2019

Essays in applied political economy

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ESSAYS IN APPLIED POLITICAL ECONOMY

by

JUAN DELFIN CONDE CARVAJAL

B.A., University of Vigo, 2011
M.Sc., CEMFI, 2013

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2019

Approved by

First Reader

Johannes Schmieder, Ph.D.
Associate Professor of Economics

Second Reader

Daniele Paserman, Ph.D.
Professor of Economics

Third Reader

Raymond Fisman, Ph.D.
Professor of Economics

To my mother, always.

Acknowledgments

First, I would like to thank my advisor, Prof. Johannes Schmieder for his continuous support throughout my Ph.D, for his patience, and his help. Without his guidance, this project would not be possible. I would also like to thank the rest of my committee, Prof. Daniele Paserman and Prof. Raymond Fisman for their time, motivation and unmeasurable help.

I would like to show my gratitude to my fellow Ph.D candidates at Boston University, and in particular those who participated weekly in the Labor and Political Economy reading groups. You listened to my presentations and projects as if it was the first time you heard about them, every single time. Christoph, Bruno, Phil, Jesse, Yeseul, Kevin, and many others, thank you.

Last, but not least, I am particularly grateful for my friends, my mother and my sister. Christoph, Roisin, Nick, Megan, Mili, Angeles, you made these last six years of my life fly by, and without your support, these thesis would not be.

ESSAYS IN APPLIED POLITICAL ECONOMY

JUAN DELFIN CONDE CARVAJAL

Boston University, Graduate School of Arts and Sciences, 2019

Major Professor: Johannes Schmieder, Ph.D., Associate Professor of Economics

ABSTRACT

The first chapter analyzes the impact of gender quota regulation on women's participation in politics. Gender quotas are the main policy tools used to encourage participation in politics. A natural experiment in Spanish municipal elections is exploited to study the success of such reforms. Gender quotas are found to improve the number of women candidates, but due to strategic reaction from political parties, much fewer women are being elected. Political parties disproportionately allocate women to the lowest possible position while still complying with the law. Parties have a propensity to assign women candidates to positions where they have relatively low chance of being elected. There is also no shift in public policy toward spending preferred by women.

The second chapter presents empirical evidence in support of the Leviathan model of government. In Spain, the number of politicians chosen in local elections depends on the population of the municipality. Using a data set that covers over two decades of municipal elections, I present two main results. First, there is an unusual concentration of municipalities (bunching) with reported populations just above the threshold that increases the number of local representatives. I present compelling evidence that elected officials manipulate population figures in advance of upcoming elections in order to maximize the size of the council. Second, I use machine learning techniques to construct an unbiased

measure of population based on luminosity data and census population figures, and study which municipalities are more likely to misreport based on the quality of the democratic institutions. Based on those measures, I conclude that misreporting is more likely to happen in municipalities with higher turnout and less parties in their council.

The final chapter studies the impact that World War II fatalities had on political preferences during the twentieth century in the United States. We document enlistment and fatalities at the county level and use this variation to study the hypothesis that fatalities permanently shifted U.S. political preferences. In particular, we test whether the proximate casualties theory, which states that voters punish incumbents in the short run after a war, affected United States counties after World War II. We conclude that there is not enough evidence in our analysis to determine that fatalities during World War II significantly impacted long term political preferences.

Contents

List of Tables

List of Figures

List of Abbreviations

GOP	....	Grand Old Party; Republican Party
ICPSR	....	Consortium for Political and Social Research
INE	....	Instituto Nacional de Estadistica
OLS	....	Ordinary Least Squares
RD	....	Regression Discontinuity
US	....	United States

Chapter 1

Gender Quotas and Women's Political Empowerment. Evidence from Spain

1.1 Introduction

Gender quota legislation has been the most common policy tool used to increase women's participation in political activities. The 2014 Atlas of Gender Quotas states that 118 countries (about 60%) in the world imposed some sort of gender quota system in their elections (Dahlerup, Hilal, Kalandadze, and Kandawasvika-Nhundu (2013)). From an economic and policy perspective, it is important to understand the impact of such widespread reforms, in terms of whether they help correct gender imbalances, the extent of their effectiveness and what makes them successful.

