.044**
		(1.806)		(0.514)
Population		0.090***		0.031***
		(0.025)		(0.004)
Starting value		0.746***		0.104***
		(0.093)		(0.017)
Female mayor		0.011		0.131
		(0.848)		(0.307)
Age of the mayor		0.036		0.024**
		(0.031)		(0.010)
N. terms in office (party)		-0.882		-0.110
		(0.582)		(0.193)
Observations	12,687	12,687	12,687	12,687
R-squared	0.005	0.255	0.005	0.464
Number of Mayors	2195	2195	2195	2195
Avg. outcome	21.52	21.52	12.26	12.26

Table 11: Tenure in Office and the level, RD - Table 5 with polynomial degree 4

VARIABLES	(1) Winner local RDD	(2) Winner local RDD	(3) Max % wins same firm OLS	(4) Max % wins same firm RDD
N. terms in office	4.459**	3.419	5.056***	5.831***
	(1.880)	(2.171)	(1.841)	(2.195)
Term limit binding		-1.378		-3.371
		(3.214)		(2.739)
Population		0.003		-0.112**
		(0.015)		(0.044)
Starting value		-0.992***		-0.159***
		(0.105)		(0.036)
Female mayor		3.042*		-4.666**
		(1.729)		(1.842)
Age of the mayor		-0.079		-0.219***
		(0.075)		(0.067)
N. terms in office (party)		0.567		1.310
		(1.332)		(1.146)
Observations	12,687	12,687	11,099	11,099
R-squared	0.002	0.100	0.026	0.309
Number of Mayors	2195	2195	1814	1814
Avg. outcome	70.20	70.20	22.37	22.37

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Tenure in Office and the nature, RD - Table 6 - Polynomial degree 2

VARIABLES	(1) Winner local RDD	(2) Winner local RDD	(3) Max % wins same firm OLS	(4) Max % wins same firm RDD
N. terms in office	5.360***	3.901*	6.787***	5.738***
	(1.984)	(2.169)	(2.031)	(2.206)
Term limit binding		-0.521		-2.457
		(3.372)		(2.871)
Population		0.005		-0.108**
		(0.015)		(0.044)
Starting value		-0.990***		-0.157***
		(0.105)		(0.036)
Female mayor		3.155*		-4.634**
		(1.730)		(1.839)
Age of the mayor		-0.074		-0.222***
		(0.075)		(0.067)
N. terms in office (party)		0.604		1.367
		(1.330)		(1.146)
Observations	12,687	12,687	11,099	11,099
R-squared	0.002	0.100	0.033	0.309
Number of Mayors	2195	2195	1814	1814
Avg. outcome	70.20	70.20	22.37	22.37

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 13: Tenure in Office and the nature, RD - Table 6- Polynomial degree 4

Interpretation

Model	*Variable*	*Degree of polynomial*		
		2	3	**4**
Table 5	*Number of bidders*	**11.21%**	11.48%	**11.37%**
	Winning rebate	**5.8%**	5.7%	**6.2%**
Table 6	*Winner local*	**4.9%**	5%	**5.6%**
	Max wins same firm	**26%**	25.6%	**25.6%**

Table 14: Summary interpretation tables 5 and 6 with different degree of polynomial

The implications of the results when changing the polynomials are not significant. As can be seen in Table 14, the results do not vary substantially.

Now, the difference between using polynomials of degree 2, 3, or 4 in local regressions lies in the flexibility of the model to capture more complex and nonlinear patterns in the data around the cutoff point. As the degree of the polynomial increases, a greater curvature is allowed in the relationship between the outcome variable and the threshold variable.

The choice of polynomial degree depends on the shape of the data and the expected complexity in the relationship. Here are some implications of using polynomials of degree 2, 3, and 4 in local regressions:

1. Degree 2 polynomial: A degree 2 polynomial models a quadratic relationship. It can capture patterns of change in the form of a parabola around the cutoff point. This implies that the causal effect can have a concave or convex curve, depending on the direction of the parabola's slope.

2. Degree 3 polynomial: A degree 3 polynomial models a cubic relationship. It adds greater flexibility by allowing more pronounced and nonlinear changes in the relationship around the cutoff point. It can capture more complex curves, such as S-shaped curves or changes in the slope direction.

3. Degree 4 polynomial: A degree 4 polynomial adds even more flexibility by allowing even more complex changes in the relationship. It can capture more complicated curve shapes and abrupt changes in the slope around the cutoff point.

The choice of polynomial degree depends on the exploration of the data and the interpretation of its shape. It is important to find a balance between model flexibility and the risk of overfitting (overly fitting the specific sample data and losing generalizability).

In summary, the degree of the polynomial in the local regression determines the model's flexibility to capture nonlinear patterns in the relationship between the outcome variable and the threshold variable. Higher-degree polynomials allow capturing more complex shapes, but they should be used with caution to avoid overfitting the model to specific sample data.

In these models, the original estimation was done with a degree 3 polynomial. As can be seen, the results of the degree 2 and degree 3 polynomials are quite similar, even statistically significant. In that sense, an important conclusion is that for achieving greater simplicity, we should stick with the degree 2 polynomial despite the fact that the degree 3 polynomial provides more flexibility.

g. IV and assumptions

The authors employ instrumental variables (IV) analysis to confirm their results and address potential endogeneity concerns. They use a unique quasi-experiment created by the introduction of a two-term limit

on the mayoral office in March 1993 as their instrumental variable. This quasi-experiment provides a natural source of exogenous variation in mayoral tenures.

The identification assumptions for the instrumental variable strategy are as follows:

- Exclusion restriction: The instrumental variable (time of first election relative to the reform) must only affect the outcome variable (procurement outcomes) through its impact on the tenure of mayors and should not have a direct effect on the outcome. In this case, the assumption is that the 1993 reform primarily affected the length of tenure and not the administration of public procurement auctions directly.
- Relevance: The instrumental variable should be correlated with the endogenous variable (mayoral tenure) to ensure its relevance in explaining the variation in tenure. The introduction of the two-term limit creates a clear distinction between mayors elected before the reform (who could serve two additional terms) and those elected after the reform (who could serve only one additional term).

To address concerns related to the changing electoral rule introduced by the 1993 reform, the authors adopt a fuzzy regression discontinuity design (fuzzy-RD). They focus on mayors elected immediately before and after the reform, minimizing the potential bias from the changing electoral system. This approach helps mitigate any confounding effects caused by the electoral rule change on the outcomes of public procurement auctions.

The authors estimate their equations within a two-stage least squares (2SLS) framework. In the first stage, they regress the number of terms in office on the instrumental variable (indicator for election before March 1993, augmented with the distance of the first election from the reform). To account for the discrete nature of election timing, they include time dummies as flexible functions. The second stage involves estimating the relationship between mayoral tenure and procurement outcomes using the predicted values from the first stage as the instrumental variable for tenure.

To ensure the validity of the exogeneity assumption, the authors compare a set of pre-intervention characteristics for the treated and control groups. If there were nonrandom selections around the 1993 reform, these characteristics would exhibit systematic differences. This analysis helps assess the potential bias resulting from any nonrandom selection and strengthens the identification strategy.

Overall, the instrumental variable analysis allows the authors to address endogeneity concerns and provide additional evidence to support their findings.