Kent Academic Repository

School of Economics
University of Kent

Essays on the impact of politicians' characteristics and speeches on policy outcomes

Alessio Mitra

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Economics at the School of Economics of the University of Kent - Feb, 2023

Abstract

This thesis consists of three self-contained essays. The first substantive chapter, entitled *"Should you want an educated mayor? Evidence from close elections in Italy"*, studies the impact of politicians' education on policy choices and public finance. A probabilistic voting model with candidates' education level and public expenditure decomposition is used to explain how a politician's education matters for policy decisions over public budget allocation. Empirically, I use Italian municipality data on electoral results, balance sheets, and mayor candidates' educational attainment from 2000 to 2015. To estimate the causal effect of mayors' education on public finance I rely on regression discontinuity design focusing on close elections. Overall, I find that educated mayors boost public investment, especially in the education sector, without compromising the fiscal stability of the municipalities.

The second substantive chapter, entitled *"The impact of women's political leadership on institutional quality"*, studies the impact of female political leadership on future women's participation in politics and hiring in public administration, as well as on different measurements of institutional quality. Empirically, I use Italian municipal and provincial data on elections, politicians, and institutional characteristics from 2000 to 2015. To estimate the causal effect of electing female mayors I rely on regression discontinuity design, and instrumental variable approach, focusing on close elections. Overall, I find that women's political leadership boosts women's participation as politicians and public servants while improving a wide set of institutional quality metrics, reducing corruption, increasing government effectiveness, the rule of law, and political accountability. Furthermore, the positive impact of female leadership on institution quality is stronger in areas with more established gender norms, and if the female politician is holding a tertiary education degree.

The third substantive chapter, entitled *"Do speeches at the UN General Assembly affect international aid allocation?"*, studies whether countries' speeches at the United Nations General Assembly matter for international aid allocation. I use a supervised machine learning algorithm called *wordscore* and 7,533 different statements from 198 countries made in the period from 1975 to 2018 to measure countries' preferences from text. I find considerable evidence that countries' preferences derived from speeches affect aid allocation. Both the US and Russia provide more aid to those countries whose speeches please their current agenda and preferences the most.

Acknowledgements

I would like to express my gratitude to my supervisors, Prof. Irma Clots-Figueras and Prof. Anirban Mitra, for their invaluable academic and technical guidance throughout the writing of this thesis. Their constructive feedback and transparent approach have been instrumental in equipping me with the right mindset to conduct meaningful research. Their constant and generous support has been a source of motivation for me.

Additionally, I wish to extend my thanks to the following professors, Olle Folke, Luca Repetto, Johanna Rickne, Guy Tchuente, Nicole Rae Baerg, and Charles R. Dannehl, for their valuable suggestions and inputs.

During my Ph.D., I had the privilege of presenting my work at numerous conferences and seminars. I am grateful to all those who provided me with feedback and suggestions to improve my work. I would also like to express my gratitude to my colleagues and friends at Kent for engaging in thought-provoking research discussions, and especially to my office-mate and friend Werner Pena for his invaluable feedback and engaging conversations.

Finally, I am deeply grateful to my family for their unwavering love and support.

Contents

List of Tables

List of Figures

Chapter 1

1 Introduction

Politics is everywhere, shaping most dimensions of our economy and society. In Western democracies, politicians affect the lives of millions with their decisions. In the literature, politicians have been represented in the most diverse ways, from the self-interested individuals that pursue wealth and fame (Downs et al. 1957, Buchanan et al. 1965), to the benevolent actors that exclusively care about maximizing their community well-being (McCain 2019). As is often the case, truth is probably in the middle, with politicians being individuals like all of us, shaped by their personal characteristics and life experiences, acting on the basis of both their self-interest and beliefs, constrained by the laws and norms of society (Persson & Tabellini 2002).

Starting from this base, this Ph.D. thesis begins by analysing how politicians' characteristics affect their decisions as policy-makers. I focus on two relevant personal features such as education and gender. Subsequently, I move into exploring how the words used by politicians during speeches can shape other politicians' policy decisions. Indeed, if personal characteristics can be seen as the building blocks of one political identity, speeches are the vehicle often employed to transmit such an identity.

Politicians' level of education is a relevant topic in the public debate, with many claiming the up-most criticality of having well-educated politicians, and numerous critics downplaying its relevance. Supporters of the importance of education often highlight the competence and managerial requirements that are needed to effectively manage and allocate public resources. At the same time, education sceptics depict education as a social-class feature, which brings politicians far away from the needs of the common people. Being both mechanisms theoretically reasonable, it became an empirical question to verify their veracity.

Politicians' gender is another relevant dimension to be considered, indeed women are still under-represented in all political positions. Female participation in politics may not only be relevant for concerns of gender equality but also due to the very specific richness of perspectives that women may bring to the table of policy-making. Since birth, women face very different life challenges than men, potentially making them more aware of certain social issues. Again, an empirical approach is required to find a satisfactory answer; does gender matter for policy-making?

Words are vehicles of meaning, yet words spoken by political leaders can be seen as either cheap talk or sincere vision. However, regardless of their nature, they are likely to exert power and influence, perhaps influencing policy-making.

In order to conduct the first and second substantive chapters of this Ph.D. thesis, I take advantage of a very detailed dataset on politicians that contested for a mayoral seat in Italy during 2000-2015. In the first substantive chapter, I combine this data with municipal balance sheet data. In the second substantive chapter, I combine it with provincial indexes of institutional quality. To conduct the third substantive chapter, instead, I take advantage of a very rich dataset containing the corpus of all the speeches performed by national representatives at the United Nations General Assembly during 1975-2018. I combine this data with international aid allocation information.

In the Ph.D. thesis, I face a common identification problem, which is to identify the causal impact of politicians' characteristics and speeches on policy outcomes. As an example, there may be an omitted variable that is related to the politicians' characteristics being examined, and the policy outcome of interest. In order to tackle this problem, in the first two substantive chapters, I use a common identification strategy, namely a regression discontinuity design focused on close elections (Lee & Lemieux 2010, Calonico et al. 2014a). I define close elections as those elections in which the winner won by a relatively small amount of votes at the ballot. Being the outcome of a close election largely random, the identification strategy quasi-randomises the municipality characteristics allowing to compare municipalities that differ only on a specific feature of their mayors, with respect to some outcome variables. In the third substantive chapter, instead, I employ a machine learning algorithm to extract political orientation from speeches, and a panel regression model with various controls and fixed effects to relate the extrapolated political orientations to aid allocation. Here, I do not claim causality, however, I perform an additional synthetic control approach to support my findings.

I find that educated mayors significantly increase public investments, particularly in the education sector, at the expense of current expenditure, without increasing taxes or making more debt. Such effect is consistent regardless of the political affiliation of the politician or other personal characteristics, such as gender or age. I also find that female mayors boost future female participation in politics and public administration while improving institutional quality. The positive effect on institutional quality is larger in more conservative areas, and when the elected woman is more educated. Finally, I find that political alignment emerged from political speeches performed at the United Nations affects international aid allocation, with the US and Russia granting more aid to those countries that showcase speeches more in line with their own.

Most of the research on politicians' educational background and policy choices relies on control variable approaches in panel data (Dreher et al. 2009, Hayo & Neumeier 2014), hence not fully addressing the endogeneity issue. In contrast, the few papers that address the

aforementioned issue (Rocha et al. 2018, Freier & Thomasius 2016) by employing a similar identification strategy to the one used in this thesis, focus on politicians' experience, with less convincing results on education. The first substantive chapter contributes to this literature by finding causal evidence of the impact of politicians' education on public finance, and by proposing a simple theoretical model to explain the mechanism through which the effect may be coming.

There is a large literature on politicians' gender and policy choices, yet most results that find an impact of gender on policy outcomes come from developing countries, such as India and Brazil (Hessami & da Fonseca 2020, Clots-Figueras 2011, 2012, Chattopadhyay & Duflo 2004). Furthermore, most studies focus on the impact that female political leadership has on female political participation, public spending, or corruption (Cowper-Coles 2020). The second substantive chapter contributes to this literature in three ways. Firstly, by finding causal evidence of the impact of politicians' gender on institutional quality. Secondly, by providing evidence of the causal impact of female political representation on future female political and work participation. And thirdly, by unveiling how such causal impact interacts with the level of education of the female politicians and the gender norms of the local environment.

The literature on the relationship between international aid allocation and countries' political preferences is broad, yet it has overwhelmingly relied on UN General Assembly voting records (Bernstein & Alpert 1971, Rai 1980, Wittkopf 1973, Wang 1999, Alesina & Dollar 2000, Alesina & Weder 2002, Gates et al. 2004). Such literature disregards almost completely an arguably richer source of information: speeches. The third substantive chapter contributes to this literature by computing the evolution of UN members' preferences from 1975 to 2018, in terms of the relative distance between US and Russian preferences, using politicians' speeches text as data, and showing how government preferences derived from politicians' speeches matter for aids allocation.

Chapter 2

Should you want an educated mayor? Evidence from close elections in Italy

2.1 Introduction

Education is widely accepted as a fundamental element of human society. It produces an impact on individuals' behaviour as well as on economic outcomes. Empirical evidence links it to higher individual earnings (Card 1999) and pro-social behaviours (Dee 2004, Milligan et al. 2004). However, from a political perspective, there is little convincing evidence of its impact on politicians' behaviour and decision-making.

The aim of this paper is to answer the question: does the education level of politicians matter for policy choices and public finance? Due to the rise of populism across the world, this question has experienced rising popularity in mass media and political debate. Indeed, populist parties tend to narrow the importance of politicians' education (Clarke & Newman 2017, Collins et al. 2020). Highly qualified candidates are often depicted with suspicion, portrayed as being part of selfish elites, and claimed to be far away from the interests of the common people (Canovan 1984, Rooduijn 2014, Merkley 2020).

To add context to the discussion, it is crucial to understand and verify the policy-making differences between more and less educated politicians. The paper brings novelty in two ways. Firstly, it finds causal evidence of the impact of politicians' education on public finance. Secondly, it proposes a simple theoretical model to explain the mechanism through which the effect may be coming.

I match rich and detailed Italian data on mayors candidates' level of qualification, mayoral electoral results, and municipality balance sheets from 2000 to 2015. To overcome endogeneity issues I rely on regression discontinuity design focusing on close elections. The idea is that since politicians cannot perfectly predict the result of an election, the results of closely contested elections can be taken as good as random. The identification strategy quasi-randomises the municipality characteristics allowing me to compare municipalities that differ only in the level of education of their mayors with respect to some outcome variables. Mayors with or without tertiary education are quasi-randomized to municipalities, then I test and verify the identifying assumption that there is no difference in traits around the threshold.

From a theoretical perspective, politicians' education level is usually included in models as a desired quality that voters are looking for. In the post-politics career concern model of

Persson & Tabellini (2002) voters can select politicians on the basis of their ideology, competence, and behaviour as an incumbent. Voters reward competence, hence incumbents try to signal a higher level of competence to win elections. Besley (2005, 2006) defines competence as a valence issue: every citizen wants more of it regardless of the policy choice being implemented. Caselli & Morelli (2004) employs the citizen-candidate framework of Osborne & Slivinski (1996) and Besley & Coate (1997) to investigate how competent politicians are selected into office. I contribute to this literature by defining politicians' education as a means to an end rather than a desired goal. I build upon Persson & Tabellini (2002) pre-politics probabilistic voting model and focus on public budget allocation. I define a politician's education as a required technology to engage in different kinds of policy choices. As a result, politicians will frame their electoral platform to suit best their level of competence. In this regard, policy choices may be reduced to a simple decision between current expenditure and public investment. From the political candidate's perspective, relying greatly on current expenditure is an easier task than planning, mediating, and implementing public investments. As a result, low-quality politicians will likely shift their electoral platform toward current expenditure, at the expense of public investment.

Empirically, most of the research on politicians' backgrounds and policy choices relies on control variable approaches in panel data. Dreher et al. (2009) employ a panel over the period 1970–2002 covering the profession and education of more than 500 political leaders from 72 countries and found mixed results on the impact of politicians' education and professional background on market-liberalizing reforms. Hayo & Neumeier (2014), employs a panel with information regarding German Prime Ministers' backgrounds and shows that Prime Ministers from poorer socio-economic backgrounds are associated with higher levels of public spending.

In the attempt of excluding the presence of omitted variable bias Rocha et al. (2018) performed a regression discontinuity design focusing on close elections in Brazilian municipalities with data from 2000 to 2008. The research found that more experienced mayors reduce current expenditure and deficit during their term. Freier & Thomasius (2016) applied a fuzzy regression discontinuity design to German municipalities with data from 1950 to 2009 showing similar results: experienced mayors reduce current expenditure and deficit, however, no impact of mayor's education level is detected. Worth noting that Freier & Thomasius (2016) did not have exact information on politicians' education, hence they had to resort to using the probability of having a higher education level by looking at candidates' income. Besley et al. (2011) exploits random leadership transitions due to natural death or terminal illness of around 120 political leaders and find that leaders' educational attainment matter for countries' economic growth.

Thanks to the large Italian data set I observe more than 18,000 mayoral elections, 1,211

of which with a margin of victory of less than 5%. This allows to use the optimal bandwidth of Calonico et al. (2014a) as well as a conservative bandwidth of 5% for the regression discontinuity design.

I find that more educated mayors (defined as holding a university/technical degree or higher) boost public investment, particularly in the education sector, after the election and reduce other current related expenditure types. Electing by chance an educated mayor increases public investment by 3 percentage points of total expenditure as opposed to electing her less educated counterpart. Educated mayors increase investment in education by 1 percentage point of total expenditure, almost one-third of the total investment increase. Immovable assets dedicated to nurseries are the sub-category of investment that receive the biggest increment of resources. Furthermore, the increase of public investment is not performed at the expense of fiscal stability, indeed, the fiscal deficit does not result to be statistically different from zero.

As a result, the promotion of an educated political representation may help shift important resources toward long-term investments in education and early schooling facilities. This would not only shift political attention toward infrastructural projects but also enhance access to early schooling services.

2.2 Theoretical considerations

The presented model has its roots in the probabilistic voting model introduced by Lindbeck & Weibull (1987) and builds upon the Persson & Tabellini (2002) version employed to explain public expenditure. The model departures from the probabilistic voting model of Persson & Tabellini (2002) in two ways. Firstly, public expenditure is divided into two elements: public investment (g_I) and current expenditure (g_C). Secondly, it includes a component of competence (e) on which the two candidates can differ. Since it has been shown that education is a key determinant of adult competencies (Becker 1964, OECD & Canada 2000, Desjardins 2003, Reder 2008, Park & Kyei 2011, OECD 2013), I will use formal qualifications as proxy for competence and skills.

The model starts from the idea that public investments represent a more challenging task to perform than current expenditure, hence politicians' competence plays a role in its implementation. Surveys among U.S. and Italian mayors depicts increases in investment as one of the most desirable and, however, challenging task to achieve as it involves more planning, presentation of detailed projects, and dealing with intermediary bodies (Einstein et al. 2018, Einstein & Glick 2017, Carreri 2019).

The model is populated with two types of players: voters and candidates. Candidates

are opportunistic, without partisan preferences, and can commit themselves to an electoral platform that will be implemented as an economic policy program in case of electoral victory. Voters choose the candidate to vote for depending on both economic policy announcements and each candidate's ideology. The absence of partisan preferences shall not be considered a drawback of the model, on the contrary, it represents a useful feature. By allowing those candidates to care only about winning the election, it is depicted a mechanism that avoids valence statements. Indeed, it would be convenient to argue that educated politicians highly regard public investment over current expenditure because of personal preference. However, this is not the channel that will be explored. Conversely, a more conservative approach where all candidates' choice over the policy platforms is exclusively driven by electoral interest is chosen.

Finally, despite the credible commitment assumption may present some degree of unrealism, it is encouraging to observe that local elections usually present a higher degree of political accountability and fidelity to the political proposals than the national ones (Seabright 1996, Besley & Coate 2003, Persson & Tabellini 2002).

As in the baseline model, society is populated by different citizens indexed by i. The population is divided into three distinct groups, $J = P, M, R$, representing the poor, middle class and rich, respectively. Every citizen i, in each group J, has the same income y^J such that $y^P < y^M < y^R$. The population share of each group is α^J, with $\sum_j \alpha^J = 1$. y^i is distributed in the population according to a step function with $F(.) = 0$ for $y^i < y^P$, $F(.) = \alpha^P$ for $y^P \leq y^i < y^M$ and so on. The income per capita y is the expected value of individuals' income: $y = E(y^i)$. All voters i have the same quasi-linear preference over private consumption c^i and total public expenditure g_I, g_C.

$$w^i = c^i + H(g_I, g_C) \tag{1}$$

Individuals' utility gain from public expenditure $H(g_I, g_C)$ is represented by a concave function, increasing in g. Public investments g_I and current expenditure g_I, however, differ in the mechanism through which tax-dollar expenditure is converted into citizen utility.

$$H(g_I, g_C) = ln(g_C) + \frac{1}{\gamma} \cdot e \cdot ln(g_I) \tag{2}$$

Engaging in public investment is a difficult task, hence its outcome is competence dependent. Politician competence (e) can be seen as a technology for converting tax dollar investment expenditure into citizens' utility. On the other hand, current expenditure is seen as an in-kind monetary transfer, utility gain has no connection with candidate competence. Finally, given the forward-looking nature of investment spending, part of its outcome will

be realized in the future, hence the immediate utility gain from public investment is subject to a temporal discount factor $\gamma \in [1, \infty]$.

Given the delineated design, for the same level of competence, current expenditure results to be electorally more convenient than public investment. This feature makes the model in line with the related literature on public spending composition of Kennett (1990), Schuknecht (2000), Block (2001), Vergne (2009), Katsimi & Sarantides (2012), Gupta et al. (2016), where electoral incentives are stated to induce shifts of public spending toward more visible government consumption and away from long-term public investment.

At the time of election[1], voter i, in group $J \in P, M, R$, prefers candidate A over candidate B if:

$$W^J(g_I^A, g_C^A) > W^J(g_I^B, g_C^B) + \sigma^{i,J} + \delta \tag{3}$$

The timing of events is as follow:

1. The two candidates announce, independently and simultaneously, their electoral platforms (g_I^A, g_C^A) and (g_I^B, g_C^B). When announcing the platforms, candidates know voters' policy preferences and the distribution of the ideological bias $\sigma^{i,J}$ and relative popularity δ, but not their realized values. Competence e is common knowledge.

2. Before an election, a popularity shock δ may occur. Thus, from the perspective of the candidates, the electoral outcome can be regarded as a random event.

3. Elections are held and voters choose their favourite candidate.

4. After the election, the winner implements her policy platform.

Voters care about candidates' competence only to the extent to which this is relevant to the policy proposal. On the other hand, politicians know their ability in succeeding in different tasks in case of victory and shape their political platform accordingly. Noticeably, the possibility that candidates will fake their level of competence during the electoral campaign by promising non-optimal political platforms is ruled out by the credible platform commitment assumption: candidates do not stand for platforms that they would not be able to perform in case of victory.

Hence, politician candidate A maximizes her probability of being elected (see in Appendix A: Derivation 1), subject to the government budget constraint, by choosing her platform $(g_I^A; g_C^A)$

---

[1]It is assumed that ideological bias $\sigma^{i,J}$ has a group specific uniform distribution on $\left[-\frac{1}{2\phi^J}, \frac{1}{2\phi^J}\right]$ with density ϕ^J and average ideology $\phi = \sum_J \alpha^J \phi^J$. The candidate relative popularity δ has a uniform distribution on $\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$ with density ψ.

$$\underset{g_I^A, g_C^A}{\text{Max}} \frac{1}{2} + \frac{\psi}{\phi} \left[\sum_J \alpha^J \phi^J \left[\left(\left(\frac{y^J}{y} \right) (y - (g_I^A + g_C^A)) + ln(g_C^A) + \frac{1}{\gamma} \cdot e^A \cdot ln(g_I^A) \right) - \right. \right.$$

$$\left. \left. \left(\left(\frac{y^J}{y} \right) (y - (g_I^B + g_C^B)) + ln(g_C^B) + \frac{1}{\gamma} \cdot e^B \cdot ln(g_I^B) \right) \right] \right]$$

subject to

$$ty = g_I^A + g_C^A$$

It can be shown (see in Appendix A: Derivation 2) that, at equilibrium, as candidate A competence e^A rises, public investment increases while current expenditure decreases. The same will occur, symmetrically, also for candidate B. Intuitively, the two candidates share the same concave preference, however, they can differ in the technology (e) used to convert tax dollars into expected votes. Hence, if they differ in their level of competence, they will end up formulating different optimal policy platforms. The equilibrium is such that both candidates converge to the same platform $(g_I^A, g_C^A) = (g_I^B, g_C^B)$ only if $e^A = e^B$.

$$g_I^A = \frac{ty}{\left(\frac{1}{\frac{1}{\gamma} \cdot e^A} + 1 \right)} \tag{4}$$

in equilibrium as $e^A \uparrow$, $g_I^A \uparrow$

$$g_C^A = ty - \frac{ty}{\left(\frac{1}{\frac{1}{\gamma} \cdot e^A} + 1 \right)} \tag{5}$$

in equilibrium as $e^A \uparrow$, $g_C^A \downarrow$

An educated candidate finds public investments comparatively more convenient for electoral purposes than a less educated one. Given the diminishing marginal return nature of both current expenditure and public investment, both educated and uneducated candidates will allocate part of their platform to both goods. The following main proposition can be made:

Proposition 1. *More educated candidates will engage in higher public investment than less educated ones.*

16

2.3   Institutions and data

2.3.1   Institutional framework

Each Italian municipal government is composed of an elected mayor, an executive body (*Giunta comunale*) appointed by the mayor, and an elected city council (*Consiglio comunale*). In 2015, there were 8.046 municipalities in Italy, although this number has changed slightly over the years because of mergers and separations. Mayors are elected through a majoritarian electoral rule and enjoy a considerable amount of executive power with a substantial degree of discretion over the budget allocation (Bracco et al. 2019, Fabbrini 2001, Carreri 2019). Municipal governments manage around 10% of Italian public expenditure and oversee the provision of public goods and services to citizens, such as public transportation, waste management, water supply, infrastructure, construction and maintenance of school buildings, elderly care, public housing, and other related social services (Gagliarducci & Paserman 2012, Repetto 2018). Local expenditure is categorized as capital expenditure and current expenditure. Capital expenditure refers to investments in projects such as building new infrastructures, schools, hospitals, roads, etc, whereas current expenditure represents the operating costs incurred, in other words, wages and utilities. Municipalities finance their expenses by levying taxes (on income, real estate, and services) and through regional or national transfers.

Since 1999, the Internal Stability Pact (*Patto di Stabilità Interno*) has imposed the principle of a balanced budget on each municipality. Deviations from a balanced budget towards a deficit are subject to sanctions. Overall, municipalities that have incurred deficits represent a minority.

The electoral rule applying to municipalities with fewer than 15,000 inhabitants is "first past the post" (majoritarian with one round), whereas in more populous cities the electoral rule is "second round runoff" (majoritarian with two rounds). The vast majority of the municipalities fall into the first declination, as only a few hundred cities exceed 15,000 inhabitants. Municipal elections select both mayor and city chancellors and are held every 5 years on a Sunday between May and June. During municipal elections in smaller cities, each candidate running for mayor is linked to a list of candidate councillors. The most voted mayor candidate is elected to the office and the corresponding list obtains two-thirds of the city council seats while the remaining positions are distributed proportionally among the other lists. In the most populous municipalities, the mayor candidate who obtains either the absolute majority in the first round or more votes in the second round is elected mayor. At least three-fifths of the seats are assigned to the coalition of the elected mayor within the city council. Consequently, the majoritarian design of Italian local electoral rule allows for

a significant council majority of the elected mayor even in the presence of close victories.

Cities exert forces that shape the political framework of a country in profound ways, hence the question of who leads cities, and how, could not be more important not only for the health and well-being of urban residents but also for the economy and society as a whole. Boosting public investments without undermining fiscal stability is not an easy task. In the context of city governance, qualitative research has highlighted some elements that stand out as critical for a mayor's effective governance. Firstly, the capacity to build, maintain, and activate networks. Secondly, the expertise to turn the city into a "learning organization"[2] that absorbs and digests information about the changing environment and the performance of city operations (Harkness 2017). To sum up: the capacity to plan, implement, evaluate, and innovate on programs, projects, and services (Peter et al. 1990, Garwin 1993).

Education background plays a crucial role in all of these capacities: I expect mayors' education to be a decisive factor in the ability and wiliness of engaging in higher levels of public investments.

2.3.2 Data

The research uses Italian municipal data from all available municipalities from 2000 to 2015. Since cities with more than 15,000 citizens have a different electoral system than the remaining majority, there is the risk that the likelihood of winning a contested election under a "first past the post" electoral rule differs from a "second round runoff" rule. In order to have all elections experiencing the same majoritarian electoral system I focus on municipalities with fewer than 15,000 residents. The large majority (around 90%) of Italian cities have less than 15,000 inhabitants.

The collected data can be grouped into four categories: politicians' characteristics, electoral results, municipality balance sheets, and socioeconomic controls. Among politicians' characteristics are included age, gender, place of birth, level of education, and professional background of all politicians that run for the position of mayor from 2000 to 2015. Electoral results include candidates' and political parties' performances at municipal elections from 2000 to 2015. The municipality balance sheets comprise all relevant information on the public finances and general state of the cities, from revenue sources and expenditure types to population, city surface and street length. I obtained the electoral and balance sheet data from the Italian Ministry of the Interior[3]. Additional geographical controls such as altitude,

---

[2]From book *The Fifth Discipline* of Peter Senge that popularized the term "learning organization" in the 1990s.