The present study considers one particular country where gender quota legislation was introduced as part of a larger effort by the government to establish equality between women and men in many aspects of society, Spain. On March 22nd, 2007, a new law was introduced to help correct gender imbalances in the Spanish economy and government. Article 44 of this law states that all lists of candidates presented to municipal elections must have a balanced gender representation, such that either of the genders must compose at least 40% of the list. Studying the consequences of this reform is relevant for several reasons.

First, extensive literature has been remarkably insightful in studying the role of women in government in developing countries (Duflo (2012)) but work on the impact of gender quotas in developed economies is more scarce. Several studies have focused on a natural experiment in India, where the head position in a few randomly selected villages was reserved for women. This was a perfect scenario for studying the impact of having a female politician heading the the government. Chattopadhyay and Duflo (2004) conclude that policy was enacted to better reflect women's desire in those villages with a female head of council.

A few studies focused on developed countries and investigated a short lived gender reform quota in Italy: De Paola, Scoppa, and Lombardo (2010) find that gender quotas improve female politician's election probabilities, and Weeks and Baldez (2015) and Baltrunaite, Bello, Casarico, and Profeta (2014) exploit identical natural experiment to argue that the quality of politicians remained unchanged after the reform. In the same Italian context, Gagliarducci and Paserman (2011) argue that networks in the political arena, predominantly male dominated, impair women's odds of success, measured in probabilities of reelection. In Sweden, a fre researchers studied gender quotas at the party level, and they found evidence that gender quotas may improve efficiency not just through policy. Besley, Folke, Persson, and Rickne (2017) show that a gender quota reform in fact improved the quality of politicians, by driving low-quality male politicians out of the electoral process.

Second, in this particular application, the policy mandated that parties present a minimum number of women as candidates in their lists, rather than having council seats reserved for women. This type of gender quota legislation the most common one. As of 2014, 60 countries had implemented candidate quotas, whereas 36 had enforced "reserved seats" legislation (Dahlerup et al. (2013)). A third type of common gender quota legislation is voluntary party quotas, where own parties enforce the policy of a minimum number of women in their lists. O'Brien and Rickne (2016) presented a recent study where they analyzed a gender quota reform imposed within the Swedish Social Democratic Party, and found that gender quotas can increase women's political representation.

Third, the Spanish reform lasted longer than other natural experiements, which allowed me to study the treatment effect over time. Morevoer, in some cases, how long the reform has been in place has been shown to be very relevant. Beaman, Chattopadhyay, Duflo, Pande, and Topalova (2009) and Beaman, Duflo, Pande, and Topalova (2012) show that girls' inspirations after gender reform legislation change only after the gender reform has been in effect for two terms, not just one. This means that the short term impact of gender quotas may be different than the long term effect, and this particular natural experiment allows me to distinguish between the two.

This policy reform helps to cleanly identify treatment effects. According to the law, not all public elections are subject to the reform, and there is variation in the timing of implementation. In 2007, when the reform first passed, only municipalities with more than 5,000 people were affected by the law. In 2011, this threshold was reduced to 3,000, except for those municipalities in islands where 5,000 was retained. This legal design provides a perfect natural experiment where the causal effect of the reform can be estimated using a regression discontinuity design. I observe candidates in three types of municipalities: those affected by the reform since 2007, those when the reform was enacted in 2011, and those where the gender quotas have never been in effect. I use this variation to understand how the reform affected women's political representation, how parties reacted strategically to this reform, and whether the reform had an effect on policy choices in terms of public spending.

The results of my analysis show that this reform helped improve women's political representation but it failed to fully empower them. It increased the share of female candidates, both at the 5,000 and 3,000 thresholds by around 7-8% when the reform was first introduced. However, these gains translated in smaller increases in the proportion of women elected, due to two reasons. First, parties reacted by assigning women to the lowest possible position while still in compliance with the gender quotas. Seconds, parties allocate women to positions where they had fewer chances of being elected. Therefore, the strategic behavior of political parties prevents the reform from empowering women. I also investigate whether the reform impacted public spending toward areas preferred by women. I find no evidence that this is the case. Generally, the reform was unsuccessful in creating a permanent impact that increased women's political power. On the contrary, the dynamic effects of the reform are decreasing over time: even when gender quotas had a positive impact, the magnitude of the impact decreased over time, and was mostly negligible by 2015. For instance, the proportion of women elected increases by 5% when the reform is first introduced, but it falls to 2% by 2015.

This paper builds upon the results from (Bagues and Campa, 2017), who studied this same reform and its ability to empower women. The present study however is able to determine why gender quota legislation is not as successful as it is expected. Public perception or voter behavior does not hold women back, but the strategic behavior by political institutions that does not allow women to achieve positions of power. Other papers study political gender dynamics in Spain. Casas-Arce and Saiz (2015) used the same reform to test whether gender quotas negatively affected parties that were most affected by them. The findings suggest that they in fact helped those in the elections. This is particularly relevant to the present study results, because my results indicate that political parties are obstructing women. The study results show that this behavior is inefficient if political parties main goal is to win elections. Campa (2011) uses this reform to conclude that there is no change in public spending. Esteve-Volart and Bagues (2012) focus on Senate elections and show that party dynamics play a big role in women's lack of political power: even when they make the ballot, parties tend to nominate women to lower positions. Besides the political arena, Bagues and Esteve-Volart (2010) study gender discrimination in the selection of judges.

Throughout the sample period, the political position of women in untreated (smaller) municipalities is improving over time. This motivates me to test the hypothesis that the convergence pattern between treated and untreated municipalities could be because of the reform itself. I explore whether geographical spillovers could have caused part of the growth in the untreated municipalities. However, the results suggest the contrary: in 2007, when parties had little time to react to the reform, municipalities that are not affected by the law put less women on their lists if they are surrounded by more "treated" neighboring municipalities (municipalities where the gender quotas were in effect). This is some compelling evidence that, in 2007, political parties may have "poached" some female candidates from nearby municipalities when they experience a shortage in the supply of candidates.

Finally, the study analyzes whether the reform shifts public spending to issues preferred by women. It is important to note that, according to the median voter or the Coasian model, the personal views (or gender) of the politician do not affect equilibrium policy. However, if elected public officials do not perfectly represent the preferences of their constituency, but at least partially their own, and women's preferences differ from those of men, then an imbalance in gender representation will also result in inefficient policies. The literature shows support for both parts of this theory. First, several studies confirm that both personal preferences matter, and that women's preferences differ: Levitt (1996) shows that US senators' personal ideology accounts for half of the weight in their decisions. Lott and Kenny (1999), Miller (2008) and Edlund and Pande (2002) are good examples of the evidence in favor of women having diverse political preferences. The first two show that suffrage coincided with both an increase in state government expenditures and a decrease in infant mortality rate, better representing women's policy preferences. The last investigates how these different preferences are related to the decrease in marriage rates in the United States. Second, Svaleryd (2009) shows that gender quotas in Sweden resulted in a shift in public policy, showing increased spending in childcare and education. However, in the United States Ferreira and Gyourko (2014) find that having a woman as mayor has no impact on policy outcomes, although the authors do find women to perform better as politicians, which is measured by a higher incumbency rate.

The rest of the paper is organized as follows: section 2 explains the institutional setting and reform. Section 3 covers the data sources and the empirical strategy, focusing on its validity. The main study results are presented in section 4. A couple of extensions are presented in section 5, and section 6 concludes.

1.2 Institutional Setting

Spain has been a democratic country since 1975. It is a highly decentralized country, governed through a parliamentary monarchy system. It has three main levels of government: the central government, which would be the equivalent of the United States federal government, the autonomous communities, which control areas of government such as education or health provision, and the local municipalities. There are 16 autonomous communities and about 8,100 municipalities.

Municipal elections are held every four years. All the elections are conducted simultaneously, usually in late April or early May. It is a multiparty system: any political party can submit a list of candidates. The list must be numbered and the number of candidates must be equal to the size of the political council. Council size is an odd number, which is determined solely by the population of the municipality. Figure 2·1 shows this relationship in detail. The first person in each list is the candidate for mayor post for that particular political party. Votes are cast to a political party and not particular candidates, and seats are assigned by the Jefferson method (also known as the D'Hondt method), an approximately proportional mechanism[1]. Consequently, positioning in the list is very important: candidates near the top of the list have a much higher likelihood of being elected than their counterparts farther down the list. Because many parties present candidates lists, it is often the case that one party does not reach at least half plus one of the council size seats, which would grant this party majority, and its first candidate the position of mayor. In such occasions, the choice of mayor is through negotiations between the different parties are represented. In the majority of cases, the first candidate in the strongest member of the coalition will be chosen as the mayor.