[3]Data on electoral results from the 5 out of 20 Italian special autonomy regions are not available. The 5 special autonomy region are Sicily, Sardinia, Trentino–Alto Adige, Friuli–Venezia Giulia, and Valle d'Aosta.

longitude and altitude of the municipalities have been retrieved from the Italian National Agency for New Technologies, Energy and Sustainable Economic Development[4] (ENEA).

The final sample consists of 6,016 municipalities (out of the 8,046 existing municipalities in 2015) for the years 2000-2015. I observe the results of 18,082 elections, 1,211 of which are close elections with a margin of victory of no more than 5% for the winner. Table 1 summarizes the number of municipalities, elections and unique candidates of the sample.

Table 1: Number of municipalities, elections and candidates

	Municipalities	Elections	Candidates
Full sample	6,016	18,082	26,238
Optimal bandwidth	2,998	4,176	7,198
5 % bandwidth	1,081	1,211	2,320

The table shows the samples (2000–2015) used in the regression discontinuity design (RDD). Only those elections in which the winning mayor and the best opponent differ in terms of education are counted within the bandwidth. Optimal bandwidth (CCT) of 21%.

The study focuses on elections in which the margin of victory is small and the winner and the best opponent have different levels of education. Previous studies did not directly observe such information and assessed the education level of each candidate indirectly from previous earnings or profession. I gather education attainments directly from the registry of the Minister of Interior which details the education level of the candidates from primary school to PhD. I define a candidate as educated if she holds a university degree or higher. Overall, 40% of the mayors hold at least a degree. Women are strongly under-represented, only 10% of majors are female. The average age is 49 years old. Table 2 presents an overall picture of the elected mayors' main characteristics.

---

[4]Ente per le nuove tecnologie, Energia e Ambiente (ENEA).

Table 2: Descriptive statistics of mayors characteristics

	Observations	Mean	SD	Min	Max
Mayor characteristics					
Mayor holding a degree	17477.0	0.4	0.5	0.0	1.0
Female mayor	17973.0	0.1	0.3	0.0	1.0
Mayor age	18079.0	49.6	10.2	19.0	86.0
Mayor political affiliation					
Mayor left wing	16548.0	0.1	0.3	0.0	1.0
Mayor right wing	16548.0	0.1	0.3	0.0	1.0
Mayor of centre party	16548.0	0.0	0.1	0.0	1.0
Mayor of civic party	16548.0	0.8	0.4	0.0	1.0

Note: The table shows the descriptive statistics of the mayors characteristics at the election years for all 6,016 municipalities of our full sample.

With regard to municipalities' public finances, figure 1 gives an overview of the financial status of municipalities over the sample period. Current expenditure typically represents the biggest expenditure item with an average value of 60% of total public expenditure. Public investments account for 26% of total expenditure, 6% for loan repayments, and 8% for external services' charges. Noticeably, public investments are consistently declining over time. On the revenue side, on average, tax revenue corresponds to 34% of total income, whereas the remaining part is financed through transfers, alienation of assets, and charges for services. Table 3 shows detailed municipality balance sheet descriptive statistics for the entire time period of the sample. While table 4 summarizes geographical and demographic municipality characteristics.

Figure 1: Evolution of municipalities revenues and expenditures
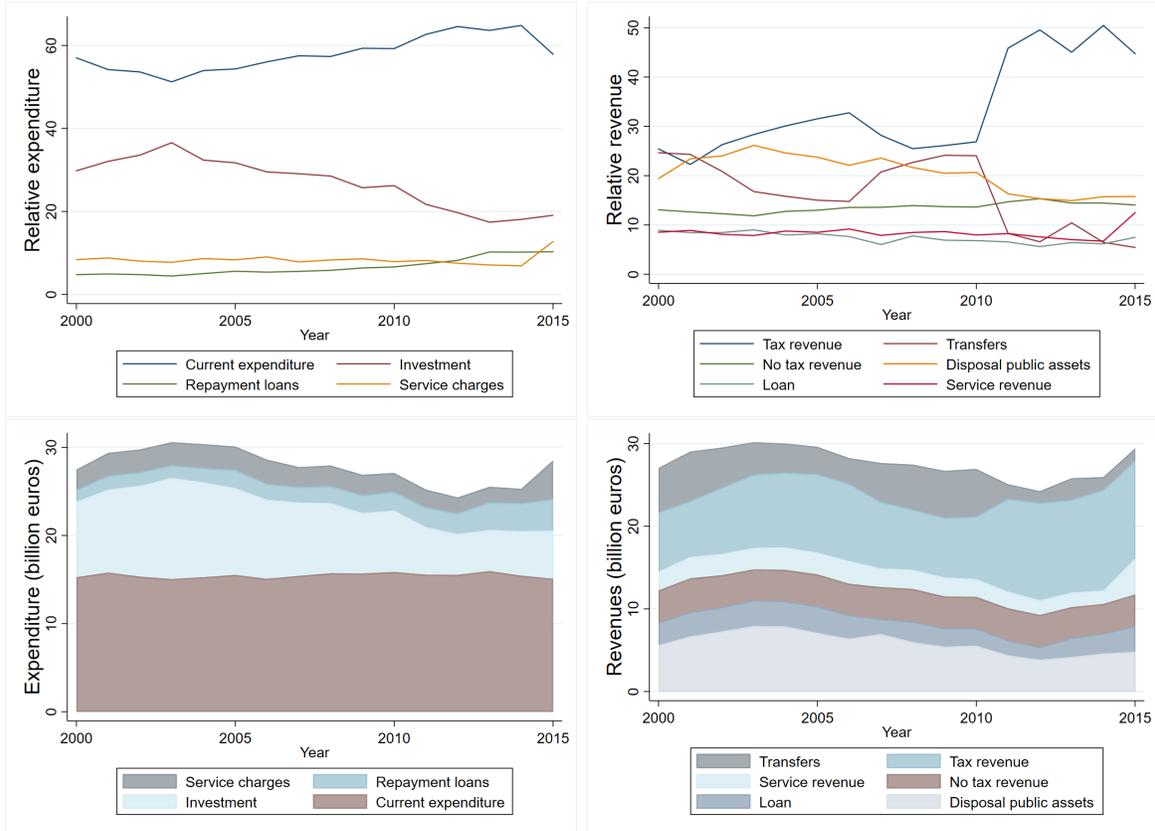
Note: The sample is composed of 6,016 municipalities, it excludes municipalities from special regions and with more than 15,000 population. Values expressed in billion of euros have been deflated using the St. Louis FED GDP deflator for Italy (2005 Euros). Values expressed in relative terms are percentages of total expenditure or revenue. The slight difference between revenues and expenses is due to the existence of unplotted balance sheet deficits or surpluses (the majority of municipalities have a balanced budget, with an average of 0.1% fiscal surplus).

Table 3: Descriptive statistics of fiscal data

	2000-2007	2008-2015	2000-2015
Balance sheet expenditures			
Current expenditure	54.8	61.2	58.0
	(15.4)	(16.2)	(16.1)
Repayment loans	5.1	8.1	6.6
	(5.8)	(9.7)	(8.2)
Service charges	8.4	8.4	8.4
	(5.8)	(5.7)	(5.7)
Investment	31.8	22.1	26.9
	(17.4)	(17.0)	(17.9)
Investment in education	2.5	2.2	2.3
	(4.8)	(5.1)	(5.0)
Balance sheet revenues			
Tax revenue	28.1	39.3	33.7
	(14.3)	(17.7)	(17.0)
No tax revenues	12.8	14.3	13.6
	(8.0)	(8.7)	(8.4)
Transfers	19.1	13.5	16.3
	(11.6)	(11.1)	(11.7)
Disposal public assets	23.4	17.6	20.5
	(17.0)	(15.7)	(16.6)
Loan	8.1	6.7	7.4
	(10.8)	(11.4)	(11.1)
Service revenue	8.5	8.4	8.4
	(5.8)	(5.7)	(5.7)
Observations	47,811	48,045	95,856

Note: The table shows the averages of fiscal data taken over the periods specified in the column headings for all 6,018 municipalities of our full sample (standard deviations in parentheses). All variables are expressed in relative terms. Relative expenditures are measured as: (Expenditure type)/(Total Public Expenditure) · 100. Relative revenues are measured as: (Revenue type)/(Total Public Revenue) · 100.

Table 4: Descriptive statistics of municipality characteristics

	Observations	Mean	SD	Min	Max
Geographic coordinates					
Latitude	17933.0	43.2	2.2	37.0	46.0
Longitude	17933.0	10.9	2.9	6.0	18.0
Altitude	17933.0	357.8	281.6	0.0	2035.0
Demographic characteristic					
Census population	17907.0	3359.5	3246.1	117.0	14996.0
Others					
Urban surface(ha)	15370.0	215.4	454.9	1.0	4700.0
Lenght internal roads(km)	15365.0	18.0	40.2	1.0	1034.0
Lenght esternal roads(km)	14913.0	51.7	106.0	2.0	3000.0

Note: The table shows the descriptive statistics of the municipality characteristics at the election years for all 6,016 municipalities of our full sample.

Since local elections are held between May and June and balance sheet data refer to the decisions made during the calendar year, the main RDD estimates are based on fiscal outcomes one year after the mayoral election. By doing so, I exclude the possibility that part of the estimated impact is due to the half-year of fiscal choices done by the previous mayor.

2.4 Empirical strategy

2.4.1 Baseline Specification

Identifying the effects of the educational background of politicians on policy choices is a complex task. Indeed, a simple comparison between municipalities with an educated mayor and those with an uneducated mayor will probably generate biased estimates due to endogeneity issues. There might be an underlying reason if a municipality has elected a more or less educated mayor, for example, bigger cities might elect more educated mayors than smaller cities. More in general, policy decisions might be correlated with municipality-specific characteristics such as attitudes toward politicians' personal backgrounds, all of which could also influence the education of the local mayor.

Another challenge that the research design must overcome is simultaneity or reverse causality. Indeed, it is complex to understand whether the educated politician is increasing investments or if such a politician is elected because of the rise in investments.

To address the described problem, a regression discontinuity design based on close elections is employed. Indeed, if the margin of victory of a candidate during an election is very small, then the results of a closely contested election can be taken as random. The formal argument for the validity of this approach is the continuity assumption as pointed out by Hahn et al. (2001) as well as Lee & Lemieux (2010). No shift in potential outcomes as we cross the threshold. In other words, the only reason that the outcome shifts at the threshold is the treatment. If reasonably close to the threshold, there are no omitted variables that are correlated with both the treatment status and the outcome. This implies that an election between a more educated candidate against a less educated opponent is likely to have been decided by random chance if the election is a very close race. The vote margin between the two candidates effectively determines the treatment (the education level of the mayor) randomly. Then, in close elections, municipalities that happen to have an educated mayor will be comparable to the ones without. Even though there is no way to directly test the identifying assumption (Lee 2008), I will perform different approaches to indirectly verify its validity.

The paper will compare municipalities in which a politician holding a degree barely defeats a politician without a degree, to those municipalities where a politician with a degree barely loses. This comparison, with respect to some particular outcomes of interest (e.g. municipality public investment), measures the causal effect of politician education on this outcome.

For the regression discontinuity to be valid, we must consider only observations in which the margin of victory is small, and the winner and the best opponent have different levels of education. This feature makes the RDD a very data-demanding empirical strategy.

The regression equation is as follows:

$$Y_{m,t+1} = \alpha + \beta edu_{m,t} + \gamma X_{m,t} + f(marg) + \varepsilon_{m,t} \tag{6}$$

$Y_{m,t+1}$ is a policy outcome of interest after the election, α is the constant, $edu_{m,t}$ is a dummy variable that takes the value of one if a politician holding a degree win a mayoral election, $X_{m,t}$ is a set of controls, $f(marg)$ is a polynomial of the margin of victory that captures the impact of the vote margin on the policy outcome and $\varepsilon_{m,t}$ is the error term. β indicates the causal effect of the presence of a more educated mayor on the outcome of interest.

A key choice that such research design requires is the definition of the RDD bandwidth to use around the threshold. The bandwidth sets up to which point an electoral victory can be defined as contested. Moving close to the threshold is a trade-off between identification and statistical power. The closer to the threshold we are, the more certain the covariates are essentially the same, on the other hand, we will also have fewer observations. Further

from the threshold, we have more observations, however, covariates might start differing. For this reason, the related literature generally uses a diversified set of bandwidths. In practice, different bandwidths are applied to show that the effects are not sensitive to the precise choice of the length of the bandwidth. Data-driven optimal bandwidth selection procedures, such as the one developed by Calonico et al. (2014b), are also widely employed. Freier & Thomasius (2016) defines close elections as those in which the winner beats the runner by up to 10% or fewer votes. I employ optimal bandwidth following Calonico et al. (2014a) as well as manual bandwidths of 5%. As a robustness check, I also replicate the main result with all 160 possible bandwidths from 0.25% to 40% as well as with all 160 possible thresholds from -40% to +40% with a small 3% bandwidth.

2.4.2 Validity of the RD Design

For the identification strategy of the regression discontinuity design to be valid, its continuity assumption needs to be holding: no shift in potential outcomes as we cross the threshold. An important feature of any RD Design is that such an assumption can be empirically verified through a series of tests.

Discontinuity in Density. It checks whether there is a discontinuity in the density of the running variable at the threshold. If there is a discontinuity, it could be argued that politicians are able to sort themselves above and below the running variable, calling into question the assumption of local randomization. Figure 4 shows the histogram of the running variable, while figure 5 presents the rddensity test Cattaneo et al. (2018, 2019) to investigate this issue. I find no significant discontinuity in the density of the margin of victory at the threshold.

Municipality characteristics balance. To confirm the success of the local randomization, I estimate the treatment effects on numerous placebo outcomes that are predetermined and should not be affected. Municipalities where a more educated mayor barely won or barely lost the election should not be statistically different in their predetermined characteristics. Table 9 indicates no significant discontinuity in any municipality characteristics.

Lagged placebo outcome. Another way to verify the success of the local randomization is to estimate the treatment effects on the lagged placebo outcome. Since the treatment effect cannot have a retroactive effect, if the quasi-randomization process has been successful, we should not detect a statistically significant treatment effect on the lagged placebo outcome. Table 10 shows no significant discontinuity in our main lagged outcome variable.

Placebo thresholds test. It checks whether the treatment effect is present excursively at the true threshold of 0% margin of victory. If the treatment effect is found at a fake threshold, it raises questions about the validity of the RD design. Figure 6 shows that there is no significant treatment effect at all 159 possible fake thresholds from -40% to +40% margin of victory.

2.4.3 Robustness tests of the RD Design

Bandwidth sensitivity test. It checks that the RDD estimates are not sensitive to the choice of a particular bandwidth. Figure 3 shows the treatment effects with all 160 possible bandwidths from 0.25% to 40% margin of victory.

Kernel and Polynomial sensitivity test. It checks that the RDD estimates obtained using the data-driven optimal bandwidth are not sensitive to the choice of a particular kernel function or polynomial for point estimation. Table 11 shows the main RDD results with triangular, epanechnikov and uniform kernel functions, as well as with first, second and third-order polynomial for point estimation.

Confoundedness with Mayor characteristics. A further concern is the confounding effects due to the identity of mayors. More and less educated mayors may systematically differ in some personal characteristics, such as gender or political affiliation. If this was the case, it could be that I am measuring the impact of political affiliation and not of the mayor's education. Table 12 and 13 present a series of mayors' characteristics bundling tests. Mayors' education seems to be the driving factor of the treatment effect.

2.5 Results

2.5.1 Effect of mayor education on education public investment

Table 5 presents the main results. Column 1 displays the result for relative investment in education using optimal bandwidth and local polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed in Calonico et al. (2014a) as well as the results using manual bandwidths of 5% and local linear control functions. Columns 2 and 3 replicate the same results by adding year fixed effect and regional fixed effect. Despite the addition of fixed effects, this is not strictly required by the identification strategy, it is a good sign to see that the findings are not sensitive to the inclusion of them.

Overall, estimates suggest that electing by chance an educated mayor increases public education investment by 1 percentage point of total expenditure as opposed to electing her less educated counterpart. This finding is statistically significant at 1% level using the optimal bandwidth and at 5% level in the other specifications.

Table 5: RDD estimates: the impact of mayor's education on education public investment

	Relative education public investment		
	(1)	(2)	(3)
RDD result using optimal bandwidth	0.910***	0.934***	0.949***
	(0.340)	(0.332)	(0.334)
	[6,663]	[6,663]	[6,663]
RDD result using 5% bandwidth	1.146**	1.076**	1.075**
	(0.493)	(0.499)	(0.498)
	[1,208]	[1,208]	[1,208]
Year FE	No	Yes	Yes
Region FE	No	No	Yes
Running variable	MV educated	MV educated	MV educated

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: educated margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employ local linear control functions. The outcome variable is one year after the elections.

The results can be visually inspected in figure 2. At the 0% threshold of the educated mayors' margin of victory, it can be observed how relative public investment in education sharply jumps from around 2% to about 3% of the total public expenditure.

Figure 2: RD plot for public investment in education
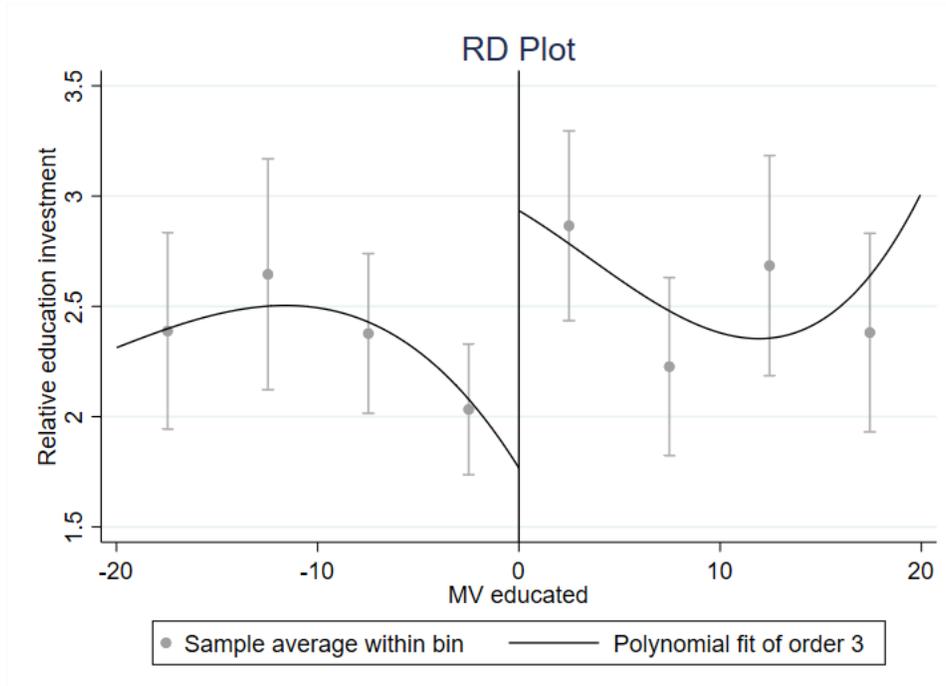
Note: IMSE-optimal evenly-spaced method using spacings estimators to select the number of bins, 95% confidence interval.

To check how this main result is sensitive to the choice of the bandwidth, I perform a bandwidth sensitivity test. To do the bandwidth sensitivity test I conduct a set of RDD estimates increasing the bandwidth length in increments of 0.25 pp, starting from 0.25% to 40%. Figure 3 shows the results employing a Local linear control function on public education investment both in relative and euro per capita terms. As expected, RDD estimates start insignificant with the smallest of the bandwidths, and as soon as statistical power increases (due to the rise in the identifying observations) results become statistically significant and steady.

Figure 3: Bandwidth sensitivity test
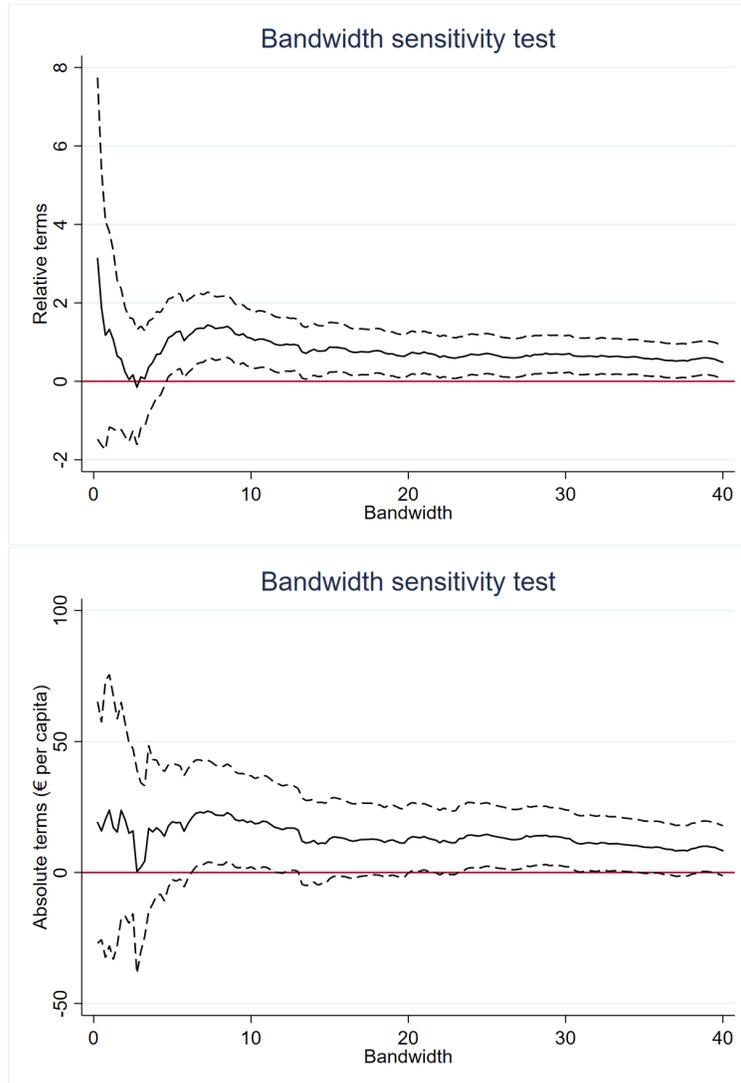
Note: 160 RDD estimations using Local linear control function and increasing bandwidth, move in increments of 0.25 pp. 95% confidence interval. Outcome variable: public investment in education. Values expressed in euro per capita have been deflated using the St. Louis FED GDP deflator for Italy (2005 Euro).

2.5.2 Effect of mayor education on public education investment by sub-component

I dig deeper into the municipal balance sheets and look into the different sub-components of investment in education. Table 6 shows the impact of mayoral education on the sub-components of investments in education. Electing by chance an educated major boosts investments in nurseries by around 0.35 percentage points of total public expenditure and in primary education by around 0.5 percentage points of total public expenditure. Furthermore,

most of the investments go into immovable assets.

Research in cognitive science and economic development has built a broad consensus on the relevance of early-age schooling and nutrition on cognitive development and future earnings (Dillon et al. 2017, Duflo 2001, Duncan 2003, Glewwe & King 2001, Whaley et al. 2003). Elango et al. (2015) find that disadvantaged children benefit the most from public investment in early childhood education and that society receives a higher return, which, in turn, exceeds the costs. Hence, such a shift of resources on early schooling infrastructures made by educated mayors might have desirable long-term consequences for the local community.

Table 6: RDD estimates: the impact of mayor's education on education investment by type

	Nurseries		Primary education		Middle education	
	(1)	(2)	(3)	(4)	(5)	(6)
RDD result (immovable assets)	0.322**	0.356*	0.490*	0.432	0.121	0.238
	(0.151)	(0.188)	(0.275)	(0.375)	(0.160)	(0.231)
RDD result (movable assets)	0.00427	0.00373	-0.0218	-0.0217	0.00592	0.0138
	(0.00720)	(0.00888)	(0.0278)	(0.0359)	(0.00650)	(0.00873)
RDD result (total)	0.363**	0.394**	0.562**	0.476	0.0879	0.228
	(0.154)	(0.190)	(0.282)	(0.378)	(0.161)	(0.233)
Observations	6,663	1,208	6,663	1,208	6,663	1,208
Bandwidth	Optimal	5%	Optimal	5%	Optimal	5%
Running variable	MV educated	MV educated	MV educated	MV educated	MV educated	MV educated

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Running variable: educated margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employ local linear control functions. The outcome variable is one year after the elections.

2.5.3 Effect of mayor education on budget composition

The economic relevance of such a boost in education investment makes it wonders how this increase is financed. Here I investigate the impact of electing a more educated mayor on the different budget components; expenditure composition, revenue composition, and on budget surplus.

Table 7 shows the impact of electing an educated mayor on the municipality expenditure composition. Overall, electing an educated mayor makes public investment rise by about 3 percentage points of total expenditure. This result is statistically significant at 5% level using both optimal and manual bandwidth. On the other hand, all the other types of public expenditures show individually insignificant, however, negative estimates, and jointly statistically significant negative estimates.

30

Table 7: RDD estimates: the impact of mayor's education on expenditure composition

	Investment (1)	Current exp (2)	Service charges (3)	Repayment loans (4)	Sum (5)=(2)+(3)+(4)	Surplus (6)
RDD result using optimal bandwidth	2.952**	-1.090	-0.429	-0.647	-2.897*	-0.00845
	(1.496)	(1.278)	(0.495)	(0.813)	(1.488)	(0.00655)
	[6,663]	[6,663]	[6,663]	[6,663]	[6,663]	[6,658]
RDD result using 5% bandwidth	3.896**	-1.328	-1.221*	-1.382	-3.932**	-0.0125
	(1.967)	(1.862)	(0.722)	(1.110)	(1.965)	(0.00861)
	[1,208]	[1,208]	[1,208]	[1,208]	[1,208]	[1,208]
Running variable	MV educated	MV educated	MV educated	MV educated	MV educated	MV educated

Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: educated margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico at al(2014). RDD results with a bandwidth of 5% employ local linear control functions. The outcome variable is one year after the elections.