The present study analyzes the impact of candidate gender quotas in municipal elections introduced by Law 3/2007 "Ley de Igualdad". Article 44 of the law states that all lists presented for any election must have a balanced gender representation. It defines "balanced gender representation" as each gender composing at least 40% of the list. Note that this policy provides both a minimum (40%) and a maximum (60%) number of female candidates (40%). Importantly, the article does not apply to municipalities with a population of less or equal to 5,000 people in 2007. In the subsequent election cycle, in 2011, the threshold was lowered to include all mainland municipalities with populations greater than 3,000. Municipalities in the Spanish archipelagos, the Canary and Baleares islands, were still subject to the threshold established in 2007. This lends itself to a regression discontinuity design, where the effect of the reform can be estimated at two different points of the population distribution. The next section, explains the data sources, and tests the validity of such design. Another important distortion that the law introduces is that the minimum quota has to be met for each of the five consecutive positions within the candidate list. The study results show that this is relevant and political parties will react strategically.

[1] If interested in the properties of this allocation mechanism, check Balinski and Young (1978).

The law was first enforced on March 22nd of 2007. This was a month and a half before the municipal elections of 2007 were conducted. This implies that parties might have been unable to strategically react to the law when it was first passed, and it raises the interesting issue of whether there are different responses in 2007 and 2011, when parties may have had more time to explore new candidates or strategically create their election lists. The study finds some evidence that this might have been the case.

One of the most attractive features of this reform is that it allows for studying not only the short term impact in 2007 and 2011, but also the medium term dynamic effects that gender quotas had in treated municipalities by 2015. The possible impacts of the reform are graphically shown in Figure 1·1. This graph, which plots time on the horizontal axis versus a theoretical outcome that would capture the intended gain in women's position. The vertical red line represents the time during which the reform takes place. The dashed line shows the counterfactual evolution of the outcome had the reform not taken place. Based on this conterfactual, both the short term impact of the law can be calculated (ΔY_1), as well as the long term effect of gender quotas (ΔY_2).

In the short term, a gender quota may or may not have an impact. If it does, this would be depicted as a sudden increase in the outcomes at the time of the reform, represented by the red vertical line in the diagrams. The present study is in line with the literature, which finds that gender quotas generally have an impact on women's candidacy and election rates. However, how things evolve in the medium or long term is less clear. Even believing that gender quotas have a short term impact, we could later observe three distinct possibilities, each shown in one the panels of the Figure. First, this change in gender ratios and women's political participation is a permanent one: it creates a wedge between treated and untreated municipalities, and this wedge remains constant over time. This would be depicted by *"One Time Permanent Impact"*. If this was the case, the short term impact of the reform would be the same as the long term impact ($\Delta Y_1 = \Delta Y_2$), and dynamics would not play a significant role. Another possibility is that the reform has a short term impact, but a smaller long term effect ($\Delta Y_1 > \Delta Y_2$), as shown by the curve labeled *"Transitory Impact"*. This would occur if women's relative position in the society is strengthening over time, and the reform decelerates the growth of women's outcomes when compared to no reform. In this case, the reform is improving the gender ratios that would have happened either way a little bit earlier in time.

Finally, the best case scenario for the gender reform is shown by the possibility labeled *"Poverty Trap"*, owing to its relationship to the theory widely studied in Development Economics. This theory states that certain situations (such as poverty) are a self reinforced. This could be the case if women had greater presence in the political arena: if there are no women candidates and the voters believe that women are inferior in skills to their male counterparts, parties have no incentive to field female candidates. In that case, voters fail to learn about the true skill of female politicians and parties are not motivated to change their behavior. However, as it is the case with poverty trap, such vicious cycle can be broken with a one-time intervention. If gender quotas add enough women on the lists, and voters learn about the women's true skill, then municipalities may be set on a new equilibrium, in which case, woman's outcomes improve faster than when the reform did not take place. This would result in a longer term effect bigger than the short term impact ($\Delta Y_1 < \Delta Y_2$). The next section describes the data collection method, and the strategy to empirically test these hypothesis.

1.3 Data and Empirical Strategy

Electoral data for each Spanish municipality from 2003 to 2015 were collected. These data are provided by the Ministerio del Interior. The Ministerio del Interior also provided data on the names and position of all candidates in each election year. Combining both of these data, a data set was created with all candidates, their name, gender, political party, position on the list and their election to the council. Later, this data set is merged to public spending data available since 2010. Although data on public spending is available for the previous years, the level of disaggregation is not relevant to the present analysis.

The study duration provides a panel of about 8,000 municipalities observed at four time points: 2003, before the reform takes place; 2007, soon after its enactment; and 2011 and 2015. Names and gender of politicians were not collected by the Ministry prior to 2003. For the purpose of identification, the main feature of this reform is that it does not affect all municipalities simultaneously. In 2007, only municipalities with population greater than 5,000 were affected. In 2011, the threshold was lowered in the mainland to 3,000. This implies that the "treated" municipalities increased in 2011 by all those that reported populations between 3,000 and 5,000. This also means that there may be dynamic effects at play. If we assume that municipalities did not change population brackets before the 2015 election, and most of them did not, some municipalities will have been treated twice, some once, some none at all.

In the four election years of data collected, 77,000 parties presented a list of candidates for municipal elections. Of these, nearly 25,000 occur in municipalities and years where gender quotas were in effect (those are considered treated elections or municipalities), and the rest occur without gender requirements. For the rest of the analysis, the sample is restricted to municipalities with population between 1,000 and 10,000. Several small municipalities have fewer than 1,000 people, but they have been excluded from the analysis for several reasons: first, they are distant from any discontinuity, so they do not play a role in helping to identify the impact of the reform. Therefore, dropping these municipalities prevents descriptive statistics from overrepresenting small municipalities. Second, some of these municipalities have a different political system, not based on party lists, but in direct democracy, so they are not comparable with the rest of the study. With regard to municipalities with population greater than 10,000, similar reasoning is followed. They are distant from either threshold, so they do not play a role in identifying the treatment effect, there are few, and much different from the rest of the sample, since they include the most urban and populated areas of Spain. Therefore, the treatment effect in relatively small municipalities, where social and gender norms might usually differ from those in bigger cities. Once those restrictions are applied, a sample of nearly to 800,000 candidates emerges, with approximately 200,000 candidates per election year. Of those, more than 50,000 are elected every election year.

Table 1.1 provides a summary of descriptive statistics. The table presents the position and evolution over time of women's participation in Spanish municipal elections. The percentage of woman candidates increases from slightly more than 30% in 2003 to 43% in 2015. The percentage of women that are elected follows a similar pattern, raising from 27 to 38%. However the proportion of women who are selected as head of the list remains surprisingly constant, at 4%. This shows that, in general, women's relative position in the political arena is improving over time, even though it does not reach equality. The percentage of women in top three positions ranges from a quarter in 2003 elections to 36% in 2015.

The average municipality in the sample used in the present analysis has a population of over 3,700 in 2003 and 4,200 in 2015. The average council size ranges between 9 and 11 seats. Finally, that turnout is quite high in municipal elections but is decreasing over time. Due to data availability, turnout is defined as total number of cast votes divided by the total population in the municipality. A more accurate definition would use the number of eligible voters as the denominator, but that information is unavailable. Therefore, this measure of turnout can be considered as a lower-bound proxy for actual voter turnout. Descriptive statistics on the overall average yearly spending in three particular categories have been found in the literature to be preferred by women in Spain: pensions, health and unemployment insurance. Owing to the economic and financial crisis in Spain from 2008 to 2014, it is not surprising that public spending trends downwards in the data.

Finally, two variables of interest considered in the present study are whether women are put in “realistic positions" or whether they are “irrelevant candidates". These two variables play a key role in the results. The rationale behind these variables is that parties have a good estimate prior to elections on how many votes, and therefore seats, they will obtain. Therefore, in response to the reform, parties might position women in situations where they are on the list, but have no chance of winning. These variables try to capture the following accurately: do parties react to the reform by allocating women to positions in which they think they will not win? Or do they put them in positions where they have a fair chance of being elected?