Table 8 depicts the same estimates on the municipality revenue composition. Overall it does not seem that educated mayors engage with the revenue side of the balance sheet in a significantly different way than the less educated mayors.

Table 8: RDD estimates: the impact of mayor's education on revenue composition

	Tax revenue (1)	Transfers (2)	Non tax revenue (3)	Public assets disposal (4)	Loans (5)	Services revenue (6)
RDD result using optimal bandwidth	0.617	-0.593	-0.983	0.764	0.666	-0.652
	(1.507)	(0.972)	(0.634)	(1.280)	(0.930)	(0.436)
	[6,663]	[6,663]	[6,663]	[6,663]	[6,663]	[6,663]
RDD result using 5% bandwidth	0.527	-0.300	-0.767	1.348	0.326	-1.170
	(1.989)	(1.291)	(0.899)	(1.824)	(1.383)	(0.716)
	[1,208]	[1,208]	[1,208]	[1,208]	[1,208]	[1,208]
Running variable	MV educated	MV educated	MV educated	MV educated	MV educated	MV educated

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: educated margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico at al(2014). RDD results with a bandwidth of 5% employ local linear control functions. The outcome variable of one year after the elections.

A relevant implication of these results is that the higher resources devoted to investments are not coming from a deficit or higher taxes but from budget reallocation. This means that educated mayors manage to boost public investment without endangering the health of the public finance of their municipality.

2.5.4 Validity of the RD Design

Discontinuity in Density.

The RDD analysis based on close elections crucially relies on randomized variations just around the threshold. This requirement does not hold in the presence of sorting around the threshold. For example, if a political candidate can manipulate a contested electoral outcome, causing sorting, the treatment would not only be determined by the forcing variable. Therefore, units on each side of the threshold would not be comparable as in the randomization case, taking us back to an endogenous world.

To test the validity of the regression discontinuity design I conduct several robustness tests. Since sorting is directly related to the idea of having more, or less, observations on either side of the threshold, I check this by visually scrutinizing the histogram of the forcing variable and checking that there are no spikes at the right or left-hand side of the threshold. I also more formally examine the sorting issue by performing the density test of the forcing variable as proposed by Cattaneo et al. (2018, 2019). The test uses a nonparametric density estimator based on local polynomial techniques and, differently from the McCrary test (McCrary 2008), does not require pre-binning or any other transformation of the data. Figure 4 and 5 show no evidence of sorting.

Figure 4: Histogram around the threshold
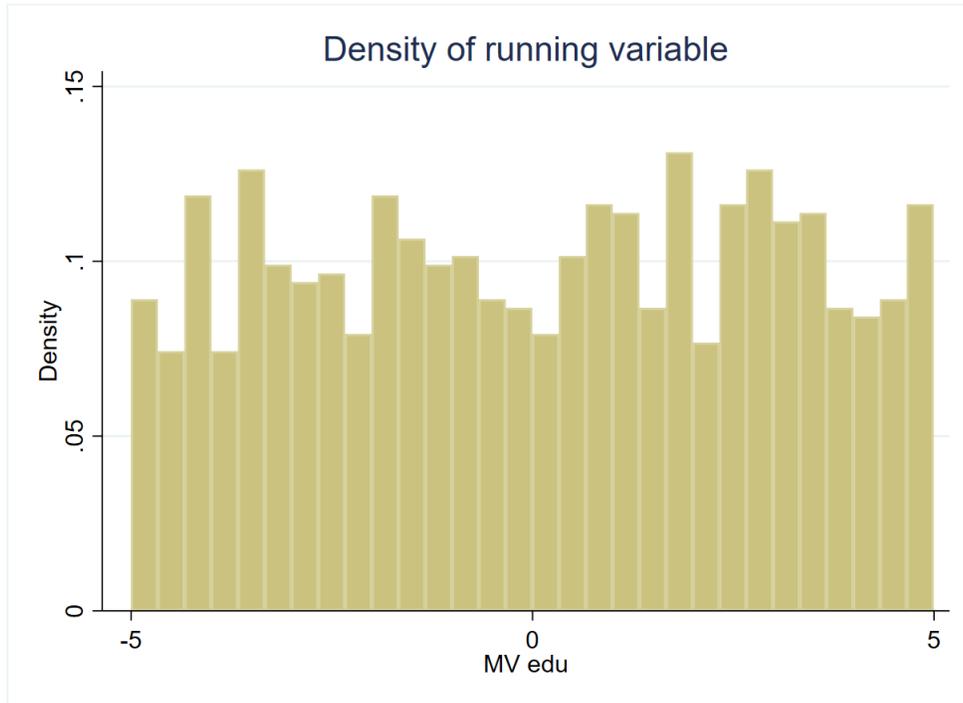
The histogram presents the frequencies of observations for the RDD analysis. The data are ordered by the assignment variable "educated mayor margin of victory" (or loss). The frequencies within a margin of victory of ± 5 percentage points are shown.

Figure 5: Density test
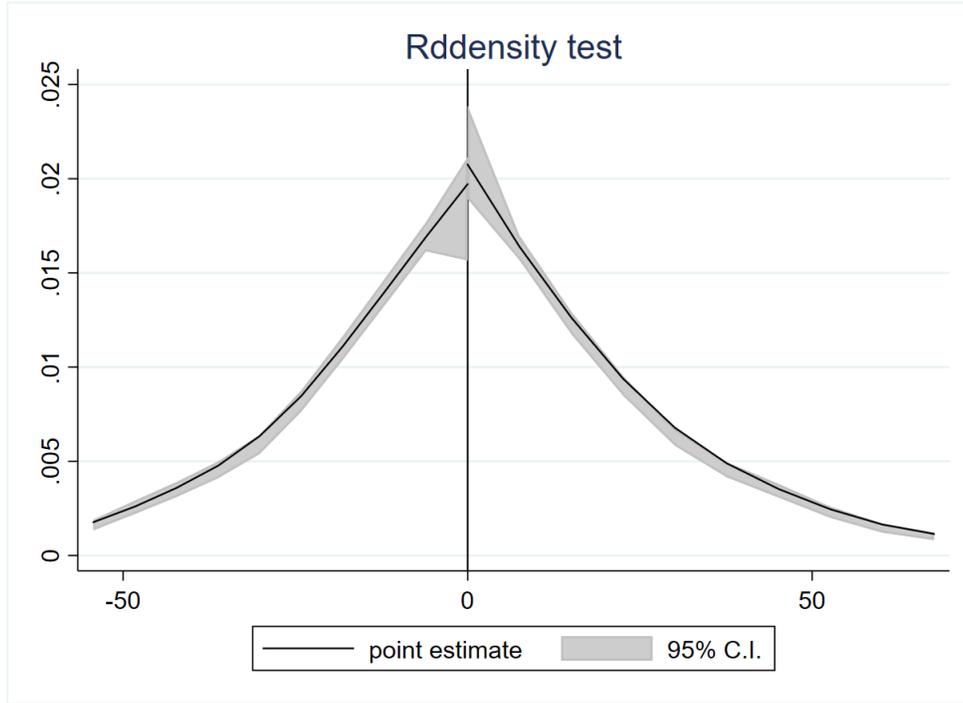
Note: The density test plots the density around the cutoff using the local polynomial density estimators as proposed by Cattaneo et al. (2018, 2019).

Municipality characteristics balance.

To confirm that the randomization holds, I also perform a municipality characteristics balance test. I run the same RDD models of my main specification on various demographic and geographical municipality-level characteristics that are predetermined and should not be affected by the treatment. Table 9 shows no significant impact for any of the municipal characteristics considered. This ensures the absence of any omitted variables, which could have survived in case of a failure of the quasi-randomization process.

Table 9: Municipality characteristics balance test

	(1)	(2)
Municipality placebo outcome	RDD result using optimal bandwidth	RDD result using 5% bandwidth
Latitude	-0.155	0.169
	(0.185)	(0.266)
	[6,765]	[1,233]
Longitude	0.245	-0.0441
	(0.238)	(0.338)
	[6,765]	[1,233]
Altitude	-32.50	-8.356
	(22.75)	(30.99)
	[6,765]	[1,233]
Census population	173.2	-190.7
	(265.8)	(391.2)
	[6,802]	[1,241]
Urban surface(ha)	-10.86	-33.46
	(37.41)	(61.19)
	[5,692]	[1,019]
Length internal roads(km)	3.596	-2.134
	(3.021)	(3.277)
	[5,692]	[1,019]
Length external roads(km)	9.196	1.989
	(9.629)	(16.11)
	[5,527]	[988]
Running variable	MV educated	MV educated

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: educated margin of victory. Threshold=0%. Placebo variables are either steady or taken in the year of the elections.

Lagged placebo outcome.

Since the increase in investment is caused by the victory of the most educated mayor, I should not observe any statistically significant difference in public investments between the cities that will have an educated mayor in the upcoming year and those that will not. In other words, the boost in investment can only be caused by the treatment and not pre-existing.

Table 10 highlights the estimated effects on the lagged outcome variable: public education investment made in the years prior to the election. Again, it is reassuring not to detect any statistically significant result.

Table 10: Lagged placebo outcome RDD estimates

	Relative education investment	
	(1)	(2)
RDD result using optimal bandwidth	-0.399	0.566
	(0.389)	(0.395)
Years before the election	2	1
Running variable	MV educated	MV educated
Observations	6,921	7,041

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Running variable: educated margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico at al(2014). Lagged placebo outcome of one/two years before the elections.

Placebo thresholds test.

To rule out the possibility that the main result arises from random chance rather than a real causal relationship, I perform a threshold placebo test. I do this by examining if there are any jumps in the relationship between the forcing variable and the outcome variable. To do the threshold placebo test I conduct a set of RDD estimates by shifting the threshold of 0.25 pp, starting from -40% to +40% (being 0% the correct threshold). In all the 160 estimates a small-moving 3% bandwidth is employed (see figure 6). Overall, almost all the placebo thresholds show statistically insignificant estimates. As expected, at threshold=0% the estimated impact of mayor education on public investment is shown positive and statistically significant at 5%.

Figure 6: Placebo threshold test
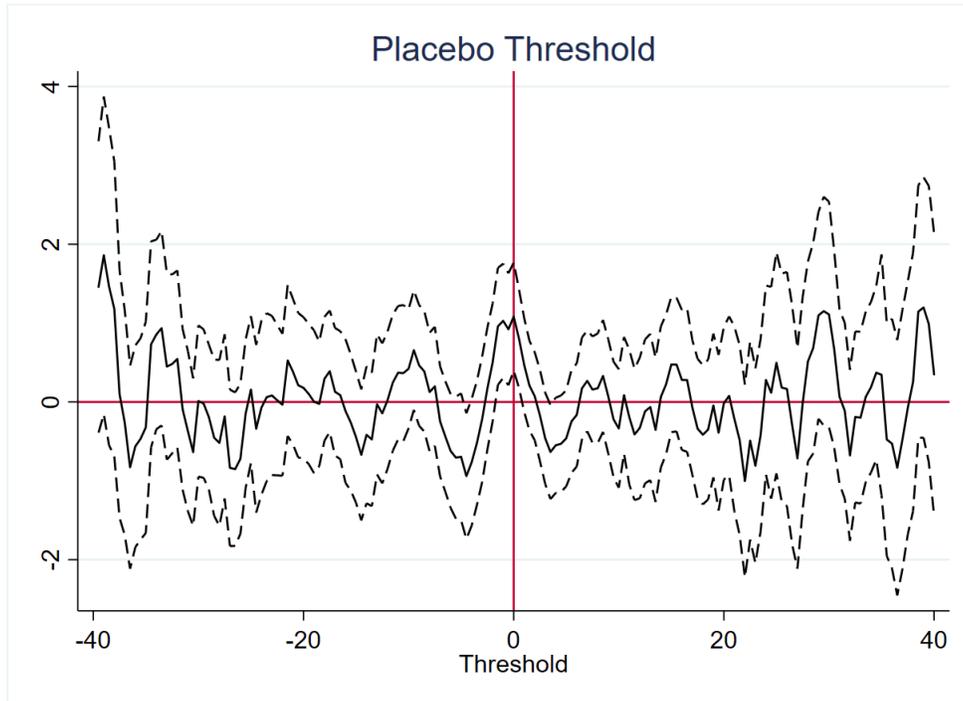
Note: 160 RDD estimations using a moving threshold of 3%, shifts in increments of 0.25 pp. 95% confidence interval. Outcome variable: Relative public investment in education.

Kernel and Polynomial sensitivity test.

Typical practice in applied research is to employ a local regression estimator, with local linear being the most common. Calonico et al. (2014*a*) have taken the polynomial order as given and the first order of the local-polynomial to construct the point-estimator by default (local linear). Furthermore, different kernel functions can be used to construct the local-polynomial estimators, with the triangular kernel being the most popular. In Table 11, I display Calonico et al. optimal bandwidth results with first, second and third-order polynomial specifications as well as using different kernel functions.

It is reassuring to observe that the impact on education public investment is not sensitive to kernel function or local-polynomial specifications. In particular, higher levels of local Polynomial for point estimator result in higher magnitude and statistical significance of the RDD estimates.

Table 11: RDD estimates: kernel function and order local-polynomial sensitivity test

	Relative education investment			Relative public investment		
	(1)	(2)	(3)	(4)	(5)	(6)
RDD result (triangular kernel)	0.910***	1.071**	1.339***	2.952**	3.946**	4.460**
	(0.340)	(0.423)	(0.501)	(1.496)	(1.757)	(2.001)
	[6,663]	[6,663]	[6,663]	[6,658]	[6,658]	[6,658]
RDD result (epanechnikov kernel)	0.869**	1.019**	1.386***	2.897*	3.795**	4.446**
	(0.337)	(0.419)	(0.508)	(1.495)	(1.708)	(2.009)
	[6,663]	[6,663]	[6,663]	[6,658]	[6,658]	[6,658]
RDD result (uniform kernel)	0.816**	0.919**	1.401***	2.967**	3.971**	3.777*
	(0.347)	(0.432)	(0.502)	(1.499)	(1.716)	(2.043)
	[6,663]	[6,663]	[6,663]	[6,658]	[6,658]	[6,658]
Bandwidth	Optimal	Optimal	Optimal	Optimal	Optimal	Optimal
Polynomial for point estimator	1	2	3	1	2	3
Running variable	MV educated	MV educated	MV educated	MV educated	MV educated	MV educated

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: educated margin of victory. Threshold=0%.Kernel function used to construct the local-polynomial estimators are triangular, epanechnikov and uniform. RDD result using optimal bandwidth employs first, second or third-order local-polynomial used to construct the point-estimator with robust bias-corrected confidence intervals and inference procedures as developed in Calonico et al(2014).

Confoundedness with Mayor characteristics.

With regard to the presence of idiosyncrasy of the treatment, it has to be noted that the treatment is "electing an educated mayor" and not the exclusive impact of education on politicians' policy choices. Indeed, the described regression discontinuity design randomizes the municipality, not the politicians. Hence, I measure the impact of electing an educated politician, with all the attached individual characteristics that usually come with education. The treatment is a bundle of elements composed of the education of the politician and other elements characterizing educated politicians. In this regard, Marshall (2022) highlights the importance of checking for bundling effects when performing politician characteristic regression discontinuity (PCRD) designs. To investigate whether other personal elements are affecting the results, I perform a mayor characteristics bundling test. I run the same RDD models as above on personal characteristics such as gender, age and political affiliation of the winning mayors. It is important to state that the test does not validate or invalidate the regression discontinuity design, on the contrary, it simply checks for the presence of bundling into the treatment.

Table 12 depicts the co-occurrences of mayors' characteristics with education level. It is reassuring to observe that no bundling is present for political views. However, educated mayors are younger and more often females.

Table 12: Mayor characteristics bundling test

Mayor placebo outcome	(1) RDD result using optimal bandwidth	(2) RDD result using 5% bandwidth
Female	0.0860***	0.126***
	(0.0227)	(0.0387)
	[6,639]	[1,206]
Age	-3.662***	-4.298***
	(0.825)	(1.239)
	[6,679]	[1,211]
Left wing party	0.0370	0.0215
	(0.0246)	(0.0329)
	[6,126]	[1,103]
Right wing party	0.0263	0.00925
	(0.0262)	(0.0357)
	[6,126]	[1,103]
Centre party	-0.00573	-0.00594
	(0.0103)	(0.0172)
	[6,126]	[1,103]
Civic party	-0.0564	-0.0248
	(0.0375)	(0.0482)
	[6,126]	[1,103]
Running variable	MV educated	MV educated

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: educated margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employ local linear control functions. Placebo variables of one year after the elections.

To verify that the effect of education is not driven by age or gender, I add age and gender as control variables, exclude female mayors from the sample, and see if there is any change in the estimated effect (table 13). All estimations show minor changes in the estimated effect, sometimes increasing in both magnitude and statistical significance with respect to the baseline result, confirming that neither age nor gender is driving the result. Furthermore, in figure 7, I restrict the sample to a moving 20 years age window and I run 30 RDD estimations by shifting the 20-length age window in increments of 1 year each time. Even though at the mean of the age distribution these estimations rely on half of the original

observations, the treatment effect is still present thought almost all the distribution.

Table 13: RDD estimates controlling for co-occurrence mayors' characteristics

| | Relative education investment | | Relative public investment | |
	(1)	(2)	(3)	(4)
RDD result	0.925***	1.087**	2.927**	4.007**
(mayor gender and age as control)	(0.338)	(0.491)	(1.478)	(1.992)
	[6,625]	[1,203]	[6,625]	[1,203]
RDD result	0.985***	1.175**	3.398**	4.352**
(mayor gender and age dummies as control)	(0.348)	(0.513)	(1.485)	(2.031)
	[6,625]	[1,203]	[6,625]	[1,203]
RDD result	0.877**	1.114**	2.911*	3.710*
(only male mayor sample)	(0.384)	(0.549)	(1.657)	(2.175)
	[5,834]	[1,042]	[5,834]	[1,042]
Bandwidth	Optimal	5%	Optimal	5%
Running variable	MV educated	MV educated	MV educated	MV educated

Note: Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: educated margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employ local linear control functions. The outcome variable is one year after the elections.

Figure 7: Age bundling window test
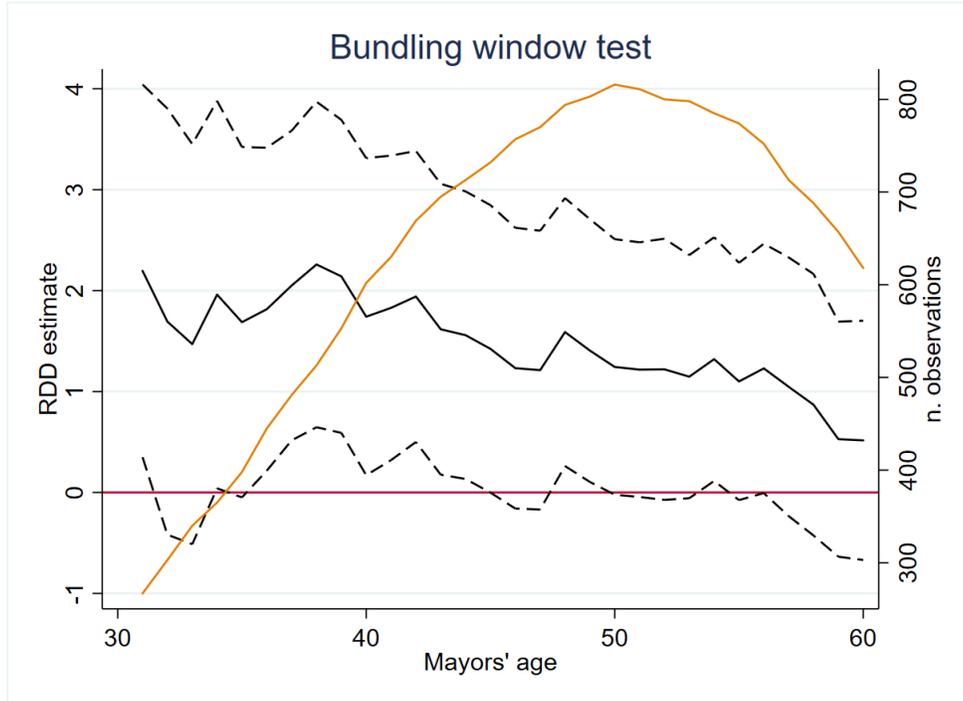
Note: I estimate 30 regression discontinuity design (RDD) models using a local linear control function approach, with a threshold of 5% and a shifting age window of 20 years. Specifically, I shift the window by 1 year from 30 to 60 years old, resulting in 30 different subsamples. I report 95% confidence intervals for all estimations. The outcome variable is the relative level of public education investment. The yellow line and the right y-axis indicate the number of observations used in each RDD estimation.

2.6 Conclusion

The purpose of this paper is to evaluate the impact of mayors' education on policy choices and public budget allocation over investments. The theoretical hypothesis is that public investment if compared with other kinds of expenditures, requires a higher level of competence to be successfully performed. As a consequence, more educated mayors will shift their political platforms toward public investments at the expense of other expenditure types. On the contrary, less educated mayors will find it convenient to rely more heavily on current expenditure-based programs. The presented mechanism is formalized using a simple probabilistic voting model. The model departures from the well-known version of Persson & Tabellini (2002) by splitting public expenditure into public investments and current expenditure, as well as by including a component of education on which the politicians can differ.

The theoretical model predicts that candidates will shape their political platforms in

order to best fit their level of competence. To maximize the probability of victory, educated candidates will find it convenient to shift their platforms toward public investment, whereas less educated ones will find it advantageous to do the opposite.

I use regression discontinuity design based on close election to identify the causal effect of politician education on policy choices. I find evidence that educated mayors choose to increase public investments after the election. This increase is realized at the detriment of other kinds of expenditures and without endangering the fiscal stability of the municipality. I also unfold the effect on its sub-components and uncover that investment in education is the driving factor of the total investment increase. Educated mayors focus their investment efforts on schooling infrastructure such as nursery and primary education facilities.

The results are robust to controlling for election-specific factors and using alternative samples, functional forms and bandwidths. Furthermore, all validity tests suggested by the literature are performed reassuring on the validity of the identifying assumption of the regression discontinuity design.

To conclude, given the rise of populism, the prominent use of current expenditure as a political tool and the lack of public investments in core sectors such as education and schooling, the promotion of an educated political representation becomes a key counterbalance element.

Chapter 3

The impact of women's political leadership on institutional quality

3.1 Introduction

A growing branch of the literature is studying the impact of women's participation in politics on policy choices and societal outcomes. Gender quotas and the election of women in leadership roles have been found to boost future women's political participation and success, as well as to improve education and health provision, while reducing corruption (Cowper-Coles 2020). Yet, most of the existing empirical evidence focuses on developing countries such as India and Brazil where gender roles are stronger and improvement in women's participation in politics are more likely to generate relevant changes in the socio-political sphere (Hessami & da Fonseca 2020).

The main aim of this paper is to answer the following questions: Do female politicians improve future female participation in politics and employment in public administration? Do female politicians increase institutions' quality metrics such as corruption, government effectiveness, rule of law and political accountability? Do cultural framework conditions and women's education play a role?

The paper presents the first empirical evidence on the impact of women's political participation on different institutional quality variables (such as corruption, government effectiveness, regulatory quality, rule of law, and political transparency) in a Western democracy. Secondly, it provides further evidence of the causal impact of female political representation on future female political and work participation. Thirdly, it showcases how such causal finding interacts with the level of education of the female politicians and the gender norms of the local environment.

I match rich and detailed Italian data on mayoral candidate gender, mayoral electoral results, and municipality/provincial socio-economic data from 2000 to 2015. To overcome endogeneity issues, I rely on regression discontinuity design as well as an instrumental variable approach, focusing on close elections. The idea is that since politicians cannot perfectly predict the result of an election, the results of closely contested elections can be considered as good as random. The identification strategy quasi-randomizes municipality characteristics, allowing for a comparison of municipalities that differ only in the gender of their mayors with respect to some outcome variables. Female and male mayors are quasi-randomized to municipalities, and then I test and verify the identifying assumption that there is no difference

in traits around the threshold.

From a theoretical perspective, the impact of female political representation on different policy outcomes can be explained with the citizen-candidate model. Indeed, regardless of party affiliation, men and women are likely to present different policy preferences due to both social pressures and personal experiences. For example, female politicians may be more interested in tackling specific issues that, by nature or due to traditional gender roles, predominantly affect women (Chattopadhyay & Duflo 2004, Prindeville 2002, Kamlongera 2008), such as legislation on female genital mutilation, domestic violence, and childcare and maternity policies (Goetz 1998, Cowper-Coles 2020)

In this paper, I find that electing more women as mayors boosts future female participation in politics and public administration, as well as improves institutional quality, particularly in terms of the rule of law and political accountability. I also find that such effects are larger and more significant in communities where gender norms are more established. The presented findings are in line with the results of Decarolis et al. (2022), which found a negative correlation between female politicians and bureaucratic corruption using two distinct datasets from Italy and China. Additionally, Baskaran et al. (2018) found, using RDD, that Indian constituencies electing women experienced higher economic growth due to women legislators being less likely to be criminal or corrupt, more effective, and less vulnerable to political opportunism. Furthermore, Acconcia et al. (2021) found, employing difference-in-differences, that gender quotas decrease the probability of early termination of legislatures in Italy. Yet, to the best of my knowledge, no study on the causal impact of female political representation on institutional quality has been performed in developed Western democracies, and the evidence on corruption is limited to India and Brazil.