To construct these variable, the study focuses on the subsample of observations for which party membership can be linked over time (around 70% of all my data). For that subsample, a candidate is defined to to be in a “realistic position" if their position for the upcoming election is one that was elected previously. Therefore, the number of seats won in the previous election is taken as a proxy for parties' expectation of the number of seats they think they will win. This is a simple and straightforward estimation of the prior. A more convoluted expectation function could be created, but this one is tractable, and it has good predicting power (65% R^2).

$$\text{E(Winner)}_{ijpt} = \mathbb{1}\left\{\text{Position}_{ijpt} \leq \text{E(Seats)}_{jpt}\right\}$$

Those who are not put in “realistic positions" are termed as “irrelevant candidates" in this study. This variable is defined for candidate i, in municipality j, from political party p in election year t. The percentage of women is higher among irrelevant candidates than realistic positions, but it seems the gap is smaller in 2015 than it was in 2007.

Given the full names and party associations of candidates, they can be linked over time. Based on those links, the proportion of women who are “returning candidates" are listed, defined as having appeared as a candidate in any previous election, and “new candidates" as women who are new to the pool of candidates. Since the study starting year is 2003, this variable considers all candidates as new in that year. For subsequent elections, around four fifths of all women are new candidates, and one fifth has ran for office before.

The present study uses a regression discontinuity design as its identification strategy, based on the two population thresholds that determine whether the gender quotas are binding in 2007 and 2011, for 5,000 and 3,0000 people respectively. The main identifying assumption of this empirical method is that whether a municipality is found at the just at the right or left of the cutoff is unrelated to the treatment and outcomes. If the population is randomly assigned, then this assumption is true. This assumption can be visually tested by showing a density plot of the assignment variable (population). The McCrary test is a more rigorous statistical test that checks for smooth density at the cutoff. The p-values for this test are 0.012 for the 5,000 threshold and 0.303 for the 3,000 threshold. This implies that this test rejects the null of smooth density only for the 5,000 threshold, and not for the 3,000 one. These results are visually confirmed in Figure 1·3a and 1·3b. In Figure 1·3a, some bunching can be observed right after the threshold, which disappears in the 1·3b. This is slightly concerning, but in a previous study Carvajal (2017) explores this behavior and concludes that the bunching is due to the desire of these municipalities for a bigger council size. Recall that council size varies at the 5,000 threshold but not at the 3,000. See Carvajal (2017) for more details.

Bunching would invalidate the regression discontinuity design if municipalities were choosing their population level in a way that was correlated to their gender dynamics. This can be at least partially tested. If municipalities are bunching, and this relates to gender politics, we should observe differences before the reform takes place between municipalities that bunch and those that do not. There always is a possibility of selection in unobservable characteristics, however, no differences in observable variables will alleviate the concerns. The test results are shown in the appendix, in the last four columns of Table **??**. These results focus on a subsample around the discontinuities (between 2,800 and 3,200 for the 3,000 threshold, and 4,600 and 5,400 for the 5,000 one) before the quotas were introduced. A large bandwidth for the 5,000 cutoff is needed to include more municipalities. The test explores whether there are significant differences in several variables in the municipalities above and below the threshold in 2003. If there were, this implies that municipalities are not randomly assigned around the threshold and the non-random assignment is related to the gender ratios in these municipalities. The averages below and above the cutoff are similar for all nonmechanical variables at both the thresholds (obviously population and total number of votes are higher above the threshold). In particular, the turnout, percentage of female candidates, proportion of lists over quota and the number of women in top three positions are almost identical at the 5,000 threshold, where there is bunching. More female politicians are elected above that threshold. The Results section will show the regression discontinuity estimates for the year 2003 as well. These estimates will work as Placebo tests: directly testing for smoothness of the outcome variables around the cutoff pre-reform. The results are encouraging in support of the identification strategy adopted in the study.

A subsequent section of the paper explores the idea that the reform affected untreated municipalities. If gender norms are changing in municipalities subject to the reform, it could be the case that neighboring municipalities are mirroring this change. To test this hypothesis, the data set is geographically located using an algorithm that searches the name of the municipality on Google Maps and retrieves its coordinates. With these coordinates, the data are merged to a shape-file that contains the geographical information of all municipalities in Spain. This procedure provides information about which municipalities share a geographical border with each other, and this information is used to test for the presence of spillovers.