3.2   Literature Review

Different studies have found a positive impact of female political representation on future female participation in politics (Martin & Borrelli 2016, Atkeson 2003, Bauer & Okpotor 2013, Campbell & Wolbrecht 2006, Dassonneville & McAllister 2018, Fridkin & Kenney 2014, Gilardi 2015, Ladam et al. 2018, Liu & Banaszak 2017, Mariani et al. 2015, McAllister 2019, Brown et al. 2021, Maitra & Rosenblum 2021, O'Connell 2020, Bhalotra et al. 2018). For example, Brown et al. (2021) used RDD to find that the contested victory of seats by women at national elections in India increases female participation for seats at the next national election. Similarly, Maitra & Rosenblum (2021) and O'Connell (2020) used exogenous timing differences in the implementation of Indian local-level gender quotas and found significant positive impacts on the political participation of women both at the national and the state

level. At the same time, Bhalotra et al. (2018) used regression discontinuity design on Indian constituency-level data and found that such increases in the share of women candidates are primarily driven by an increased propensity of the incumbent woman to contest for re-election.

Extensive evidence also supports the idea that women politicians focus on tackling specific issues. For example, Clots-Figueras (2011) exploit contested elections in India and find that women politicians invest more in health and early education. Similarly, Bhalotra & Clots-Figueras (2014) find that an increase in women's representation in India results in a reduction in neonatal mortality. Clots-Figueras (2012) use close elections in India and find that increasing female political representation increases the probability that an individual will attain primary education in that district.

Atkeson & Carrillo (2007), Ulbig (2007), and Schwindt-Bayer (2010) all find a positive correlation between female representation and feelings of trust in government. However, Fox & Lawless (2004) does not. Adler (1996, 1997), Fraga et al. (2007), Holman (2015), Rosenthal (2000), Tripp (2001) find women's style of leadership to be more democratic, cooperative, and inclusive. Bratton & Haynie (1999), Bratton (2005), Jeydel & Taylor (2003) find no difference between female and male politicians in government effectiveness, while Anzia & Berry (2011) and Fraga et al. (2007) find that women politicians often introduce and pass more priority legislation. Norris (1996) and Carroll (2001) find women in local government to be more responsive to constituents.

Different studies found a relationship between having (more) women as representatives and lower levels of corruption at both local level (Bauhr et al. 2019, Chhoeun et al. 2008, Goetz 2002, Jayal 2006, Sundström & Wängnerud 2016, Beaman et al. 2007, Baskaran et al. 2018, Brollo & Troiano 2016) and national level (Bari 2010, Dollar et al. 2001, Jha & Sarangi 2018, Schwindt-Bayer 2010, Stockemer 2011). As an example, Beaman et al. (2007) exploit randomized mandated political representation for women in local governments in India to detect a negative impact of female political representation on corruption. Baskaran et al. (2018) and Brollo & Troiano (2016) use RDD on close elections, respectively in India and Brazil, finding that women legislators are less likely to be criminal and corrupt.

Women politicians may be less corrupt for different reasons. The first explanation is that they find it harder than men to infiltrate the networks used for corrupt practices (Bauhr et al. 2019, Bjarnegård 2013). Party biases also tend to place women in less strategic positions (Esteve-Volart & Bagues 2012). Second, greater gender equality and lower levels of corruption are both a result of democratization and development, so there may be a correlation rather than a causation (Alatas et al. 2009). Third, women in power may act to reduce corruption by actively targeting it and improving the provision of services. This is

because, unlike men in power, women are more aware of the consequences of their actions. Women can face a higher risk of operating outside of political norms or engaging in corrupt practices (Esarey & Chirillo 2013), with female voters holding female politicians more accountable than male voters (Eggers et al. 2018, Esarey & Schwindt-Bayer 2018). The literature provides evidence which suggests that all three explanations may contribute to this overall pattern.

If the magnitude and significance of the presented effects cannot be denied for developing countries, less clear and sometimes contradictory results come from developed Western democracies (Hessami & da Fonseca 2020). For example, in Norway, Geys & Sørensen (2019) employs difference in differences to assess the impact of gender quotas and finds only a small increase in women's political participation, with no shifts in public policies. In Spain, Bagues & Campa (2021) uses RDD to evaluate the impact of gender quotas and finds no impact on the composition or size of public expenditures. In the United States, Broockman (2014) uses RDD on close elections and finds that electing a woman only increases the likelihood of that woman running again, not women's political participation in the same district. Ferreira & Gyourko (2014) also employs RDD on close elections in the United States and finds that gender does not matter for policy decisions on the size and composition of municipal spending. In Italy, Baltrunaite et al. (2019) uses RDD on close elections and a reform of the electoral system and finds no effect of local politicians' gender on the structure of public expenditure. In Germany, Baskaran & Hessami (2018) uses RDD and finds that female mayors increase the share of female council members. These findings may indicate that the causal impact of female participation in politics on the socio-political sphere may change depending on the cultural framework conditions.

In the presented paper, I investigate the causal impact of female political leadership on future female political participation, employment outcomes, and institutional quality, while also showing how such impacts are changing based on the cultural framework conditions in which the elections take place. I argue that this reconciles the mixed results found in the literature

3.3   Institutions and data

3.3.1   Institutional framework

Below the national level, Italy has three levels of government: regional, provincial, and municipal. There are 20 regions, around 100 provinces, and around 8,000 municipalities. Italian municipal governments are composed of an elected mayor, an executive body (*Giunta comunale*) appointed by the mayor, and an elected city council (*Consiglio comunale*). A

majoritarian electoral rule is employed to elect mayors. In cities with fewer than 15,000 inhabitants, the electoral rule is "first past the post" (majoritarian with one round), whereas in more populous cities, the electoral rule is a "second round runoff" (majoritarian with two rounds). The vast majority of municipalities fall into the first category, as only a few hundred cities exceed 15,000 inhabitants.[5]

Municipal elections select both the mayor and city councilors and are held every five years on a Sunday between May and June. All mayoral candidates are associated with a list of candidate councilors. The mayoral candidate who receives the most votes is elected into office and obtains two-thirds (in smaller cities) or three-fifths (in more populous cities) of the city council seats, while the remaining seats are allocated proportionally among the defeated lists. By holding a strong majority in the city council and personally selecting the members of the executive body, Italian mayors enjoy a considerable amount of executive power with a substantial degree of discretion over the budget allocation (Bracco et al. 2019, Fabbrini 2001, Carreri 2019).

3.3.2 Data

The research uses Italian municipal and provincial data from all available municipalities and provinces from 2000 to 2015. The collected data includes electoral results, politicians' characteristics, municipal characteristics, provincial metrics of institutional quality, and socio-economic and geographic controls.

I obtained the electoral data and municipalities' balance sheets from the Italian Ministry of the Interior.[6] Geographical controls such as altitude, longitude, and altitude of the municipalities have been retrieved from the Italian National Agency for New Technologies, Energy, and Sustainable Economic Development[7] (ENEA). Socio-economic variables are collected from the Italian National Institute of Statistics (ISTAT).

To measure institutional quality, I employ the metrics developed by Nifo & Vecchione (2014). Nifo & Vecchione (2014) provides a yearly measure of institutional quality at the provincial level (NUTS3) in Italy. Such institutional quality measure includes different dimensions: corruption, government effectiveness, regulatory quality, rule of law, and political accountability. Table 14 briefly summarises the administrative sources and professional registers used to construct each component of the institutional quality index.

---

[5]Due to the difference in electoral rules, I restrict the sample to cities with less than 15,000 inhabitants. However, the results hold when considering the full sample.

[6]Data on electoral results from the five (out of 20) Italian special autonomy regions are not available. The five special autonomy regions are Sicily, Sardinia, Trentino–Alto Adige, Friuli–Venezia Giulia, and Valle d'Aosta.

[7]Ente per le nuove tecnologie, Energia e Ambiente (ENEA).

Table 14: Components of the institutional quality index

Index	Source	Granularity
Lack of corruption		
Crimes against PA	ISTAT	Region
Golden-Picci Index	Golden & Picci (2005)	Province
Special Commissioners	Interior Ministry	Province
Government effectiveness		
Endowment of social facilities	Tagliacarne	Province
Endowment of econ. facilities	Tagliacarne	Province
Regional health deficit	MEF and MH	Region
Separate waste collection	ISTAT	Province
Urban environment index	Legambiente	Province
Regulatory quality		
Economy openness	Tagliacarne	Province
Local government employees	ISTAT	Region
Business density	Tagliacarne	Province
Business start-ups/mortality	ISTAT	Province
Business environment	Confartigianato	Province
Rule of law		
Crimes against property	ISTAT	Province
Crimes reported	ISTAT	Province
Trial times	DG-STAT, CSM	Province
Magistrate productivity	DG-STAT, CSM	Province
Submerged economy	ISTAT	Region
Tax gap	Revenue Agency	Province
Voice and accountability		
Social cooperatives	ISTAT	Province
Civic engagement	ISTAT	Province
Election participation	Interior Ministry	Province
Books published	ISTAT	Region
Invalsi	Invalsi	Province

Note: Based upon the Appendix material of the Version IV (May 2021) of the Institutional Quality Index made by Nifo & Vecchione (2014)

3.4 Empirical strategy

Identifying the causal effects of politicians' gender on socio-economic outcomes is not an easy task. To begin with, cities that elect women may be different from cities that elect

men. Hence, a simple comparison between localities with a woman mayor and those with a man mayor may generate biased estimates due to endogeneity issues. There might be an underlying reason if a municipality has elected a woman as its mayor; for example, more progressive cities might elect more women mayors than more conservative cities. In more general terms, socio-economic outcomes (such as institutional quality) might be correlated with municipality-specific characteristics, such as attitudes towards politicians' personal backgrounds, all of which could also influence the gender of the local mayor.

To address the endogeneity issue (i.e. the fact that a woman's electoral victory may well be correlated with third context-specific factors related to the outcome variable of interest), I employ two identification strategies: regression discontinuity design (RDD) on close elections and instrumental variable (IV) on close elections. The main idea behind the use of close elections as a source of an exogenous change in the gender of the political leadership is that if a politician wins an election by a small margin of victory, such a victory can be considered as good as random. Hence, the gender of a politician who won a contested election is considered orthogonal to the socio-economic characteristics of the context, allowing for causal inference. Within this framework, the running variable is the margin of victory of the elections, while the threshold is the value of 0% of the running variable.

The formal argument for the validity of this approach is the continuity assumption as pointed out by Hahn et al. (2001) as well as Lee & Lemieux (2010). There is no shift in potential outcomes as we cross the threshold. In other words, the only reason that the outcome shifts at the threshold is the treatment. If reasonably close to the threshold, there are no omitted variables that are correlated with both the treatment status and the outcome. This implies that an election between a female and a male candidate is likely to have been decided by random chance if the election is a very close race. The vote margin between the two candidates effectively determines the treatment (the gender of the mayor) randomly. Then, in close elections, municipalities that happen to have a woman mayor will be comparable to the ones with a man mayor. Even though there is no way to directly test the identifying assumption (Lee 2008), I will perform different approaches to indirectly verify its validity.

The paper will compare municipalities/provinces in which a woman politician barely defeats a man politician to those municipalities/provinces where a woman politician barely loses. This comparison, with respect to some particular outcome variable (e.g. institutional quality), measures the causal effect of politician gender on this outcome. For the approach to be valid, we must consider only observations in which the margin of victory is small and the winner and the best opponent have different genders. This feature makes the empirical strategy very data demanding.

Regression discontinuity design

The regression equation is as follows:

$$Y_{m/p,t+3} = \alpha + \beta female_{m,t} + \gamma X_{m,t} + f(marg) + \varepsilon_{m,t} \tag{7}$$

$Y_{m/p,t+3}$ is the outcome variable of interest in municipality/province m/p in year $t+3$ after the mayoral election[8] α is the constant, $edu_{m,t}$ is a dummy variable that takes the value of one if a woman wins a mayoral election, $X_{m,t}$ is a set of controls, $f(marg)$ is a polynomial of the margin of victory that captures the impact of the vote margin on the outcome variable, and $\varepsilon_{m,t}$ is the error term. β indicates the causal effect of the presence of a female mayor on the outcome of interest.

Instrumental variable design

For the instrumental variable approach, I increase the granularity of the analysis from municipality-year to province-year. In other words, I aggregate electoral information at the provincial level by counting how many women have won contested municipal elections in all Italian provinces every year. Here, I follow a similar IV approach as that of Clots-Figueras (2011) and Rehavi (2007). I define contested elections as any election with a margin of victory smaller than 3%.[9] When looking at the impact of female political representation on institutional quality, the instrumental variable approach is particularly important as it provides causal estimates where the dependent variable and independent variable of interest are at the same level of aggregation, a feature that is lacking in the regression discontinuity estimates.

First stage:

$$\begin{aligned} ShareElectedFemMayors_{p,t} = \alpha + \beta ShareFemMayorsElectedInCloseElections_{p,t} \\ +\gamma X_{p,t} + \varepsilon_{p,t} \end{aligned} \tag{8}$$

Second stage:

$$Y_{p,t+3} = \alpha + \hat{\beta} ShareElectedFemMayors_{p,t} + \gamma X_{p,t} + \varepsilon_{p,t} \tag{9}$$

---

[8]Results are also shown for 1 and 2 years after the election, as well as up to 3 years before the election. The choice of employing 3 years as a baseline is due to the fact that institutions change relatively slowly; hence, an immediate effect is not expected.

[9]Clots-Figueras (2011) defines contested elections as those in which the winner beats the runner-up by less than 3.5% of votes.

$ShareElectedFemMayors_{p,t}$ is the share of elected female mayors in province p in year t, while $ShareFemMayorsElectedInCloseElections_{p,t}$ is the share of female mayors that were elected in a close election in province p in year t. $X_{p,t}$ represents a set of controls, such as the share of close elections in province p in year t, and $\varepsilon_{p,t}$ is the error term. $Y_{p,t+3}$ is the outcome variable of interest in province p in year $t+3$ after the mayoral election. $\hat{\beta}$ indicates the causal effect of the presence of a female mayor on the outcome of interest. As a robustness test, the same specification is also replicated with the instrumented variable expressed in absolute numbers.

3.4.1 Validity of the identification strategy

The identification strategy of the regression discontinuity design will be valid if its continuity assumption holds: no shift in potential outcomes as we cross the threshold. A useful feature of any RD design is that such an assumption can be empirically verified through a series of tests. Furthermore, since the exogeneity assumption of the presented instrumental variable approach relies on the success of the local randomisation if the continuity assumption holds, then the IV exogeneity assumption also holds.

Discontinuity in Density. It checks whether there is a discontinuity in the density of the running variable at the threshold. Figure 11 presents the histogram of the running variable, as well as the rddensity test Cattaneo et al. (2018, 2019) to investigate this issue. I find no significant discontinuity in the density of the margin of victory at the threshold.

Lagged placebo outcome. A different method to check on the success of the local randomization is to estimate the treatment effects on the lagged placebo outcome. Since the treatment effect cannot have a retroactive effect, if the quasi-randomization process has been successful, we should not detect a statistically significant treatment effect on the lagged placebo outcome. Table 22 and figure 12 show no significant discontinuity on any of our lagged outcome variables.

Municipality characteristics balance. To verify the success of the local randomization, I estimate the treatment effects on numerous placebo outcomes that are predetermined and should not be affected. Municipalities where a woman mayor barely won or barely lost the election should not be statistically different in their predetermined characteristics. Table 23 and table 24 indicate no significant discontinuity in any municipality characteristics.

Bandwidth sensitivity test. It verifies that the RDD estimates are not sensitive to the

choice of a particular bandwidth. Figure 13 plots the treatment effects with all 160 possible bandwidths from 0.25% to 40% margin of victory.

Placebo thresholds test. It controls that the treatment effect is present solely at the true threshold of 0% margin of victory. If the treatment effect is found at a fake threshold, it raises questions about the validity of the RD design. Figure 14 shows that there is no significant treatment effect at all 159 possible fake thresholds from -40% to +40% margin of victory.

Confoundedness with Mayor characteristics. A further concern is the confounding effects due to the identity of mayors. Female and male mayors may systematically differ in some personal characteristics, such as education, age or political affiliation. If this was the case, it could be that I am measuring the impact of political affiliation and not of mayor gender. I find that confounding factors can enhance or damper the magnitude of the gender effects (like in the case of education), however, they do not drive the results (see Table 21 and Table 25).

3.5   Results

3.5.1   Effect of mayor gender on women political and public servant participation

Table 15 presents the results on the impact of the mayor's gender on future female political participation. Column 1 displays the results using optimal bandwidth and local polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed in Calonico et al. (2014a) as well as the results using manual bandwidths of 5% and local linear control functions. Columns 2 and 3 replicate the same results by adding year fixed effect and regional fixed effect. The addition of fixed effects is not strictly required by the identification strategy, yet it is a good sign to see that the findings are not sensitive to the inclusion of them.

Overall, estimates suggest that electing by chance a female mayor increases the likelihood that in the next election there will be at least one woman running for the position of mayor by around 40 percentage points. This finding is statistically significant at 1% level using both the optimal bandwidth and manual bandwidth. At the same time, electing by change a female mayor will reduce the probability at least one man will run at the next mayoral election, implying a 7 percentage points increase in the likelihood of having a solo female mayoral electoral competition in the next election.

Table 15: RDD estimates: the impact of mayor's gender on future female political participation

	(1)	(2)	(3)
At least one woman mayoral candidate next election			
RDD result using optimal bandwidth	0.392***	0.390***	0.364***
	(0.0795)	(0.0800)	(0.0764)
	[1,907]	[1,907]	[1,907]
RDD result using 5% bandwidth	0.334***	0.315***	0.306***
	(0.101)	(0.104)	(0.106)
	[330]	[330]	[330]
At least one man mayoral candidate next election			
RDD result using optimal bandwidth	-0.0732**	-0.0725**	-0.0707**
	(0.0323)	(0.0322)	(0.0282)
	[2,046]	[2,046]	[2,046]
RDD result using 5% bandwidth	-0.0350	-0.0273	-0.0237
	(0.0450)	(0.0467)	(0.0472)
	[351]	[351]	[351]
Year FE	No	Yes	Yes
Region FE	No	No	Yes
Running variable	MV woman	MV woman	MV woman

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: woman's margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employ local linear control functions.

Table 16 presents the results on the impact of the mayor's gender on future female participation as public servants. Electing by chance a female mayor is found to increase the share of female workers in public administration as soon as two years after their election. Overall, electing a female mayor makes the share of female employees in public administration rise by around 5 percentage points. Such a result is statistically significant at 5% level using the optimal bandwidth and at 1% level with the manual bandwidth.

Table 16: RDD estimates: the impact of mayor's gender on future female work representation in the Public Administration

	Average % of female workers (t+2)			Average % of female workers (t+3)		
	(1)	(2)	(3)	(4)	(5)	(6)
RDD result using optimal bandwidth	0.0523**	0.0554**	0.0388*	0.0468*	0.0504*	0.0343
	(0.0256)	(0.0258)	(0.0226)	(0.0259)	(0.0260)	(0.0222)
	[2,247]	[2,247]	[2,247]	[2,160]	[2,160]	[2,160]
RDD result using 5% bandwidth	0.112***	0.113***	0.107***	0.110***	0.108***	0.0958***
	(0.0362)	(0.0365)	(0.0315)	(0.0374)	(0.0378)	(0.0321)
	[390]	[390]	[390]	[368]	[368]	[368]
Year FE	No	Yes	Yes	No	Yes	Yes
Region FE	No	No	Yes	No	No	Yes
Running variable	MV woman	MV woman	MV woman	MV woman	MV woman	MV woman

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: woman's margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employ local linear control functions. The outcome variable is a two/three years average after the elections.

The presented results can be visually inspected in figure 8. At the 0% threshold of the female mayors' margin of victory, it can be observed how the likelihood of women's participation in the next election sharply jumps from around 50% to around 80%. At the same time, crossing the 0% threshold, the percentage of female workers in public administration jump from around 45% to around 55%.

Figure 8: RD plot for woman political and work participation
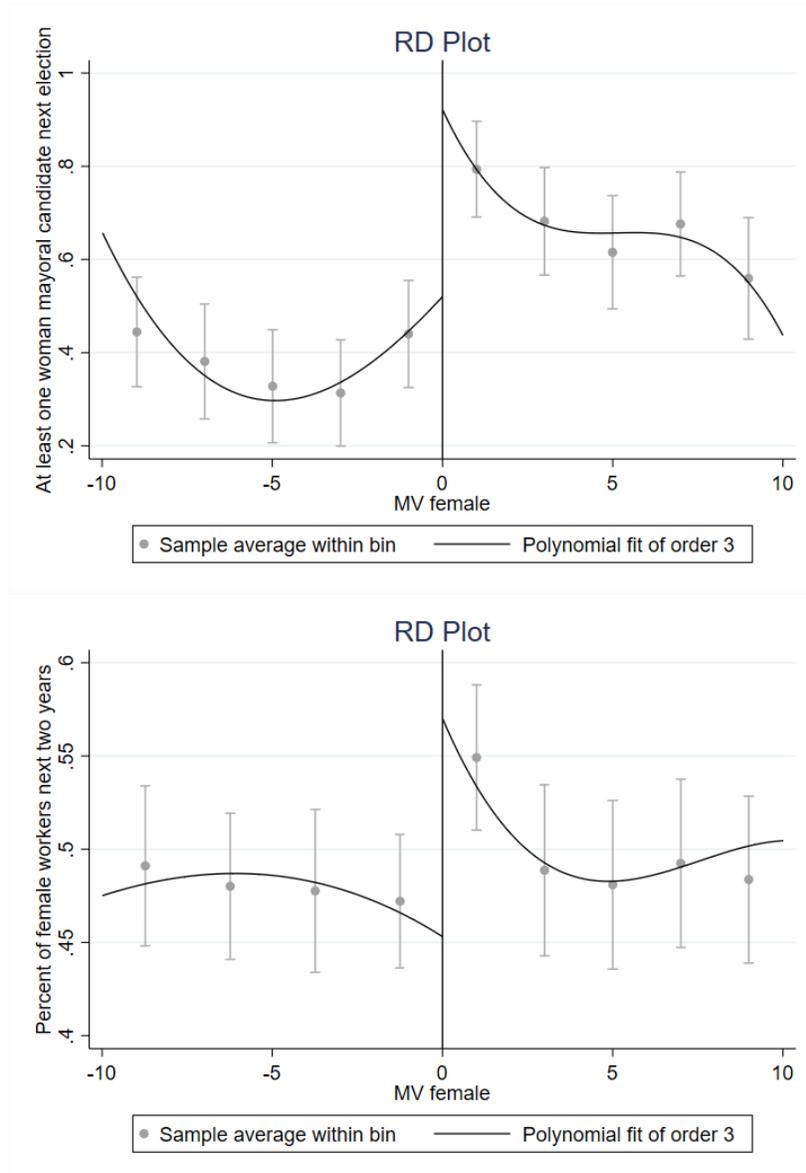
Note: IMSE-optimal evenly-spaced method using spacings estimators to select the number of bins, 95% confidence interval.

This effect is found in much of the literature (Martin & Borrelli 2016, Atkeson 2003, Bauer & Okpotor 2013, Campbell & Wolbrecht 2006, Dassonneville & McAllister 2018, Fridkin & Kenney 2014, Gilardi 2015, Ladam et al. 2018, Liu & Banaszak 2017, Mariani et al. 2015, McAllister 2019, Bhalotra et al. 2018). Yet, there are studies that do not find that having a prominent female political role model impacts levels of political interest and or engagement among women (Broockman 2014, Carreras 2017, Dolan 2006, Lawless 2004).

3.5.2 Effect of mayor gender on institutional quality

Table 17 presents the results on the impact of the mayor's gender on future institutional quality. Overall, I detect an increase in the quality of institutions (measured at the provincial level) from the second year after the election of female mayors. Electing by chance a female mayor improves the institutional quality of the province in which the mayor is elected by 5 percentage points three years after the election.

Table 17: RDD estimates: the impact of mayor's gender on future Institutional Quality

	Institutional Quality Index (t+2)			Institutional Quality Index (t+3)		
	(1)	(2)	(3)	(4)	(5)	(6)
RDD result using optimal bandwidth	0.0440**	0.0439**	0.0178	0.0554***	0.0423**	0.0253*
	(0.0201)	(0.0180)	(0.0124)	(0.0214)	(0.0180)	(0.0132)
	[3,255]	[3,255]	[3,255]	[3,124]	[3,124]	[3,124]
RDD result using 5% bandwidth	0.0374	0.0341	0.0305	0.0646**	0.0581**	0.0529***
	(0.0289)	(0.0282)	(0.019)	(0.0301)	(0.0287)	(0.0192)
	[558]	[558]	[558]	[539]	[539]	[539]
Year FE	No	Yes	Yes	No	Yes	Yes
North-South dummy	No	No	Yes	No	No	Yes
N. elections in province	No	No	Yes	No	No	Yes
N. female mayor in province	No	No	Yes	No	No	Yes
Running variable	MV woman	MV woman	MV woman	MV woman	MV woman	MV woman

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: woman's margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employ local linear control functions. The outcome variable is a two/three years average after the elections at the provincial (NUTS3) level.

Table 18 shows the impact of electing a female mayor on the different sub-component of the institutional quality index. Overall, electing a female mayor reduces corruption, increases government effectiveness, rule of law and political accountability.

Table 18: RDD estimates: the impact of mayor's gender on future Institutional features (t+3)

	Lack of corruption	Gov effectiveness	Regulatory quality	Rule of law	Voice and accountability
	(1)	(2)	(3)	(4)	(5)
RDD result using optimal bandwidth	0.0342**	0.0324*	0.0291	0.0437*	0.0423**
	(0.0165)	(0.0167)	(0.0183)	(0.0228)	(0.0188)
	[3,124]	[3,124]	[3,124]	[3,124]	[3,124]
RDD result using 5% bandwidth	0.0221	0.0435*	0.0273	0.0500	0.0433*
	(0.0211)	(0.0250)	(0.0287)	(0.0322)	(0.0244)
	[539]	[539]	[539]	[539]	[539]
Running variable	MV woman	MV woman	MV woman	MV woman	MV woman

Note: Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: woman's margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employ local linear control functions. The outcome variable is a two/three years average after the elections at the provincial (NUTS3) level.

The results can be visually inspected in figure 9. At the 0% threshold of the female mayors' margin of victory, it can be observed how the institutional quality index sharply jumps from around 0.65 to around 0.72.

Figure 9: RD plot for Institutional Quality
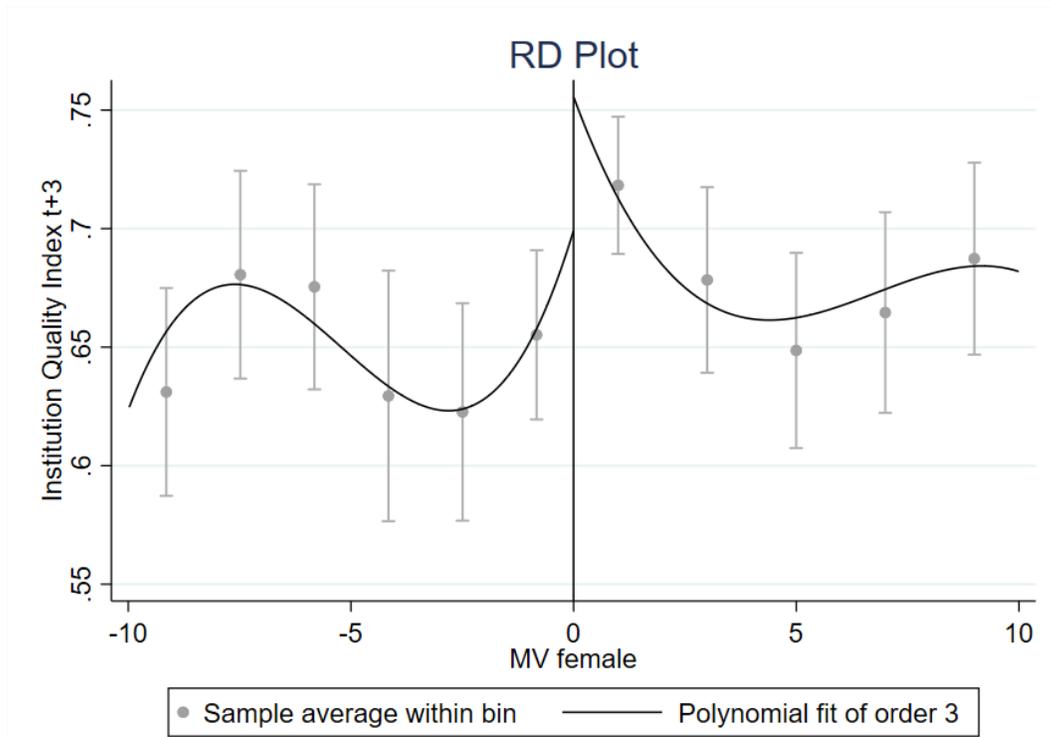
Note: IMSE-optimal evenly-spaced method using spacings estimators to select the number of bins, 95% confidence interval.

3.5.3 IV approach

Table 19 shows the results of the second stage of the IV estimates. Consistently to the RDD estimates, electing by chances more female mayors within a province increases the institutional quality of that province. The results are mostly robust also to the inclusion of year and province fixed effects. Table 20 depicts the first stage of the IV strategy for all the specifications presented. The first stage estimates result to be all highly statistically significant. The instrument passes both under-identification and weak identification tests.

Table 19: IV second stage estimates: the impact of mayor's gender on future Institutional quality (t+3)

	(1)	(2)	(3)	(4)
Institutional quality				
Instrumented share of women mayors	0.259*	0.0468**		
Instrumental share of women by chance	(0.145)	(0.0205)		
Instrumented n. women mayors			0.0405**	0.0138*
Instrumental share of women by chance			(0.0204)	(0.00770)
Share of close elections	-0.0558	-0.00830	0.0632	-0.00348
	(0.0630)	(0.0142)	(0.0683)	(0.0140)
Observations	1,145	1,145	1,145	1,145
Control for Share of CE in province	Yes	Yes	Yes	Yes
Year FE	No	Yes	No	Yes
Province	No	Yes	No	Yes

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. In columns 1 and 2 the instrumental variable is the Share of female mayors by chance (3%) in the province (out of n. mayoral elections). In column 1 the instrumented variable is the share of elected female mayors in a province (out of n. mayoral elections), while in column 2 the instrumented variable is the n. female mayors by chance (3%) in the province. In all regressions, I control for the share of close elections of a province in a year.

Table 20: IV first stage estimates: the impact of mayor's gender on future Institutional quality (t+3)

	(1)	(2)
Share elected female mayors		
Share fem mayors elected in close elections	0.907***	0.832***
	(0.0800)	(0.0811)
Share of close elections	0.0491	0.0787
	(0.0932)	(0.0904)
Observations	1,145	1,145
Control for Share of CE in province	Yes	Yes
Year FE	No	Yes
Province FE	No	Yes
Cragg-Donald Wald F statistic	108.156	93.866
Kleibergen-Paap rk LM statistic	7.787	8.381

Note: Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors are in parentheses. The share of women mayors elected in the province is instrumented using the share of the total of women mayors that got elected in that province in a close election. For the IV specification, a close election is considered and eventually counted if the margin of victory is below 3%. In all regressions, I control for the share of close elections of a province in a year.

3.5.4 Potential channels and mechanisms

Different mechanisms could explain the presented findings. Men and women may differ in policy objectives and world views, as well as in personal characteristics such as level of education and life experiences. Such differences will be greater in areas with more established gender norms, and smaller in more progressive environments. In male-dominated hostile environments, women will find it harder to succeed, having to put in more work and effort, to endure more hardships and injustices, while developing more character and resolution Cowper-Coles (2020). In such environments, once reached leadership positions, women may be more likely to challenge the status quo and change the institutions that did them wrong. Understandably, in more progressive communities, such differences between men and women in motivation and characteristics may be smaller, as well as the institutional changes that

could be done to reduce them.

To verify the aforementioned discussion, I test whether the causal marginal impact of electing a female mayor by chance on institutional quality is greater in areas with more established gender norms compared to more progressive ones. In less progressive provinces, a woman who wins an election by chance will have both more motivation (having endured greater hardship during her life) and ability (having required more education and skills to be selected as a mayor candidate) than a woman in more progressive provinces in relation to their male counterparts. At the same time, in less progressive provinces, there will be much more to do for a well-motivated and skilled woman leader than in more progressive provinces where much has already been accomplished

I proxy the degree of local progressiveness in three ways: latitude of the municipalities, number of female mayors in the province and share of female counsellors in the city council. According to survey[10] performed by the Italian National Institute of Statistics in 2018 (in which respondents were asked for their agreement with stereotyped statements about the role of women in the labour and financial sphere, family decisions and household management) gender roles stereotypes are more frequent in southern Italy. In line with such findings, southern areas of Italy present higher levels of female unemployment than the north. Hence, I interpret lower latitude, lower number of female mayors in a province and lower share of female counsellors as a different measurement of the lower level of progressiveness in society. If the presented mechanism is correct, I should observe larger and more statistically significant coefficients when running the RDD on sub-samples including cities from lower latitudes and with fewer female mayors or counsellors.

In figure 10, I restrict the sample to a moving 6 units latitude window and I run 19 RDD estimations by shifting the 6-length units window in increments of 0.25 units each time. I perform the same exercise with the other two proxies of gender norms and plot the coefficient of the causal impact of electing a woman on institutional quality based on the different gender norms windows sub-samples. In line with the hypothesized mechanism, the positive impact of female leadership on institutional quality is larger in magnitude and statistical significance in those areas that present more established gender norms. In other words, the causal coefficient decreases with the increase in community progressiveness. Such results also reassure us of the absence of endogeneity from the estimates, indeed, if the treatment would be endogenous, we would expect to find larger (correlation, not causal) coefficients between electing women and being in progressive areas, as more progressive areas are more likely to elect women.

---

[10]See https://www.istat.it/en/archivio/236678

Figure 10: Community progressiveness and marginal impact of female leadership
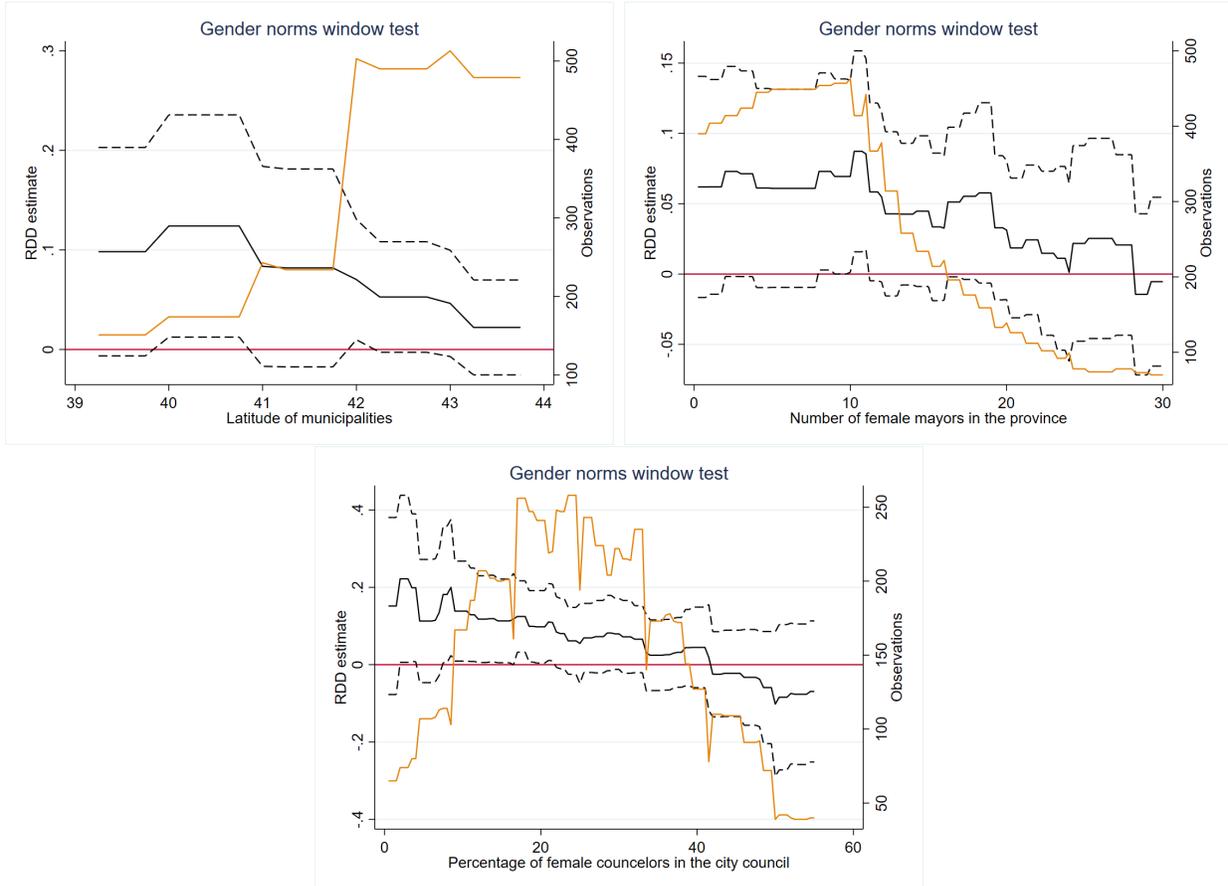
Note: The three plots represent 19, 120, and 110 RDD estimations using local linear control function with a threshold of 5% and shifting windows of 6 latitudes points, 20 number of female mayors in the province, and 16 percentage points of female counselors in the city council, shifts in increments of 0.25 latitude, 0.25 number of female mayors in the province, and 0.5 percentage points of female counselors in the city council. 95% confidence interval. Outcome variable: Institutional quality (t+3). The yellow line and the right y-axis indicate the number of observations used in each RDD estimation.

In table 21, I try to verify if women's personal characteristics (such as education) have also a role in explaining the finding. If the aforementioned hypothesis is correct, I should observe that the more educated women are the ones that if elected by luck, are more likely to challenge the status quo and effectively change institutions for the better. I measure education as holding a tertiary degree. I find that both less and more educated women that win by chance a mayoral election boost institutional quality in the upcoming years, however, such effect is around 38% larger when the woman winning the election has tertiary education.

The fact that educated women who win by chance against a man are more effective in improving institutional quality implies that education is not only a key factor for women's

empowerment and welfare (Samarakoon & Parinduri 2015, Breierova & Duflo 2004, Mocan & Cannonier 2012), but also an enhancer of their ability to actively transform the society in which they live. This is in line with the theoretical model developed by (Mitra 2020), in which politicians' education is seen as a technology that shapes politicians' self-interested policy platforms to include more challenging policies due to the increased likelihood (for such educated politicians) of succeeding with those policies. Supporting this argument, Mitra (2020) finds (using regression discontinuity design) that educated politicians improve fiscal planning by shifting resources from a simpler form of spending (current expenditure) to more complex ones (public investment). Similarly, female-educated politicians may be more likely to engage and succeed with ambitious and transformative institutional policies, thanks to their higher degree of competence and knowledge. For women, as for men, education reveals itself as a tool capable of transforming good intentions into the ability to take effective actions.

Table 21: Effect of a female winning a close election interacted with its level of education

	Institutional Quality Index (t+3)	Institutional Quality Index (t+3)
	(1)	(2)
Woman win close election (0) × winner has a tertiary degree (1)	-0.0140	-0.0152
	(0.0162)	(0.0154)
Woman win close election (1) × winner has a tertiary degree (0)	0.0414*	0.0367*
	(0.0231)	(0.0213)
Woman win close election (1) × winner has a tertiary degree (1)	0.0533**	0.0517**
	(0.0231)	(0.0210)
Observations	512	512
Year FE	No	Yes
North-South dummy	Yes	Yes
N. elections in province	Yes	Yes
N. female mayor in province	Yes	Yes
Running variable	MV woman	MV woman

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Running variable: woman's margin of victory. Threshold=0%. RDD results with a bandwidth of 5% employing local linear control functions. (0) indicates that a woman has not won a close election or/and that the winner has not a tertiary degree. (1) indicates that a woman won a close election or/and that the winner has a tertiary degree.

3.5.5 Validity of the identification strategy

To test the validity of the regression discontinuity design I conduct several robustness tests. A great advantage of the IV approach presented in the paper is that since it relies on the randomness of the close elections, if the RDD validity tests are successful, also the IV exogeneity assumption can be seen as indirectly verified.

Discontinuity in Density.

The RDD analysis based on close elections crucially relies on randomized variations just

around the threshold. Such a requisite does not hold if there is sorting around the threshold. As an example, if a political candidate can manipulate a contested electoral outcome, causing sorting, the treatment would not only be determined by the running variable. Therefore, units on each side of the threshold would not be comparable as in the randomization case, bringing us back to an endogenous world.

Since sorting is connected with the idea of having more, or less, observations on either side of the threshold, I check this by visually scrutinising the histogram of the running variable and checking that there are no spikes at the right or left-hand side of the threshold. I also more formally examine the sorting issue by performing the density test of the running variable as proposed by Cattaneo et al. (2018, 2019). The test uses a non-parametric density estimator based on local polynomial techniques and, differently from the McCrary test (McCrary 2008), does not require pre-binning or any other transformation of the data. Figure 11 shows no evidence of sorting.

Figure 11: Histogram and Rddensity test around the threshold
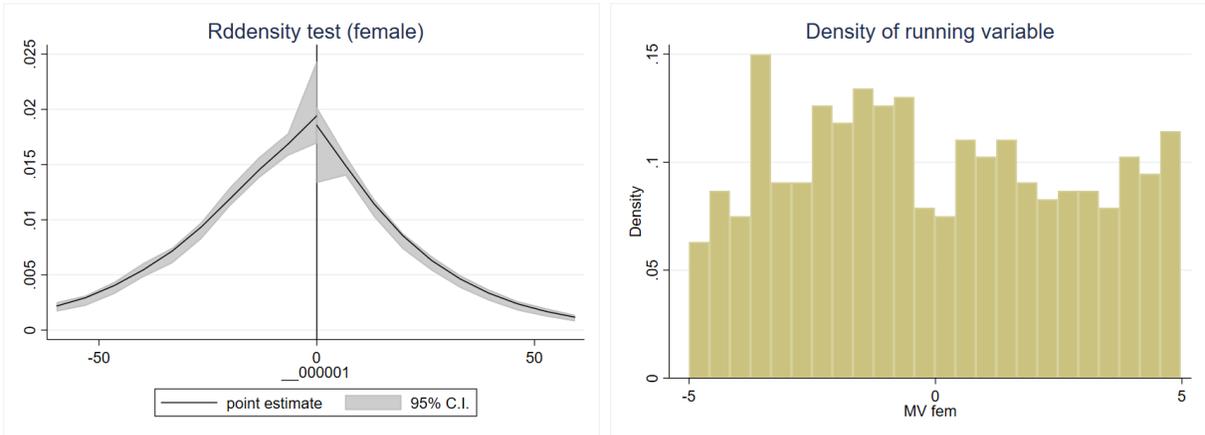
Note: The histogram presents the frequencies of observations for the RDD analysis. The data are ordered by the assignment variable "woman mayor margin of victory". The frequencies within a margin of victory of ±10 percentage points are shown. The rddensity test plots the density around the cutoff using the local polynomial density estimators.

Lagged placebo outcome.

Since the improvement in institutional quality and female political representation is caused by the victory of women mayors, I should not observe any statistically significant difference in any of these outcome variables between the cities that will have a female mayor in the upcoming year and those that will not. In other words, the increase in institutional quality(or future female participation) can only be caused by the treatment and not pre-existing.

Table 22 highlights the estimated effects on the lagged outcome variables: female political participation in the last election, and women working in the public administration. Again, it is reassuring not to detect any statistically significant result. Figure 12 plots the RDD estimates on the lagged outcome variables of institutional quality. The estimates are statistically insignificant in all the pre-electoral periods, starting to became significant two-three years after the close electoral victory of a woman mayor.

Table 22: Lagged placebo outcome RDD estimates

	At least one woman candidate (last election)	Average % of female workers (t-2)
	(1)	(2)
RDD result	-0.0252	0.0168
	(0.0601)	(0.0246)
Running variable	MV woman	MV woman
Observations	2,371	2,474

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Running variable: woman's margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). Lagged placebo outcomes either of the former election or of two years average previous the election.

Figure 12: Evolution of RDD estimates in time
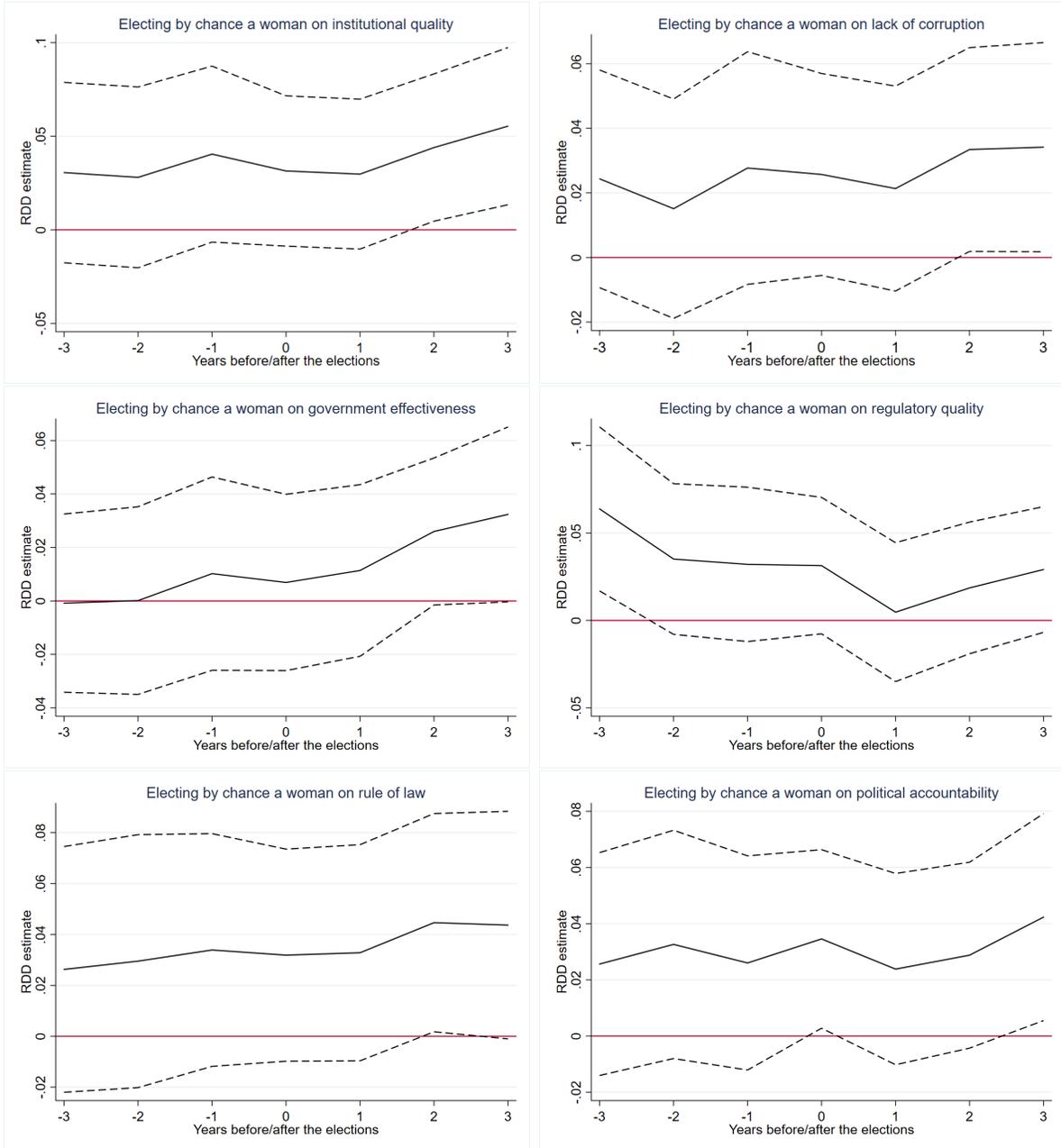
Note: Each graph represents 6 RDD results using optimal bandwidth, local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed in Calonico et al(2014), with shifting time the lag of the outcome variable. The coefficients represent the causal impact of electing by chance a woman mayor on the institutional quality metrics at different years before/after the election.

Municipality characteristics balance.

To confirm that the randomization holds, I also perform a municipality characteristics

balance test. I run the same RDD models of my main specification on various demographic and geographical municipality-level characteristics that are predetermined and should not be affected by the treatment. Table 23 shows no significant impact for any of the municipal characteristics considered.[11] This ensures us the absence of any omitted variables, which could have survived in case of a failure of the quasi-randomization process. In table 24 I repeat the same exercise using regional dummies as placebo outcome variables. The lack of statistical significance in table 24 reassures the absence of regional biases.

Table 23: Municipality characteristics balance test

Municipality placebo outcome	(1) RDD result using optimal bandwidth	(2) RDD result using 5% bandwidth
Latitude	0.413*	0.415
	(0.225)	(0.317)
	[3,120]	[540]
Longitude	-0.425	-0.172
	(0.297)	(0.461)
	[3,120]	[540]
Census population	483.9	974.5*
	(414.9)	(565.1)
	[3,487]	[606]
Urban surface(ha)	-83.16	-166.4
	(69.26)	(105.2)
	[2,516]	[439]
Length internal roads(km)	1.553	-5.474
	(3.019)	(4.132)
	[2,515]	[435]
Length external roads(km)	27.42	43.47
	(23.59)	(32.49)
	[2,422]	[422]
Running variable	MV woman	MV woman

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: woman's margin of victory. Threshold=0%. Placebo variables are either steady or taken in the year of the elections.

---

[11]Latitude and population are weekly significant at 10%, yet the result is not robust to different bandwidths, hence not raising concerns

Table 24: Municipality regional balance test

Municipality placebo outcome	(1) RDD result using optimal bandwidth	(2) RDD result using 5% bandwidth
Region Piemonte	0.0237	-0.0274
	(0.0341)	(0.0604)
Region Liguria	0.00941	-0.00217
	(0.0163)	(0.0257)
Region Lombardia	0.00624	0.00636
	(0.0163)	(0.0764)
Region Veneto	0.0216	0.0525
	(0.0348)	(0.0543)
Region Emilia-Romagna	0.00559	0.0178
	(0.0213)	(0.0335)
Region Marche	-0.00122	-0.0195
	(0.0160)	(0.0192)
Region Toscana	0.0122	-0.0127
	(0.0196)	(0.0226)
Region Umbria	0.0141	0.0200
	(0.0122)	(0.0191)
Region Lazio	-0.0196	-0.0279
	(0.0217)	(0.0373)
Region Campania	-0.0371	-0.0465
	(0.0231)	(0.0328)
Region Abruzzo	0.00470	0.0572
	(0.0219)	(0.0377)
Region Molise	-0.00765	-0.00277
	(0.0180)	(0.0250)
Region Puglia	0.00269	0.00941
	(0.0193)	(0.0307)
Region Basilicata	0.00831	-0.00724
	(0.0124)	(0.0148)
Region Calabria	-0.0283	-0.0171
	(0.0177)	(0.0245)
Observations	3,518	609
Running variable	MV woman	MV woman

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Running variable: woman's margin of victory. Threshold=0%.