1.4 Results

1.4.1 Candidacy Rates

This section presents the main study results. The estimation strategy follows a sharp regression discontinuity, where the bandwidth and robust standard errors are calculated according to the procedure outlined by (Calonico, Cattaneo, and Titiunik, 2014). In this particular application, there are two cutoffs, and four election years. In 2003, the cutoffs did not imply a change in treatment status. In other words, there are two Placebo tests in this year. The tables report the Placebo test at the 5,000 threshold for 2003, because this is the first cutoff to be affected by the reform. No statistically significant effect is found on any outcome at the 3,000 threshold in 2003. Another natural Placebo test reported in the tables is the coefficient at the 3,000 threshold in 2007, because these municipalities were not affected by the reform until 2011.

The RD estimates at the 5,000 threshold in 2007 and at the 3,000 threshold on 2011 and 2015 showcase the main effects of the reform. In these cases, political parties in municipalities to the right of the threshold are affected by the reform, and those to the left are not. The evolution of these estimates over time is relevant, since it reveals the dynamic impact of the reform.

Finally, the estimates at the 5,000 cutoff in 2011 and 2015 are of interest. In these two cases, both municipalities to the right and to the left of the threshold are subject to gender quotas. However, municipalities to the right have been subject to gender quotas for a longer duration: in 2007, larger municipalities had already been exposed to the reform, whereas those with population lower than 5,000 experienced this reform for the first time.

In 2011, larger municipalities have been exposed twice, and smaller ones only once. If these estimates were significant, that would support the hypothesis that the number of times a municipality is treated has a significant impact on gender outcomes.

The first stage impact of the reform is shown in Figure 1·5, and in Table 1.2. Figure 1·5 is a scatter plot of the dependent variable in bins of the forcing variable, population. This means the dependent variable is averaged for each bin or bracket of population. The size of the bin is set to 100. Each column in Table 1.2 corresponds to a different election year. The expected percentage of women candidates at the 3,000 population threshold for each election year is reported. Even untreated municipalities showed an increasing trend in the number of women candidates: in 2003 only 31% of all candidates were women, whereas by 2015, the number rose to 41%. The Placebo tests are reassuring. Small and statistically insignificant coefficients for both reported Placebos: the 2003 estimate at the 5,000 cutoff and the 2007 estimate at the 3,000 cutofff. In 2007, there was a significant increase at the 5,000 threshold, where the proportion of women candidates increased by 8%. This was the direct impact of the reform in 2007. These coefficients confirm the results depicted in the first two panels of Figure 1·5.

In 2011, when the municipalities above 3,000 were first subject to the reform, there was an increase in the proportion of women of 7.8%, which is similar to that observed four years earlier at the 5,000 cutoff. The coefficient for the 5,000 threshold in 2011 is zero. This implies that after being treated in 2007, the municipalities above 5,000 did not further increased their share of female candidates. In fact, the smaller treated municipalities catch up to them by 2011. Finally, column four reports the 2015 coefficients. First, there is no difference between municipalities above and below the 5,000 threshold. This fact, combined with the near zero effect for the same coefficient in 2011, reinforces the theory that the length of treatment had no impact on the proportion of female candidates. Second, the coefficient for the 3,000 threshold is still highly significant. Nevertheless, it is smaller than its 2011 counterpart: the effect has decreased from 7.8% to 6%. There appears to be some convergence from control municipalities, those with population below 3,000. These insights can be visually confirmed in the last two panels of Figure 1·5.

In general, it seems that the reform fairly successful in the short run: it increased the share of female candidates by 7-8 percentage points right after being introduced, but that effect diminishes four years after the reform. In addition, no systematic differences appear between those being treated once or twice.