Placebo bandwidth test.

To check that the main results of the paper are not sensitive to the choice of the bandwidth, I perform a bandwidth sensitivity test. To do the bandwidth sensitivity test I conduct a set of RDD estimates increasing the bandwidth length in increments of 0.25 pp, starting from 0.25% to 40%. Figure 13 shows the results employing a Local linear control function on future female political participation, future female public administration participation and institutional quality. As expected, RDD estimates start insignificant with the smallest of the bandwidths, and as soon as statistical power increases (due to the rise in the identifying observations) results become statistically significant and steady.

Figure 13: Bandwidth sensitivity test

Placebo thresholds test.

To rule out the possibility that the main result arises from random chance rather than a real causal relationship, I perform a threshold placebo test. I do this by examining if there are any jumps in the relationship between the running variable and the outcome variables. To do the threshold placebo test I conduct a set of RDD estimates by shifting the threshold of 0.25 pp, starting from -40% to +40% (being 0% the correct threshold). In all the 160 estimates a small-moving 3% bandwidth is employed (see figure 14). Overall, almost all the placebo thresholds show statistically insignificant estimates. As expected, at threshold=0% the estimated impact of mayor gender on all outcome variables is shown as positive and statistically significant at 5%.

Figure 14: Placebo threshold test
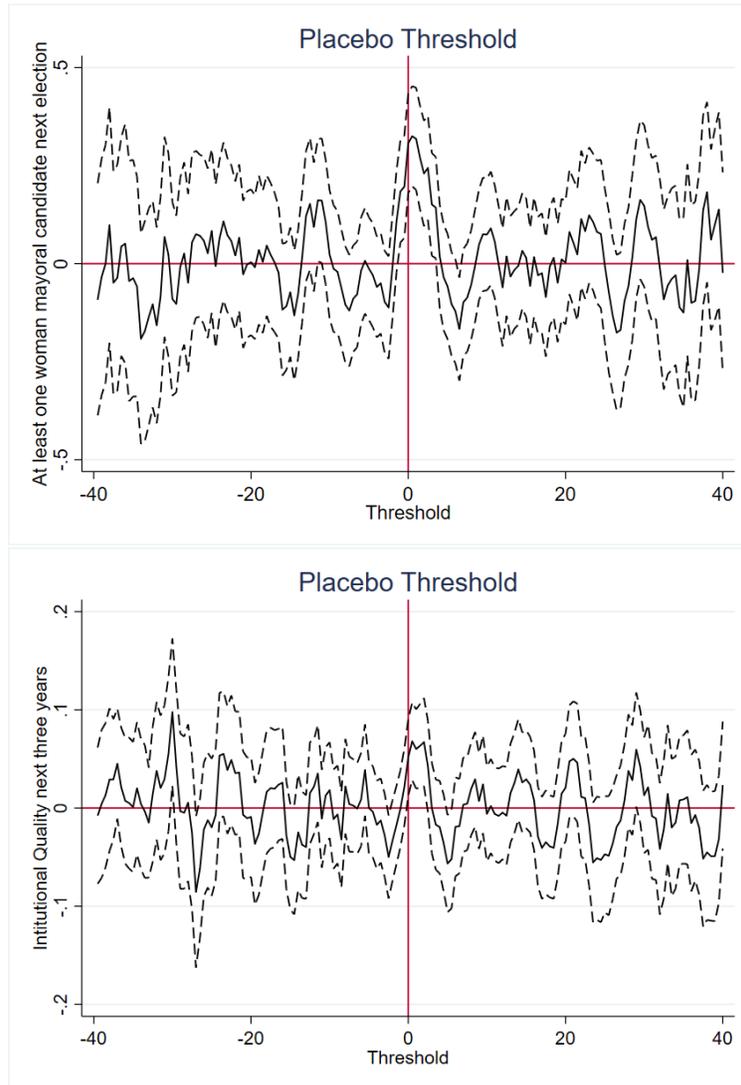
Note: 160 RDD estimations using a moving threshold of 3%, shifts in increments of 0.25 pp. 95% confidence interval.

Confoundedness with Mayor characteristics.

It has to be noted that the treatment is "electing a female mayor" with all the attached individual characteristics that usually come with being a woman. The treatment is a bundle of elements composed of the gender of the politician and other elements that come with it. As an example, women may be on average more educated or left-wing. To verify that such co-occurring characteristics are not driving the results I add age, level of education, incumbency and party affiliation as control variables, and see if there is any change in the estimated, etc. (table 25). Such a check is important to avoid misinterpreting the results of politician characteristic regression discontinuity (PCRD) designs (Marshall 2022). All

estimations show minor changes in the estimated effect with respect to the baseline result, confirming that neither age, education, incumbency, or party affiliation are biasing the results due to potential co-occurrence with the mayors gender.

Table 25: RDD estimates controlling for co-occurrence mayors' characteristics

| | Institutional Quality Index (t+3) | |
	(1)	(2)
RDD result	0.0712***	0.0531***
(mayor education and age as control)	(0.0217)	(0.0204)
	[2,998]	[512]
RDD result	0.0658***	0.0401*
(incumbent candidate as control)	(0.0248)	(0.0224)
	[2,362]	[409]
RDD result	0.0521**	0.0496**
(mayor party FEs as control)	(0.0216)	(0.0201)
	[2,846]	[491]
Bandwidth	Optimal	5%
Year FE	No	Yes
North-South dummy	No	Yes
N. elections in province	No	Yes
N. female mayor in province	No	Yes
Running variable	MV woman	MV woman

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Number of observations in square brackets. Running variable: woman's margin of victory. Threshold=0%. RDD result using optimal bandwidth employs local linear polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures as developed by Calonico et al(2014). RDD results with a bandwidth of 5% employs local linear control functions.

3.6   Conclusion

The purpose of this paper is to evaluate the impact of female political representation on politics, public administration, and institutional quality. The theoretical hypothesis is that if communities present gender norms, politicians' gender matters for policy preferences and political decisions. As a consequence, women politicians will increase future women's participation in the public sphere, both as politicians and public servants. Women's different perspective to the dominant gender-biased status quo may also bring political creative destruction capable of renovating and improving institutional quality.

I use regression discontinuity design, as well as instrumental variable approach, based on close elections to identify the causal effect of politicians' gender on future women empowerment and institutional quality. I find evidence that women mayors improve the institutional quality of the provinces where they are elected, and stimulate more women involvement in politics and public administration in the municipalities where they are elected. Such effects are stronger in areas where gender norms are more established.

The results are robust to controlling for election-specific factors and using alternative samples, functional forms, and bandwidths. Furthermore, all validity tests suggested by the literature are performed, reassuring us of the validity of the identifying assumption of the regression discontinuity design.

The paper presents the first empirical evidence on the impact of women's political participation on different institutional quality variables (such as corruption, government effectiveness, regulatory quality, rule of law, and political transparency) in a Western democracy. The findings suggest that gender can affect institutional quality, with women politicians improving it. The argument is that such impact is driven by the fact that motivated and resolute women are more likely to choose to challenge the status quo, transforming the institutions that have ill-treated them. Consistent with this argument, the paper finds that the impact of female leadership in institutions is inversely related to the level of establishment of gender norms in the community. The results contribute to the literature on gender and politics and have important implications for policymakers interested in improving institutional quality and promoting gender equality.

To conclude, given the widespread recognition on the role of institutions in defining communities well-being and prosperity, the promotion of women political participations becomes a key policy option, not only to enhance gender equality, but also to nurture socio-economic progress. While the promotion of women education emerge as not only a tool for women emancipation, but as a mean to better society.

Chapter 4

Do speeches at the UN General Assembly affect international aid allocation?

4.1 Introduction

From the beginning of the Cold War, the US and Russia have vied for their cultural and political influence, often using development aid as a tool to incentivise and reward political alignment of recipients towards their opposing ideological platforms. This phenomenon has been extensively studied using historical voting records at the United Nations.

However, there is arguably a richer source of information that has not been systematically exploited yet: the speeches and statements that hundreds of heads of state and high-level government officials make at the United Nations General Assembly every year. Such speeches, covering thousands of pages of scripts, are crucial for the functioning and internal debate of the United Nations, as well as for international politics in general.

The aim of this paper is to uncover if UN speeches matter for the allocation of international aid. I employ machine learning techniques and text as data approach to measure countries' preferences from text and establish whether countries whose speeches are lining toward the US or Russia receive more development support from either world superpower. The measure of countries' preferences is derived using the supervised machine learning algorithm called wordscore (Baturo et al. 2017). To scale the different documents, the algorithm looks at the feature-frequency matrix of the different speeches and uses the speeches of the US and Russia as reference points.

Countries' preferences exhibit both cross-sectional and time variation, making it the perfect case for panel analysis. The research uses panel data from 194 countries over the period 1975–2018 to empirically analyse the relationship between speech alignment in the UN General Assembly and US/Russia aids. Overall, I find that countries' statements at the United Nations General Debate do matter for aid allocation. The US allocates more aid to countries whose speeches leaned more toward the US preference and less to those countries that, instead, pleased more Russia. The same phenomenon applies to Russia: Moscow grants more aid to countries whose speeches are closer to its own.

The research brings novelty in three ways. Firstly, it depicts the evolution of UN members' preferences from 1975 to 2018 using an index of government preference that measures countries' preferences in terms of the relative distance between US and Russian preferences using text as data. Secondly, it shows that wordscore estimates of countries' preferences

can be successfully employed within a panel data regression framework. Thirdly, it brings original empirical evidence on the relevance of speech alignment and government preferences for aid allocation.

4.2   Literature review

Understanding the determinants of international aid is an old question of both international relations and comparative politics. Apart from altruistic reasons, the literature presents several selfish motivations such as means to buy political influence, gain economic return, promote strategic interests and reward political alignment of aid recipients (Schraeder et al. 1998, De Mesquita & Smith 2009, Hoeffler & Outram 2011).

To measure the relationship between international aid and recipients' political alignment with the donor country, scholars have overwhelmingly relied on UN General Assembly voting records. Bernstein & Alpert (1971), Rai (1980), Wittkopf (1973), Wang (1999), Alesina & Dollar (2000), Alesina & Weder (2002), Gates et al. (2004) look at the relationship between bilateral aid and UN voting behaviour and find a positive relationship between aid and voting similarity.

On the other hand, other studies find that aid is ineffective in influencing the voting behaviour of recipients (Kato 1969, Kegley Jr & Hook 1991, Morey & Lai 2003).

Lundborg (1998) focuses on the aid provision of the United States and the Soviet Union from 1948–1979. His analysis shows that recipients voted in line with the USA or USSR to stimulate aid provision from them. Alesina & Dollar (2000) argued that the US rewards recipients' voting compliance by granting more aid while Alesina & Weder (2002) and Gates et al. (2004) find that the same aid reward behaviour applies to other G7 countries.

Despite its wide use, UN votes' records have been shown to be an imperfect solution to the task of constructing measures of government preferences and political alignment (Voeten 2013). Measures based on votes in the UN General Assembly often fail to separate changes in the UN's agenda from shifts in state preferences. At the same time, vote choice is frequently constrained by existing alliances, strategic voting, the topic of the issue to be voted and external pressures.

To overcome the aforementioned drawbacks, Baturo et al. (2017) proposed an alternative measure of government preferences that do not rely on UN General Assembly votes but on the speeches that the countries' representatives give at the UN General Debate. At the UN General Debate countries' representatives are free to comment on any topics that they feel of relevance without any legally binding consequences. Using the application of text analytic methods to these statements Baturo et al. (2017) introduces a new measure of government

preference suited to capture latent preferences.

Words are vehicles of meaning and, with far more shades of significance than votes, are of crucial relevance in international relations. This paper contributes to the existing literature on international aid by using quantitative text analysis techniques to uncover the relationship between speeches in the UN General Debate and aid provision.

4.3 Data and Method

4.3.1 Text as data

Every year at the United Nations (UN), member states deliver statements during the General Debate (GD) discussing major issues in world politics. High-level country representatives and head of states fly from all over the world to gather in New York and express their political views at the United Nation General Assembly. I collect the text data on countries' statements at the UN General Debate from the United Nations General Debate Corpus (Jankin Mikhaylov et al. 2017). In total, I gather 7,533 different speeches of 198 countries from 1975 to 2018. On average, speeches are 4,154 words long and contain 1,130 unique words.

Following the standard text analysis pre-processing approach, all speeches undergo tokenization, removal of English stop words, stemming and trimming of the document-feature matrix. Figure 15 shows the word cloud for the speeches of the US and Russia in 1980 and 2002. In 1980, the US statement focuses more on concepts such as nation, peace and humanity, while Russia relies more on terms such as state, nuclear, soviet and military. In 2002, both superpowers' statements reflect the current international urgency caused by the 9/11 attack. However, it can be noticed how the choice of words differs radically between the two countries. If the US focuses on Iraq, weapons and discusses about authoritarian regimes, Russia takes a more neutral stance with words such as "international" and "terror".

Figure 15: Word clouds of USA and Russian speeches at the UNGD in 1980 and 2002

(a) USA (1980)

(b) Russia (1980)

(c) USA (2002)

(d) Russia (2002)

Note: Word cloud plots do not include English stop words. Furthermore, all words are reduced to their primitive form. As an example: humanity, humanitarian, inhumane, humanise, etc. are reduced to human.

To transform such text data into numerical information I employ a scaling supervised machine learning algorithm called wordscore. Wordscore (Laver et al. 2003) is the most popular text scaling method to estimate policy positions on dimensions of interest in political science. It differs from other text scaling methods such as Wordfish (Slapin & Proksch 2008) as it requires some reference texts, already labelled, to use as training documents. In other words, the texts of reference, for which the position on the dimension of interest is known,

are used by the algorithm as a guide to scale the other text documents[12]. One drawback of such an approach is that, as for any supervised method, the researcher has to feed the algorithm with some a priori information. On the other hand, this allows to clearly define the dimension measured, while with unsupervised scaling methods, such as wordfish, it is complex to understand at which angle of the multidimensional space the scaling is being performed.

To measure how the US and Russia change their aid allocation depending on the UN members' statements, I set the US and Russian speeches at the General Debate as opposing reference documents. For each year, I rank the relative distance of each country statement to the US and Russian statements using the wordscore algorithm. I interpret these wordscore measures as the relative distance between the US and Russia on the policy and ideological dimension for the current year of each country's preferences. The US is given a reference score (-1) and Russia (+1). Results are rescaled as proposed by Benoit & Laver (2008). In this way, all countries' wordscore is comprised between -1 and +1. The lower the wordscore is, the closer that country's preferences are toward the USA's own preferences and vice versa.

Despite the government preferences and policy positions of the US and Russia can vary through time, getting closer and far apart from them, depending on the year, the analysis allows for time comparisons. Indeed, our analysis is centred on the countries' preferences closeness to the two superpowers, whatever these preferences may be. For each year, the algorithm re-scales the UN members' preferences based on the US and Russian statements of that session.

Figures 16 maps wordscore estimates for the 1980 UN General Debate. As expected, the world is divided into two strong areas of influence with Eastern Europe much under the influence of Russia and Anglo-Saxon countries much closer to the US. The rest of the world, instead, resembles an open ground for clashes of political influences. To see how much the wordscore evolves over time I plot the time trends of UN countries' wordscore. Figure 17 reveals that government preferences meaningfully vary both through time and cross-country.

---

[12]For more details on how the Wordscore algorithm works see Appendix A

Figure 16: Wordscores map 1980

RUS vs USA: Wordscores 1980
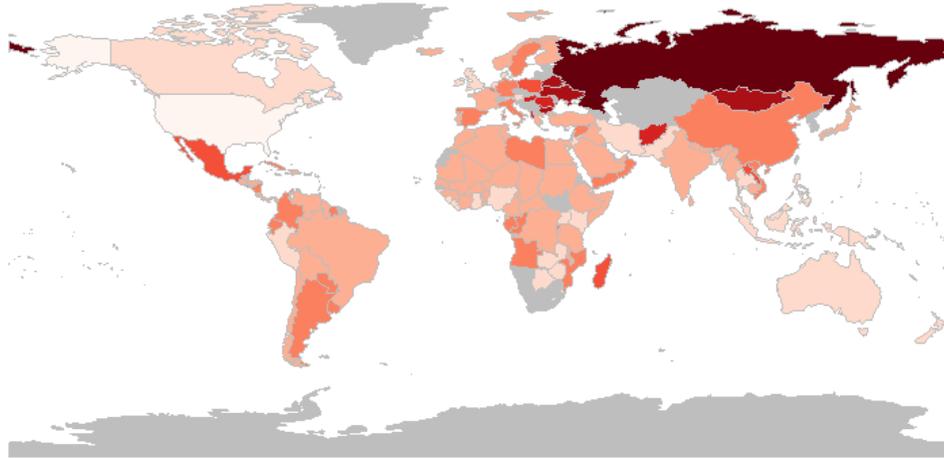
Note: The scores are estimated in the quanteda package (version '2.0.1') and quanteda.textmodels package (version '0.9.1') in R (Benoit et al.). I follow standard preprocessing during the tokenisation stage, remove English stopwords and perform stemming. I also trim the document-feature matrix to have features that appear at least ten times in five documents. The US is given a reference score (-1) and Russia (+1). Results are rescaled as proposed by Martin and Vanberg (2007), hence predicted scores are within the range (-1 ; +1).

Figure 17: Wordscores plot
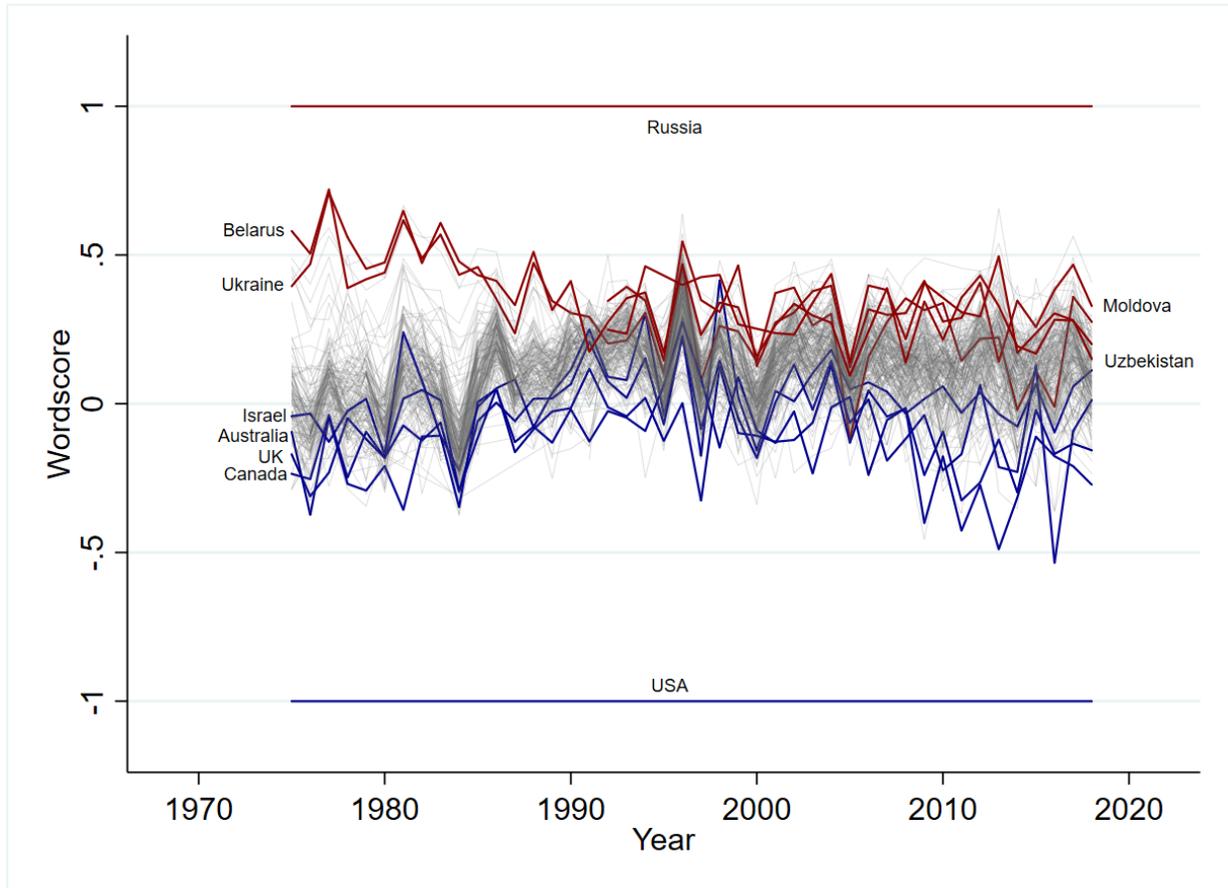
The graph plots the wordscore of 194 countries over a period of 43 years. To allow for time comparisons, raw wordscores are rescaled as proposed by Martin and Vanberg (2007). Hence, estimates shall be interpreted as relative distances of each country's speeches from the US and Russia.

4.3.2 Bilateral aids

I collect aid data on the US and other DAC donors from the World Bank. Specifically, I use net bilateral aid flows (current US$) and convert them into real dollars using the consumer price index (2010 = 100) of the World Bank.

Finding comprehensive and structured information on Soviet and Russian development aid is a challenging task. A complete account is yet to be produced and will probably be available only if Moscow authorizes an official release of documents. During the Soviet years, scholars have resorted to reports produced by the United States Central Intelligence Agency and to the replication data for Dannehl (1995) book. With the end of the Soviet Union, Russia ceased its development programs altogether. Only in 2007, under the presidency of Vladimir Putin, Russia started to rebuild its aid ambitions, beginning to report its ODA

statistics to the OECD aid community in 2011.

I collect aid data on Russia from Dannehl (1995) book and Development OECD data. Such information is less comprehensive if compared to its US counterpart, nevertheless, it covers from 1975 to 1989 for the Soviet aid and from 2011 to 2018 for Putin's modern Russia.

4.3.3 Other controls

Data on socio-economic variables such as GDP, PPP (constant 2017 international $), total population and trades are collected from the Development OECD Data. Data about the level of electoral democracy, liberal democracy, freedom of expression, and share of the population with suffrage are from the Democracy/Autocracy Dataset of Ulfelder (2012), while data on the militaristic political leaderships is from the Database of Political Institutions (Cruz et al. 2018). To reduce the sparsity of observations on Russia, I also collect GDP historical data from the Maddison Project and historical data from Charles Dannehl's 1995 book.

4.3.4 Method of estimation

The research uses panel data for all available UN countries from 1975 to 2018. In order to test whether countries' statements at the UN General Debate matter for the allocation of American and Russian aid, I estimate a set of regressions using ordinary least squares (OLS).

$$Aid_{c,t}^{USA} = \alpha_{c,t} + \beta wordscore_{c,t} + \gamma X_{c,t} + \vartheta Aid_{c,t-1}^{USA} + \varepsilon_{c,t} \tag{10}$$

$$Aid_{c,t}^{Russia} = \alpha_{c,t} + \beta wordscore_{c,t} + \gamma X_{c,t} + \vartheta Aid_{c,t-1}^{Russia} + \varepsilon_{c,t} \tag{11}$$

$Aid_{c,t}$ represents the amount of bilateral aid (constant US$, millions) that a country c in year t received from the USA or Russia. $\alpha_{c,t}$ represents country and year fixed effects while $X_{c,t}$ is a list of socio-economic controls such as GDP, population, trades, level of democracy, freedom and government type. β will measure the relationship between the UN countries' alignment with the US/Russia and the aid received from the US/Russia.

Since some of the data are not available for all countries or years, the number of observations depends on the choice of explanatory variables. For transparency, the results will be shown using both the baseline model and the extended model with all control variables included.

Finally, to decompose the effect by geographical area, I perform the same regression analysis including only recipients' countries from one continent mong: Asia, Africa, Americas, Oceania and Europe.

Given the empirical strategy employed, the estimated coefficients are not necessarily causal, however, I present a set of robustness tests that tend to indicate the effect is indeed going from the political affinity captured by the speeches and the allocation of the aid.

4.4 Results

4.4.1 UN speeches and aid recipiency

Table 26 shows the results of regression equations 1 and 2. All estimates include country and year fixed-effects, and standard errors are clustered at the country level. The estimates show that countries with a higher index receive more aid from Russia and less from the US and vice versa. In other words, the closer a country's speech is to the speech of one of the two world superpowers, the more aid these superpowers will allocate to such a country.

Table 26: Relationship between International Aids and Wordscore (1975-2018)

	USA aids			USSR aids		
	(1)	(2)	(3)	(4)	(5)	(6)
Wordscore Index	-175.9**	-94.87**	-100.8**	32.87**	35.83**	48.72**
	(85.66)	(45.58)	(46.46)	(13.76)	(15.18)	(24.10)
constant	126.3***	3.677	-24.09	5.357***	24.71***	30.29***
	(10.26)	(63.61)	(50.14)	(0.548)	(3.726)	(3.492)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Socio-economic controls	No	Yes	Yes	No	Yes	Yes
Lag of Aids included	No	No	Yes	No	No	Yes
N. country	162	148	148	50	44	44
N. observations	4993	3311	3224	701	628	586

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Standard errors (in parentheses) are clustered at the country level. Observations are available for the US from 1975 to 2018. Observations are available for USSR from 1975 to 1989. Socio-economic controls include GDP, population, trade, level of electoral democracy, level of liberal democracy, level of freedom of expression, the share of the population with suffrage, and militaristic political leadership.

Overall, an increase of 1 unit in the index corresponds to a decrease of around 100 million US dollars in aid from the US. On the other hand, the same 1 unit increase in the index coincides with an increase of around 50 million dollars in aid from Russia. The coefficients

are significant at 5% level in all specifications. It is reassuring to observe that the estimated result is not sensitive to the inclusion of control variables such as democracy and freedom (see columns 3 and 6), suggesting that the use of text as data approach is indeed capturing a more subtle and multidimensional representation of countries' preferences.

Table 27 presents the results of regression equations 1 and 2 decomposed by geographical area. Panel A shows that the US rewards countries' alignment in Africa, Asia, and the Americas, but not in Europe and Oceania. As such, despite the fact that European countries' statements were typically among the closest to the US ones, they are not the driver of the overall effect. Panel B shows that Russia rewards countries' alignment only in Asia.

Table 27: Wordscore effect on International Aids decomposed by geographical area (1975-2018)

	Africa	Asia	Europe	Americas	Oceania
	(1)	(2)	(3)	(4)	(5)
Panel A: USA aids					
Wordscore Index	-119.3*	-414.7*	75.91	-119.3**	-7.684
	(62.54)	(206.6)	(78.72)	(49.79)	(14.74)
	[2027]	[1253]	[327]	[1116]	[257]
Constant	112.7***	235.3***	61.77**	65.45***	18.63***
	(7.769)	(24.66)	(19.62)	(3.802)	(1.730)
Panel B: USSR aids					
Wordscore Index	12.41	91.36*		1.187	
	(10.70)	(46.62)		(26.65)	
	[238]	[181]	[28]	[231]	[23]
Constant	2.126***	13.08***		2.906***	
	(0.485)	(2.861)		(0.457)	

Note: Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors (in parentheses) are clustered at the country level. Number of observations in square brackets. Observations are available for the US from 1975 to 2018. Observations are available for USSR from 1975 to 1989.

4.4.2 Robustness tests

To check that the results are not given by construction, figure 18 shows the results of the main estimation using placebo outcome variables such as aids from other DAC donors. Countries

such as France, Italy, Sweden, Japan, Poland etc. should not allocate more or less aid to recipient countries depending on whether their speech pleases more the US or Russia. It is reassuring to observe that none of the placebo estimates are statistically significant, confirming the goodness of the analysis.

Figure 18: Placebo outcome estimates (1975-2018)
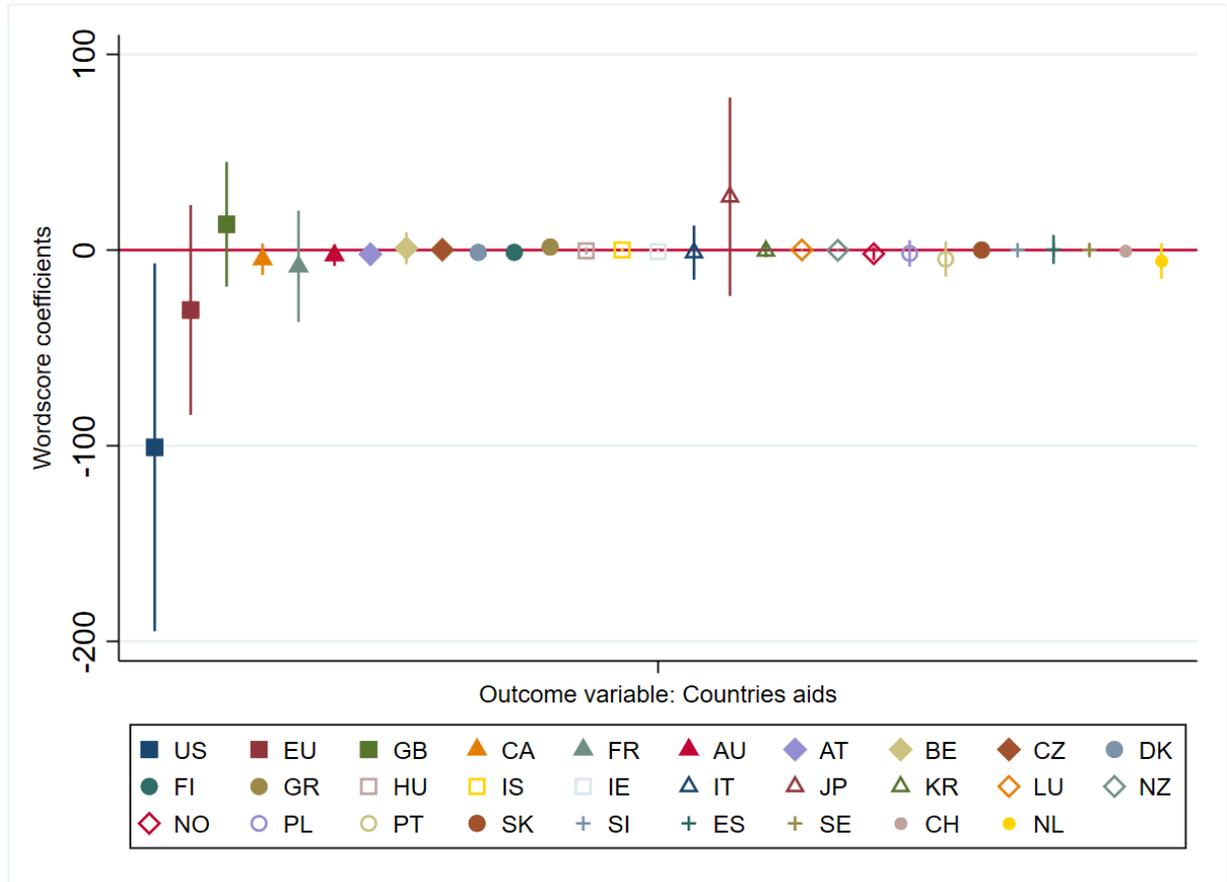
Note: The graph plots the main regression coefficients of the different placebo regressions, with each regression having as an outcome variable the aid received from a different placebo country. All estimations include country and year fixed effects, as well as socio-economic controls and lagged outcome variable. Confidence interval at 95%.

To verify that the results are not sensitive to some of the text analysis pre-processing choices, table 28 shows the main results using an index built with different trimming decisions.

Table 28: Trimming sensitivity test

	USA aids			
	(1)	(2)	(3)	(4)
Wordscore Index	-179.0**	-186.5**	-167.7**	-161.5**
	(85.39)	(88.90)	(82.78)	(80.99)
Constant	125.8***	126.0***	125.3***	124.9***
	(9.828)	(9.891)	(9.906)	(9.858)
Min term frequency	8	4	12	15
Min document frequency	3	2	8	10
N. country	162	162	162	162
N. observations	4993	4993	4993	4993

Note: Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors (in parentheses) are clustered at the country level. Min term/document frequency indicates the trimming details employed to construct the Wordscore index.

Despite the use of a large and long panel dataset, with a generous set of control variables and fixed effects, it is not possible to exclude the threat typically posed by inverse causality. Table 29 has the objective to argue that the estimated effect is indeed going from the UN speeches to aid allocation. To investigate the directionality of the discovered effect, table 29 shows the results using a symmetric-inverted version of regression equations 1 and 2 (by flipping bilateral aids with the Wordscore index). It is reassuring to observe that by inverting our dependent and independent variables of interest, we lose statistical significance on the main coefficient. This is likely to indicate that UN speeches affect aid allocation and not vice versa. The US/URSS aids of the present and of the past do not affect the Wordscore index.

Table 29: Directionality test

	(1)	(2)	(3)	(4)
		Wordscore Index		
Panel A: USA				
USA aids	-0.00000752			
	(0.00000677)			
	[3251]			
L.USA aids		-0.00000379		
		(0.00000274)		
		[3245]		
L2.USA aids			-0.00000615*	
			(0.00000328)	
			[3231]	
L3.USA aids				-0.00000332
				(0.00000587)
				[3210]
Panel B: USSR				
URSS aids	0.000183*			
	(0.0000950)			
	[567]			
L.URSS aids		0.0000321		
		(0.0000615)		
		[567]		
L2.URSS aids			0.0000341	
			(0.000107)	
			[528]	
L3.URSS aids				-0.0000235
				(0.000124)
				[488]

Note: Significance levels: *** p<0.01, ** p<0.05, * p<0.1. symmetric-inverted version of regression equations 1 and 2. Standard errors (in parentheses) are clustered at the country level. Observations are available for the US from 1975 to 2018. Observations are available on USSR from 1975 to 1989. Country and year fixed-effects included. Socio-economic controls and lagged outcome variable as a control included.

Furthermore, to check that there is no effect coming from future speeches into present aid allocation, I perform the main regressions 1 and 2 including forward-lagged variables of the Wordscore index in the same regression as ulterior covariates. The rationale of the test is that in the absence of inverse causality, speeches performed in the future should not affect the aid allocation of the past. Figure 19 plots the coefficients of the forward-lagged Wordscore indexes. It is reassuring to notice that all forward-lagged coefficients of the independent variable of interest (Wordscore index) are statistically insignificant.

Figure 19: Forward lag placebo test
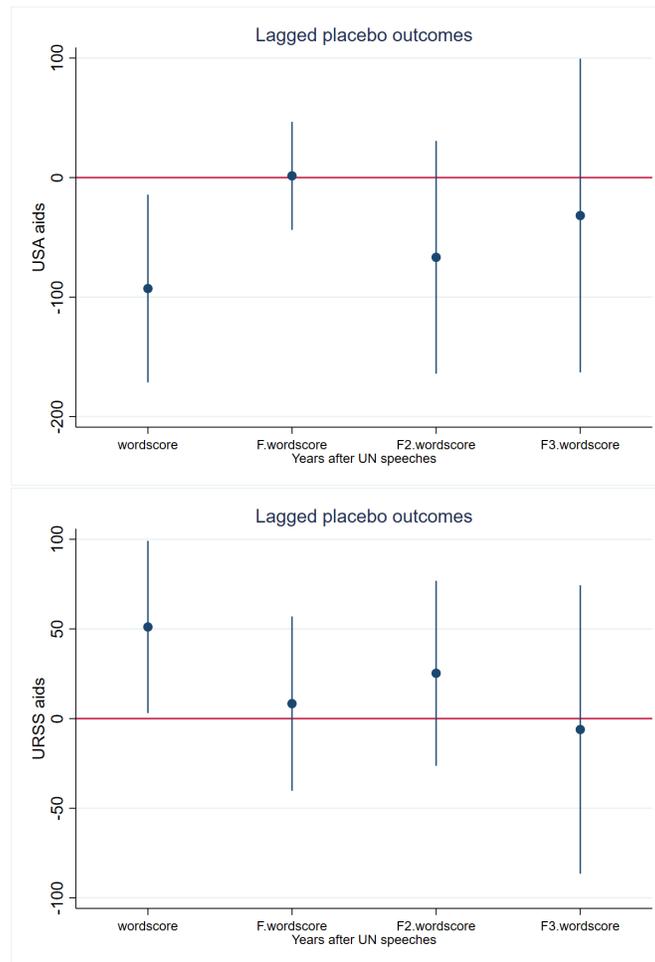
Note: Coefficients of the forward lag of the independent variable wordscore are shown. The dependent variable is international aid. Country and year fixed effects included. Socio-economic controls included. Standard error clustered at the country level. Confidence interval at 95%.

4.4.3 Case study: the 2010 Haiti earthquake

To explore the found relationship between UN speeches and Aid allocation from a more counterfactual standpoint, I focus on a case study, namely the 2010 Haiti earthquake, by employing the synthetic control method (SCM) of Abadie et al. (2015, 2010).

In 2010, a magnitude 7.0 earthquake struck Haiti, killing 220,000 people, forcing 1.5 million individuals into homelessness, and bringing the entire country to its knees. Such catastrophic and exogenous event generated one of the largest disaster relief efforts in modern history launched by the international community (Kligerman et al. 2015).

The magnitude of such relief effort can be observed in figure 20, depicting 2010 as the historical absolute peak in the US aids toward Haiti. On the other hand, the Russian relief support has been more modest, constrained by the country's financial constraints of the 2000s.

Figure 20: Historical plot of US aids toward Haiti
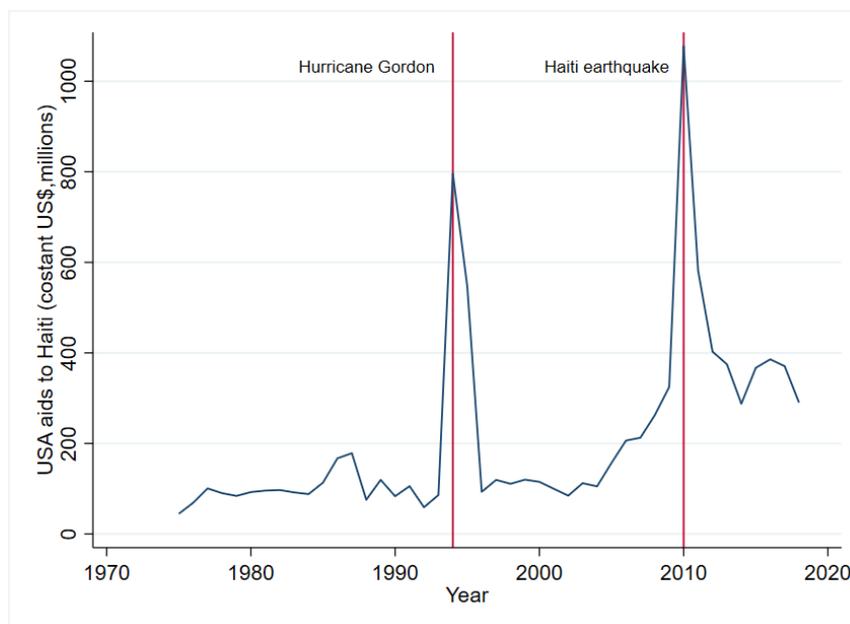
Note: The line plot shows the trend of US bilateral aid to Haiti from 1975 to 2018. Main events represented by the vertical lines: the Hurricane Gordon (1994) and the Haiti earthquake (2010).

The SCM is particularly suitable for this case as I have only one treated country (Haiti) and many non-treated countries (all other UN countries). Hence, I can evaluate the impact of the 2010 Haiti earthquake on the Wordscore Index. If the paper's argument stands, the 2010 Haiti earthquake should have had a negative effect on the Wordscore Index (speeches closer to the USA) as a means to attract as much US aid as possible. Furthermore, since the treatment is temporary and not permanent, the effect should disappear a few years after

the earthquake, being the urgent need of attracting aid ceased.

Figure 21 presents the results of the SCM. The synthetic Haiti almost perfectly matches the real Haiti in the pre-treatment period, diverging then sharply in 2010. The negative effect of the Haiti earthquake on the Wordscore index disappear around 2015, as things go back to normal. It is reassuring to observe that the treatment effect matches our expectations. The fact that the effect on the speeches does not last much longer after the earthquake makes it likely that the speech change was indeed to attract aids and not a consequence of it.

Figure 21: Trends in Wordscore index: Haiti versus synthetic Haiti
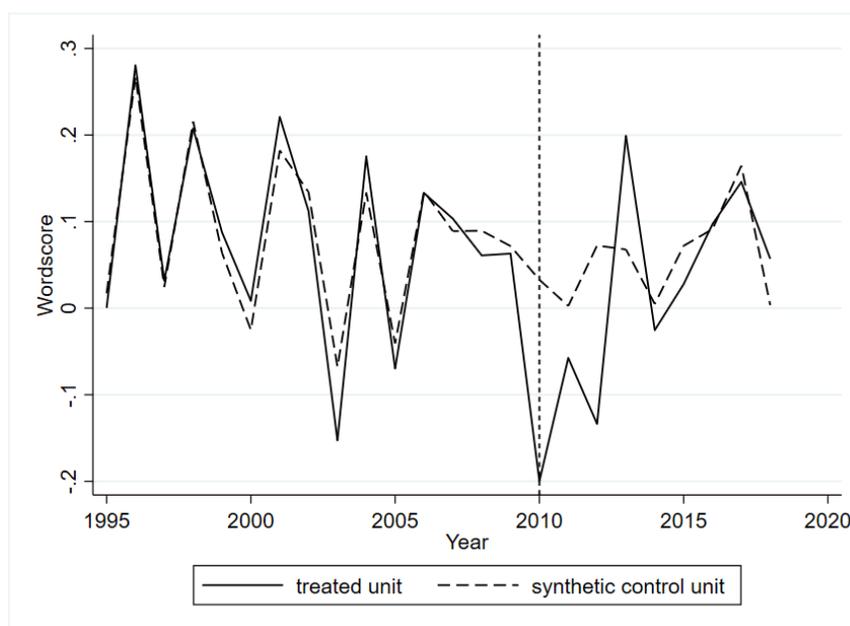
Note: The graph plots the actual versus the synthetic wordscore of Haiti. The 2010 Haiti earthquake is defined as the exogenous treatment. The wordscore index lags are used as predictors. All countries for which the predictor is available in the entire pre-treatment period (149 countries) are employed for the synthetic control method donor pool.

4.5   Conclusions

The presented project originated from the idea that the United Nations General Debate is a world platform where countries' leaders can freely express their country's point of view on any topic they deem necessary. Such an open platform can be used for a multitude of purposes: from expressing government preferences on established issues to bringing unheard matters to the world's attention.

I use machine learning and text as data approach to build a quantitative measure of UN countries' preferences based on their statements at the UN General Debate. Thanks to the supervised machine learning algorithm "wordscore", countries are scaled based on their

speech similarities to the US and Russia for all years from 1975 to 2018. Panel analysis shows that UN statements have a relevant role in international aid allocation: countries whose speeches please the US receive more aid from them, while countries whose speeches please Russia receive more aid from Russia.

The magnitude of the effect is economically relevant as a reduction of 1 unit in the index equates to an increase of 100 million dollars in aid from the US. The same reduction of 1 unit in the index is linked with a drop of 50 million dollars in aid from Russia. Results are robust to the inclusion of country-year fixed-effects as well as control variables such as GDP, population, trade, democracy level, freedom level and government type.

Looking at the effect decomposed by geographic area, it is observed that the US rewards speech alignment, particularly towards countries located in Africa, Asia and Latin America. Russia, instead, focuses on Asia.

To the best of my knowledge, this is the first research that studies international aid allocation by employing a "wordscore" index as a measure of countries' preferences within a panel regression framework. This research contributes to the existing literature on international aid by finding empirical evidence on the relevance of UN members' speeches (as a measurement of government preferences) for aid allocation decisions. Existing literature has shown that countries can signal political alignment through strategic voting and be rewarded. The presented work also suggests that speech alignment can explain aid recipiency.

A   Appendix

Derivation 1 (chapter 2):

The swing voter in group J is:

$$\sigma^J = W^J(g_I^A, g_C^A) - W^J(g_I^B, g_C^B) - \delta \tag{1}$$

Given our distributional assumption of ideological bias, the amount of votes for candidate A in group J will be:

$$\left(\sigma^J + \frac{1}{2}\phi^J \right) \phi^J$$

$$= \frac{1}{2} + \sigma^J \phi^J$$

$$= \frac{1}{2} + \phi^J \left[W^J(g_I^A, g_C^A) - W^J(g_I^B, g_C^B) \right] - \delta\phi^J \tag{2}$$

The total votes for candidate A in all groups will be:

$$\pi_A = \sum_J \alpha^J \frac{1}{2} + \sum_J \alpha^J \phi^J \left[W^J(g_I^A, g_C^A) - W^J(g_I^B, g_C^B) \right] - \delta \sum_J \alpha^J \phi^J \tag{3}$$

Candidate A wins the election if she obtains more than half of the votes $\pi_A > \frac{1}{2}$:

$$\pi_A = \frac{1}{2} + \sum_J \alpha^J \phi^J \left[W^J(g_I^A, g_C^A) - W^J(g_I^B, g_C^B) \right] - \delta\Phi > \frac{1}{2}$$

$$\Rightarrow \sum_J \alpha^J \phi^J \left[W^J(g_I^A, g_C^A) - W^J(g_I^B, g_C^B) \right] - \delta\Phi > 0$$

Candidate A does not know her popularity δ. Hence, she maximizes the probability of being elected subject to popularity shocks. Given our distributional assumption of popularity, candidate A probability of winning becomes:

$$P_A = Pr(\delta < \underline{\delta}) = \left[\pi_A > \frac{1}{2} \right] = \frac{1}{2} + \frac{\psi}{\phi} \left[\sum_J \alpha^J \phi^J \left[W^J(g_I^A) - W^J(g_I^B) \right] \right] \tag{4}$$

Since $w^i = c^i + H(g_I, g_C)$, $c^i = \left(\frac{y^i}{y} \right)(y - (g_I + g_C))$ and $H(g_I, g_C) = ln(g_C) + \frac{1}{\gamma} \cdot e \cdot ln(g_I)$. The overall voter i's utility function is:

$$W^i(g_I, g_C) = \left(\frac{y^i}{y} \right)(y - (g_I + g_C)) + ln(g_C) + \frac{1}{\gamma} \cdot e \cdot ln(g_I) \tag{5}$$

By substituting (A5) in (A4), I obtain:

$$P_A = \frac{1}{2} + \frac{\psi}{\phi}\left[\sum_J \alpha^J \phi^J \left[\left(\left(\frac{y^J}{y}\right)(y - (g_I^A + g_C^A)) + ln(g_C^A) + \frac{1}{\gamma} \cdot e^A \cdot ln(g_I^A)\right) - \right.\right.$$
$$\left.\left.\left(\left(\frac{y^J}{y}\right)(y - (g_I^B + g_C^B)) + ln(g_C^B) + \frac{1}{\gamma} \cdot e^B \cdot ln(g_I^B)\right)\right]\right]$$

Derivation 2 (chapter 2):

$$\underset{g_I^A, g_C^A}{\text{Max}}\frac{1}{2} + \frac{\psi}{\phi}\left[\sum_J \alpha^J \phi^J \left[\left(\left(\frac{y^J}{y}\right)(y - (g_I^A + g_C^A)) + ln(g_C^A) + \frac{1}{\gamma} \cdot e^A \cdot ln(g_I^A)\right) - \right.\right.$$
$$\left.\left.\left(\left(\frac{y^J}{y}\right)(y - (g_I^B + g_C^B)) + ln(g_C^B) + \frac{1}{\gamma} \cdot e^B \cdot ln(g_I^B)\right)\right]\right]$$

subject to

$$ty = g_I^A + g_C^A$$

The FOCs of the Lagrangian are as following:

$$\frac{\vartheta L}{\vartheta g_I^A} = 0$$

$$\Rightarrow \frac{\psi}{\phi}\sum_J \alpha^J \phi^J \left[-\frac{y_J}{y} + \frac{\frac{1}{\gamma} \cdot e^A}{g_I^A}\right] + \lambda = 0$$

$$\frac{\vartheta L}{\vartheta g_C^A} = 0$$

$$\Rightarrow \frac{\psi}{\phi}\sum_J \alpha^J \phi^J \left[-\frac{y_J}{y} + \frac{1}{g_C^A}\right] + \lambda = 0$$

It follows that:

$$\lambda = \lambda$$

$$\Rightarrow \frac{\psi}{\phi}\sum_J \alpha^J \phi^J \left[-\frac{y_J}{y} + \frac{\frac{1}{\gamma} \cdot e^A}{g_I^A}\right] = \frac{\psi}{\phi}\sum_J \alpha^J \phi^J \left[-\frac{y_J}{y} + \frac{1}{g_C^A}\right]$$

$$\Rightarrow -\sum_J \frac{\alpha^J \phi^J y^J}{y} + \sum_J \alpha^J \phi^J \frac{\frac{1}{\gamma} \cdot e^A}{g_I^A} = -\sum_J \frac{\alpha^J \phi^J y^J}{y} + \sum_J \alpha^J \phi^J \frac{1}{g_C^A}$$

$$\Rightarrow \frac{\frac{1}{\gamma} \cdot e^A}{g_I^A} = \frac{1}{g_C^A}$$

$$\Rightarrow g_C^A = \frac{g_I^A}{\frac{1}{\gamma} \cdot e^A} \tag{6}$$

By substituting it (A6) in the budget constraint

$$ty = \frac{g_I^A}{\frac{1}{\gamma} \cdot e^A} + g_I^A$$

$$ty = g_I^A \left(\frac{1}{\frac{1}{\gamma} \cdot e^A} + 1 \right)$$

$$g_I^A = \frac{ty}{\left(\frac{1}{\frac{1}{\gamma} \cdot e^A} + 1 \right)} \tag{7}$$

By substituting it (A7) in the budget constraint

$$ty = \frac{ty}{\left(\frac{1}{\frac{1}{\gamma} \cdot e^A} + 1 \right)} + g_C^A$$

$$\Rightarrow g_C^A = ty - \frac{ty}{\left(\frac{1}{\frac{1}{\gamma} \cdot e^A} + 1 \right)} \tag{8}$$

Algorithm explanation (chapter 4):

How does Wordscore work?

I define $C_{i,j}$ as a document-features matrix of reference document i and word j. A document feature matrix is a matrix containing information on how many times all words are repeated in each document. I also define a_i as the score associated with reference document i (known by the researcher). In our case, the reference documents are the USA and Russia speeches and their score known by the researcher is their diametrical opposite ideological position. The algorithm starts working on the reference documents by computing $F_{i,j} = C_{i,j}/C_i$, where $C_i = \sum_{j=1}^{J}(C_{i,j})$. It normalizes the document-feature matrix $P_{i,j} = \frac{F_{i,j}}{\sum_{i=1}^{I}(F_{i,j})}$ and estimate the probability that given the observation of word j , we are reading a certain reference document i with $S_j = \sum_{i=1}^{I}(a_i \times P_{i,j})$.

Then the algorithm moves to the unlabelled documents (k) and compute the Wordscore Index following: $Wordscore_k = \sum_j(F_{k,j} \times s_j)$, where $F_{k,j} = C_{k,j}/C_k$. In our case, the unlabelled documents are the speeches of all countries (the US and Russia excluded).

References

Abadie, A., Diamond, A. & Hainmueller, J. (2010), 'Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program', *Journal of the American statistical Association* **105**(490), 493–505.

Abadie, A., Diamond, A. & Hainmueller, J. (2015), 'Comparative politics and the synthetic control method', *American Journal of Political Science* **59**(2), 495–510.

Acconcia, A., Ronza, C. et al. (2021), Women's representation in politics and government stability, Technical report, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.

Adler, N. J. (1996), 'Global women political leaders: An invisible history, an increasingly important future', *The Leadership Quarterly* **7**(1), 133–161.

Adler, N. J. (1997), 'Global leadership: Women leaders', *MIR: Management International Review* pp. 171–196.

Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N. & Gangadharan, L. (2009), 'Gender, culture, and corruption: Insights from an experimental analysis', *Southern Economic Journal* **75**(3), 663–680.

Alesina, A. & Dollar, D. (2000), 'Who gives foreign aid to whom and why?', *Journal of economic growth* **5**(1), 33–63.

Alesina, A. & Weder, B. (2002), 'Do corrupt governments receive less foreign aid?', *American Economic Review* **92**(4), 1126–1137.

Anzia, S. F. & Berry, C. R. (2011), 'The jackie (and jill) robinson effect: why do congresswomen outperform congressmen?', *American Journal of Political Science* **55**(3), 478–493.

Atkeson, L. R. (2003), 'Not all cues are created equal: The conditional impact of female candidates on political engagement', *The Journal of Politics* **65**(4), 1040–1061.

Atkeson, L. R. & Carrillo, N. (2007), 'More is better: The influence of collective female descriptive representation on external efficacy', *Politics & Gender* **3**(1), 79–101.

Bagues, M. & Campa, P. (2021), 'Can gender quotas in candidate lists empower women? evidence from a regression discontinuity design', *Journal of Public Economics* **194**, 104315.

94

Baltrunaite, A., Casarico, A., Profeta, P. & Savio, G. (2019), 'Let the voters choose women', *Journal of Public Economics* **180**, 104085.

Bari, F. (2010), 'Women parliamentarians: Challenging the frontiers of politics in pakistan', *Gender, Technology and Development* **14**(3), 363–384.

Baskaran, T., Bhalotra, S. R., Min, B. K. & Uppal, Y. (2018), 'Women legislators and economic performance'.

Baskaran, T. & Hessami, Z. (2018), 'Does the election of a female leader clear the way for more women in politics?', *American Economic Journal: Economic Policy* **10**(3), 95–121.

Baturo, A., Dasandi, N. & Mikhaylov, S. J. (2017), 'Understanding state preferences with text as data: Introducing the un general debate corpus', *Research & Politics* **4**(2), 2053168017712821.

Bauer, G. & Okpotor, F. (2013), '"her excellency": An exploratory overview of women cabinet ministers in africa', *Africa Today* **60**(1), 77–97.

Bauhr, M., Charron, N. & Wängnerud, L. (2019), 'Exclusion or interests? why females in elected office reduce petty and grand corruption', *European Journal of Political Research* **58**(4), 1043–1065.

Beaman, L. et al. (2007), Women politicians, gender bias, and policy-making in rural india, Technical report.

Becker, G. (1964), 'S. human capital: A theoretical and empirical analysis, with special reference to education', *Chicago: University of Press* **19**(6), 4.

Benoit, K. & Laver, M. (2008), 'Compared to what? a comment on" a robust transformation procedure for interpreting political text" by martin and vanberg', *Political Analysis* **16**(1), 101–111.

Bernstein, S. J. & Alpert, E. J. (1971), 'Foreign aid and voting behavior in united nations-admission of communist china', *Orbis-A Journal of World Affairs* **15**(3), 963–977.

Besley, T. (2005), 'Political selection', *Journal of Economic Perspectives* **19**(3), 43–60.

Besley, T. (2006), *Principled agents?: The political economy of good government*, Oxford University Press on Demand.

Besley, T. & Coate, S. (1997), 'An economic model of representative democracy', *The Quarterly Journal of Economics* **112**(1), 85–114.

Besley, T. & Coate, S. (2003), 'Centralized versus decentralized provision of local public goods: a political economy approach', *Journal of public economics* **87**(12), 2611–2637.

Besley, T., Montalvo, J. G. & Reynal-Querol, M. (2011), 'Do educated leaders matter?', *The Economic Journal* **121**(554), F205–227.

Bhalotra, S. & Clots-Figueras, I. (2014), 'Health and the political agency of women', *American Economic Journal: Economic Policy* **6**(2), 164–97.

Bhalotra, S., Clots-Figueras, I. & Iyer, L. (2018), 'Pathbreakers? women's electoral success and future political participation', *The Economic Journal* **128**(613), 1844–1878.

Bjarnegård, E. (2013), *Gender, informal institutions and political recruitment: Explaining male dominance in parliamentary representation*, Springer.

Block, S. A. (2001), 'Elections, electoral competitiveness, and political budget cycles in developing countries', *CID Working Paper Series* .

Bracco, E., Porcelli, F. & Redoano, M. (2019), 'Political competition, tax salience and accountability. theory and evidence from italy', *European Journal of Political Economy* **58**, 138–163.

Bratton, K. A. (2005), 'Critical mass theory revisited: The behavior and success of token women in state legislatures', *Politics & Gender* **1**(1), 97–125.

Bratton, K. A. & Haynie, K. L. (1999), 'Agenda setting and legislative success in state legislatures: The effects of gender and race', *The Journal of Politics* **61**(3), 658–679.

Breierova, L. & Duflo, E. (2004), 'The impact of education on fertility and child mortality: Do fathers really matter less than mothers?'.

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

Broockman, D. E. (2014), 'Do female politicians empower women to vote or run for office? a regression discontinuity approach', *Electoral Studies* **34**, 190–204.

Brown, R., Mansour, H. & O'Connell, S. D. (2021), 'Does local female political representation empower women to run for higher office? evidence from state and national legislatures in india', *The World Bank Economic Review* .

Buchanan, J. M., Buchanan, J. M. & Tullock, G. (1965), *The calculus of consent: Logical foundations of constitutional democracy*, Vol. 100, University of Michigan press.

Calonico, S., Cattaneo, M. D. & Titiunik, R. (2014a), 'Robust data-driven inference in the regression-discontinuity design', *The Stata Journal* **14**(4), 909–946.

Calonico, S., Cattaneo, M. D. & Titiunik, R. (2014b), 'Robust nonparametric confidence intervals for regression-discontinuity designs', *Econometrica* **82**(6), 2295–2326.

Campbell, D. E. & Wolbrecht, C. (2006), 'See jane run: Women politicians as role models for adolescents', *The Journal of Politics* **68**(2), 233–247.

Canovan, M. (1984), "people', politicians and populism 1', *Government and Opposition* **19**(3), 312–327.

Card, D. (1999), The causal effect of education on earnings, *in* 'Handbook of labor economics', Vol. 3, Elsevier, pp. 1801–1863.

Carreras, M. (2017), 'High-profile female executive candidates and the political engagement of women: a multilevel analysis', *Political Research Quarterly* **70**(1), 172–183.

Carreri, M. (2019), 'Can good politicians compensate for bad institutions? evidence from an original survey of italian mayors', *Evidence from an Original Survey of Italian Mayors (November 01, 2019)* .

Carroll, S. J. (2001), *The impact of women in public office*, Indiana University Press.

Caselli, F. & Morelli, M. (2004), 'Bad politicians', *Journal of Public Economics* **88**(3-4), 759–782.

Cattaneo, M. D., Jansson, M. & Ma, X. (2018), 'Manipulation testing based on density discontinuity', *The Stata Journal* **18**(1), 234–261.

Cattaneo, M. D., Jansson, M. & Ma, X. (2019), 'Simple local polynomial density estimators', *Journal of the American Statistical Association* pp. 1–7.

Chattopadhyay, R. & Duflo, E. (2004), 'Women as policy makers: Evidence from a randomized policy experiment in india', *Econometrica* **72**(5), 1409–1443.

Chhoeun, T., Sok, P. & Byrne, C. (2008), "citadel of women': strengthening female leadership in rural cambodia', *Gender & Development* **16**(3), 535–547.

Clarke, J. & Newman, J. (2017), "people in this country have had enough of experts': Brexit and the paradoxes of populism', *Critical Policy Studies* **11**(1), 101–116.

Clots-Figueras, I. (2011), 'Women in politics: Evidence from the indian states', *Journal of public Economics* **95**(7-8), 664–690.

Clots-Figueras, I. (2012), 'Are female leaders good for education? evidence from india', *American Economic Journal: Applied Economics* **4**(1), 212–44.

Collins, H., Evans, R., Durant, D. & Weinel, M. (2020), 'Experts and the will of the people', *Cham: Palgrave Macmillan* .

Cowper-Coles, M. (2020), 'Women political leaders: the impact of gender on democracy', *The Global Institute for Women's Leadership* .

Cruz, C., Keefer, P. & Scartascini, C. (2018), 'The database of political institutions 2015 (dpi2015)', *Washington, DC: Inter-American Development Bank, Washington, DC* .

Dannehl, C. R. (1995), *Politics, Trade, and Development: Soviet Economic Aid to the Non-communist Third World, 1955-89*, Dartmouth Publishing Group.

Dassonneville, R. & McAllister, I. (2018), 'Gender, political knowledge, and descriptive representation: The impact of long-term socialization', *American Journal of Political Science* **62**(2), 249–265.

De Mesquita, B. B. & Smith, A. (2009), 'A political economy of aid', *International Organization* pp. 309–340.

Decarolis, F., Fisman, R., Pinotti, P., Vannutelli, S. & Wang, Y. (2022), 'Gender and bureaucratic corruption: evidence from two countries', *The Journal of Law, Economics, and Organization* .

Dee, T. S. (2004), 'Are there civic returns to education?', *Journal of public economics* **88**(9-10), 1697–1720.

Desjardins, R. (2003), 'Determinants of literacy proficiency: A lifelong-lifewide learning perspective', *International Journal of Educational Research* **39**(3), 205–245.

Dillon, M. R., Kannan, H., Dean, J. T., Spelke, E. S. & Duflo, E. (2017), 'Cognitive science in the field: A preschool intervention durably enhances intuitive but not formal mathematics', *Science* **357**(6346), 47–55.

Dolan, K. (2006), 'Symbolic mobilization? the impact of candidate sex in american elections', *American Politics Research* **34**(6), 687–704.

Dollar, D., Fisman, R. & Gatti, R. (2001), 'Are women really the "fairer" sex? corruption and women in government', *Journal of Economic Behavior & Organization* **46**(4), 423–429.

Downs, A. et al. (1957), 'An economic theory of democracy'.

Dreher, A., Lamla, M. J., Lein, S. M. & Somogyi, F. (2009), 'The impact of political leaders' profession and education on reforms', *Journal of comparative economics* **37**(1), 169–193.

Duflo, E. (2001), 'Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment', *American economic review* **91**(4), 795–813.

Duncan, G. J. (2003), 'Modeling the impacts of child care quality on children's preschool cognitive development', *Child development* **74**(5), 1454–1475.

Eggers, A. C., Vivyan, N. & Wagner, M. (2018), 'Corruption, accountability, and gender: Do female politicians face higher standards in public life?', *The Journal of Politics* **80**(1), 321–326.

Einstein, K. L. & Glick, D. M. (2017), 'Cities in american federalism: Evidence on state–local government conflict from a survey of mayors', *Publius: The Journal of Federalism* **47**(4), 599–621.

Einstein, K. L., Glick, D. & Palmer, M. (2018), '2017 menino survey of mayors'.

Elango, S., García, J. L., Heckman, J. J. & Hojman, A. (2015), Early childhood education, *in* 'Economics of Means-Tested Transfer Programs in the United States, Volume 2', University of Chicago Press, pp. 235–297.

Esarey, J. & Chirillo, G. (2013), '"fairer sex" or purity myth? corruption, gender, and institutional context', *Politics & Gender* **9**(4), 361–389.

Esarey, J. & Schwindt-Bayer, L. A. (2018), 'Women's representation, accountability and corruption in democracies', *British Journal of Political Science* **48**(3), 659–690.

Esteve-Volart, B. & Bagues, M. (2012), 'Are women pawns in the political game? evidence from elections to the spanish senate', *Journal of public Economics* **96**(3-4), 387–399.

Fabbrini, S. (2001), 'Features and implications of semi-parliamentarism: the direct election of italian mayors', *South European society & politics* **6**(2), 47–70.

Ferreira, F. & Gyourko, J. (2014), 'Does gender matter for political leadership? the case of us mayors', *Journal of Public Economics* **112**, 24–39.

Fox, R. L. & Lawless, J. L. (2004), 'Entering the arena? gender and the decision to run for office', *American Journal of Political Science* **48**(2), 264–280.

Fraga, L. R., Lopez, L., Martinez-Ebers, V. & Ramirez, R. (2007), 'Gender and ethnicity: Patterns of electoral success and legislative advocacy among latina and latino state officials in four states', *Journal of Women, Politics & Policy* **28**(3-4), 121–145.

Freier, R. & Thomasius, S. (2016), 'Voters prefer more qualified mayors, but does it matter for public finances? evidence for germany', *International Tax and Public Finance* **23**(5), 875–910.

Fridkin, K. L. & Kenney, P. J. (2014), 'How the gender of us senators influences people's understanding and engagement in politics', *The Journal of Politics* **76**(4), 1017–1031.

Gagliarducci, S. & Paserman, M. D. (2012), 'Gender interactions within hierarchies: evidence from the political arena', *The Review of Economic Studies* **79**(3), 1021–1052.

Garwin, D. A. (1993), 'Building a learning organization', *Harvard business review* **71**(4), 73–91.

Gates, S., Hoeffler, A. et al. (2004), Global aid allocation: Are nordic donors different?, Technical report, Centre for the Study of African Economies, University of Oxford.

Geys, B. & Sørensen, R. J. (2019), 'The impact of women above the political glass ceiling: Evidence from a norwegian executive gender quota reform', *Electoral Studies* **60**, 102050.

Gilardi, F. (2015), 'The temporary importance of role models for women's political representation', *American Journal of Political Science* **59**(4), 957–970.

Glewwe, P. & King, E. M. (2001), 'The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter?', *The world bank economic review* **15**(1), 81–113.

Goetz, A. M. (1998), 'Women in politics & gender equity in policy: South africa & uganda', *Review of African Political Economy* **25**(76), 241–262.

Goetz, A. M. (2002), 'No shortcuts to power: constraints on women's political effectiveness in uganda', *The Journal of Modern African Studies* **40**(4), 549–575.

Golden, M. A. & Picci, L. (2005), 'Proposal for a new measure of corruption, illustrated with italian data', *Economics & Politics* **17**(1), 37–75.

Gupta, S., Liu, E. X. & Mulas-Granados, C. (2016), 'Now or later? the political economy of public investment in democracies', *European Journal of Political Economy* **45**, 101–114.

Hahn, J., Todd, P. & Van der Klaauw, W. (2001), 'Identification and estimation of treatment effects with a regression-discontinuity design', *Econometrica* **69**(1), 201–209.

Harkness, A. (2017), 'Leading beyond limits: Mayoral powers in the age of new localism.'.

Hayo, B. & Neumeier, F. (2014), 'Political leaders' socioeconomic background and fiscal performance in germany', *European Journal of Political Economy* **34**, 184–205.

Hessami, Z. & da Fonseca, M. L. (2020), 'Female political representation and substantive effects on policies: A literature review', *European Journal of Political Economy* **63**, 101896.

Hoeffler, A. & Outram, V. (2011), 'Need, merit, or self-interest—what determines the allocation of aid?', *Review of Development Economics* **15**(2), 237–250.

Holman, M. (2015), *Women in politics in the American city*, JSTOR.

Jankin Mikhaylov, S., Baturo, A. & Dasandi, N. (2017), 'United Nations General Debate Corpus'.
URL: *https://doi.org/10.7910/DVN/0TJX8Y*

Jayal, N. G. (2006), 'Engendering local democracy: The impact of quotas for women in india's panchayats', *Democratisation* **13**(1), 15–35.

Jeydel, A. & Taylor, A. J. (2003), 'Are women legislators less effective? evidence from the us house in the 103rd-105th congress', *Political Research Quarterly* **56**(1), 19–27.

Jha, C. K. & Sarangi, S. (2018), 'Women and corruption: What positions must they hold to make a difference?', *Journal of Economic Behavior & Organization* **151**, 219–233.

Kamlongera, A. P. (2008), 'Malawian women's participation in state politics: what are the constraints?', *Gender & Development* **16**(3), 471–480.

Kato, M. (1969), 'A model of us foreign aid allocation: an application of a rational decision-making scheme', *Approaches to Measurement in International Relations, Appleton Century Croft, New York* .

Katsimi, M. & Sarantides, V. (2012), 'Do elections affect the composition of fiscal policy in developed, established democracies?', *Public Choice* **151**(1-2), 325–362.

Kegley Jr, C. W. & Hook, S. W. (1991), 'Us foreign aid and un voting: Did reagan's linkage strategy buy deference or defiance?', *International Studies Quarterly* **35**(3), 295–312.

Kennett, R. (1990), 'Equilibrium political budget cycles', *The American Economic Review* **80**.

Kligerman, M., Barry, M., Walmer, D. & Bendavid, E. (2015), 'International aid and natural disasters: a pre-and post-earthquake longitudinal study of the healthcare infrastructure in leogane, haiti', *The American journal of tropical medicine and hygiene* **92**(2), 448–453.

Ladam, C., Harden, J. J. & Windett, J. H. (2018), 'Prominent role models: High-profile female politicians and the emergence of women as candidates for public office', *American Journal of Political Science* **62**(2), 369–381.

Laver, M., Benoit, K. & Garry, J. (2003), 'Extracting policy positions from political texts using words as data', *American political science review* **97**(2), 311–331.

Lawless, J. L. (2004), 'Politics of presence? congresswomen and symbolic representation', *Political research quarterly* **57**(1), 81–99.

Lee, D. S. (2008), 'Randomized experiments from non-random selection in us house elections', *Journal of Econometrics* **142**(2), 675–697.

Lee, D. S. & Lemieux, T. (2010), 'Regression discontinuity designs in economics', *Journal of economic literature* **48**(2), 281–355.

Lindbeck, A. & Weibull, J. W. (1987), 'Balanced-budget redistribution as the outcome of political competition', *Public choice* **52**(3), 273–297.

Liu, S.-J. S. & Banaszak, L. A. (2017), 'Do government positions held by women matter? a cross-national examination of female ministers' impacts on women's political participation', *Politics & Gender* **13**(1), 132–162.

Lundborg, P. (1998), 'Foreign aid and international support as a gift exchange', *Economics & Politics* **10**(2), 127–142.

Maitra, P. & Rosenblum, D. (2021), 'Upstream effects of female political reservations', *European Journal of Political Economy* p. 102061.

Mariani, M., Marshall, B. W. & Mathews-Schultz, A. L. (2015), 'See hillary clinton, nancy pelosi, and sarah palin run? party, ideology, and the influence of female role models on young women', *Political Research Quarterly* **68**(4), 716–731.

Marshall, J. (2022), 'Can close election regression discontinuity designs identify effects of winning politician characteristics?', *American Journal of Political Science* .

Martin, J. M. & Borrelli, M. (2016), *The gendered executive*, Philadelphia, PA: Temple University Press.

McAllister, I. (2019), 'The gender gap in political knowledge revisited: Australia's julia gillard as a natural experiment', *European Journal of Politics and Gender* **2**(2), 197–220.

McCain, R. A. (2019), *Welfare Economics: An Interpretive History*, Routledge.

McCrary, J. (2008), 'Manipulation of the running variable in the regression discontinuity design: A density test', *Journal of econometrics* **142**(2), 698–714.

Merkley, E. (2020), 'Anti-intellectualism, populism, and motivated resistance to expert consensus', *Public Opinion Quarterly* **84**(1), 24–48.

Milligan, K., Moretti, E. & Oreopoulos, P. (2004), 'Does education improve citizenship? evidence from the united states and the united kingdom', *Journal of public Economics* **88**(9-10), 1667–1695.

Mitra, A. (2020), 'Should you want an educated mayor? evidence from close elections in italy', *Evidence from Close Elections in Italy (November 2020)* .

Mocan, N. H. & Cannonier, C. (2012), Empowering women through education: Evidence from sierra leone, Technical report, National Bureau of Economic Research.

Morey, D. & Lai, B. (2003), Liberalism, realism, and united nations voting: An empirical test of contending theories, department of political science, university of iowa, Technical report, mimeo.

Nifo, A. & Vecchione, G. (2014), 'Do institutions play a role in skilled migration? the case of italy', *Regional Studies* **48**(10), 1628–1649.

Norris, P. (1996), 'Women politicians: transforming westminster?', *Parliamentary Affairs* **49**(1), 89–103.

O'Connell, S. D. (2020), 'Can quotas increase the supply of candidates for higher-level positions? evidence from local government in india', *Review of Economics and Statistics* **102**(1), 65–78.

OECD (2013), *OECD Skills Outlook 2013*.
   URL: *https://www.oecd-ilibrary.org/content/publication/9789264204256-en*

OECD & Canada, S. (2000), *Literacy in the Information Age.*
  URL: *https://www.oecd-ilibrary.org/content/publication/9789264181762-en*

Osborne, M. J. & Slivinski, A. (1996), 'A model of political competition with citizen-candidates', *The Quarterly Journal of Economics* **111**(1), 65–96.

Park, H. & Kyei, P. (2011), 'Literacy gaps by educational attainment: A cross-national analysis', *Social Forces* **89**(3), 879–904.

Persson, T. & Tabellini, G. E. (2002), *Political economics: explaining economic policy*, MIT press.

Peter, S. et al. (1990), 'The fifth discipline', *The Art & Practice of Learning Organization. Doupleday Currence, New York* .

Prindeville, D.-M. (2002), 'A comparative study of native american and hispanic women in grassroots and electoral politics', *Frontiers: A Journal of Women Studies* **23**(1), 67–89.

Rai, K. B. (1980), 'Foreign aid and voting in the un general assembly, 1967—1976', *Journal of Peace Research* **17**(3), 269–277.

Reder, S. (2008), 'The development of literacy and numeracy in adult life', *Tracking Adult Literacy and Numeracy Skills: Findings from longitudinal research, Routledge, New York and London* pp. 59–84.

Rehavi, M. M. (2007), Sex and politics: Do female legislators affect state spending?, Technical report.

Repetto, L. (2018), 'Political budget cycles with informed voters: evidence from italy', *The Economic Journal* **128**(616), 3320–3353.

Rocha, F., Orellano, V. I. F. & Bugarin, K. (2018), 'Elected in a close race: Mayor's characteristics and local public finances', *EconomiA* **19**(2), 149–163.

Rooduijn, M. (2014), 'The nucleus of populism: In search of the lowest common denominator', *Government and Opposition* **49**(4), 573–599.

Rosenthal, C. S. (2000), 'Gender styles in state legislative committees: Raising their voices in resolving conflict', *Women & Politics* **21**(2), 21–45.

Samarakoon, S. & Parinduri, R. A. (2015), 'Does education empower women? evidence from indonesia', *World Development* **66**, 428–442.

Schraeder, P. J., Hook, S. W. & Taylor, B. (1998), 'Clarifying the foreign aid puzzle: A comparison of american, japanese, french, and swedish aid flows', *World politics* pp. 294–323.

Schuknecht, L. (2000), 'Fiscal policy cycles and public expenditure in developing countries', *Public Choice* **102**(1-2), 113–128.

Schwindt-Bayer, L. A. (2010), *Political power and women's representation in Latin America*, Oxford University Press.

Seabright, P. (1996), 'Accountability and decentralisation in government: An incomplete contracts model', *European economic review* **40**(1), 61–89.

Slapin, J. B. & Proksch, S.-O. (2008), 'A scaling model for estimating time-series party positions from texts', *American Journal of Political Science* **52**(3), 705–722.

Stockemer, D. (2011), 'Women's parliamentary representation in africa: the impact of democracy and corruption on the number of female deputies in national parliaments', *Political Studies* **59**(3), 693–712.

Sundström, A. & Wängnerud, L. (2016), 'Corruption as an obstacle to women's political representation: Evidence from local councils in 18 european countries', *Party Politics* **22**(3), 354–369.

Tripp, A. M. (2001), 'Women and democracy: The new political activism in africa', *Journal of democracy* **12**(3), 141–155.

Ulbig, S. G. (2007), 'Gendering municipal government: Female descriptive representation and feelings of political trust', *Social Science Quarterly* **88**(5), 1106–1123.

Ulfelder, J. (2012), 'Democracy/Autocracy Data Set'.
URL: *https://doi.org/10.7910/DVN/M11WFC*

Vergne, C. (2009), 'Democracy, elections and allocation of public expenditures in developing countries', *European Journal of Political Economy* **25**(1), 63–77.

Voeten, E. (2013), Data and analyses of voting in the united nations general assembly, *in* 'Routledge handbook of international organization', Routledge, pp. 80–92.

Wang, T.-Y. (1999), 'Us foreign aid and un voting: An analysis of important issues', *International Studies Quarterly* **43**(1), 199–210.

Whaley, S. E., Sigman, M., Neumann, C., Bwibo, N., Guthrie, D., Weiss, R. E., Alber, S. & Murphy, S. P. (2003), 'The impact of dietary intervention on the cognitive development of kenyan school children', *The Journal of nutrition* **133**(11), 3965S–3971S.

Wittkopf, E. R. (1973), 'Foreign aid and united nations votes: A comparative study', *The American Political Science Review* **67**(3), 868–888.