However, this overall increase in candidacy rates hides heterogeneity across positions: the increase in the proportion of women candidates is not constant across all positions in the list. Figure 1·6 shows the proportion of woman by position in each year and population bracket. In 2003, there is a smooth distribution among positions for all types of municipalities, although women are relatively unlikely to be in one of the top two positions. In 2007, this was still true for municipalities that were not subject to gender quotas. However, a clear pattern is observed for the proportion of woman candidates in municipalities with population higher than 5,000, where gender quotas are introduced: there is significant increase in the proportion of women candidates in positions 4 and 5 relative to 6 and 7. The explanation to this pattern is obvious: since the reform forces parties to have 40% of women for each five positions, parties disproportionately place women in the lowest position while still complying with the law. This pattern can be observed again with the newly treated municipalities in 2011. However, by 2015, at least for municipalities with population above 5000 (those that have been treated twice already), this pattern is less clear. This could be due to women moving up the list as parties learn about their true ability.

It is evident parties are reacting strategically to the reform. Therefore, it is possible that parties disproportionately allocating female candidates to positions where they have no realistic shot of winning. To test this hypothesis, candidates are categorized into "realistic positions" and “irrelevant candidates". The impact that the reform these two types of candidates is very different. Table 1.3, and Figures 1·9 and 1·10 show the results. Once again, the RD estimates are significant in 2007 at the 5,000 threshold and in 2011 at the 3,000 threshold. However, the table shows that the estimate for the increase in the proportion of women candidates is twice as big for women in irrelevant positions. This implies that parties are strategically allocating women disproportionately more to positions where the possibility of being elected is remote. This disparity in the size of the effect was still present in 2015, so it did not seem to dissipate over time.

Finally, how the probability that a woman heads the list changes after the reform is implemented is explored. This is particularly relevant position, since being first on the list that the individual is the mayoral candidate for that particular political party. Therefore, if the explicit goal of the reform was to empower women, this is one dimension where policy makers would like to see an effect. The results are shown in Table 1.4. While the reform did have a small impact in 2007 (a 4% increase), the effect decreased by half in 2011 (2%) and it was completely insignificant by 2015.

In conclusion, while this reform helped increase women's candidacy rates, parties reacted strategically by placing them both in the lower position possible such that they would still comply with the law, and in positions where they were not expected to win. In addition, the reform failed in increasing the number of female mayoral candidates in the medium run.

1.4.2 Election Rates

This section attempts to understand whether this reform is successful in placing women in office, not only on the ballot. Table 1.4 shows the regression results for the percentage of the elected women candidates. These results are shown in Figure 1·8.

The main inferences are as follows: first, the percentage of women in office is increasing over time. From 28% in 2003 to 40% by 2015. The reform was successful in the short run in increasing the number of female candidates and placing some them in council. The percentage of women elected increased by 5.6% in 2007 at the 5,000 cutoff and by 4.2% in 2011. The coefficient for the 5,000 cutoff is not statistically significant in 2011. which implies that the newly treated municipalities between 3,000 and 5,000 completely catch up to those above 5000. This is evidence against the reform placing municipalities in a higher growth pattern. Finally, the coefficient for the 3,000 threshold in 2015 is insignificant. This supports the hypothesis of convergence of smaller municipalities over time, and a small impact of the reform 8 years after its introduction.

It is not surprising, giving the findings in the previous section, that the overall size of the effect is lower in election than candidacy rates: if women are not placed in positions where they are expected to win, they are being elected less often. Figure 1·7 displays the election rates for women for each position. As expected, the same pattern is observed as in the candidacy rates: the proportion of elected women is higher at positions 4 and 5 with respect to positions 5 and 6. The results are quite noisy at very low positions in the list, since the number of observations of elected candidates whose position was lower than 6 or 7 is very small, due to low probability of being elected.

Therefore, the reform fails to significantly increase the number of women in positions of power: the increase in probability of election is quite small and decreasing over time, and even when more women are being elected, they are doing so in the fourth or fifth position of the ballot. Party's strategic behavior holds women back. Convergence patterns shown by smaller municipalities also seem to reduce the short term effects of the gender quotas. A later subsection I explore hypothesis by which the reform itself may have been the cause of the growth in the untreated municipalities.

1.4.3 Impact on Spending

This section of the paper discusses whether this reform shifted public spending toward areas preferred by women. Even if the reform had a small effect on election rates of women, it does not preclude a shift in policy. Women's bargaining power within the party may have increased, leading to a shift in public spending, even if the representative in power is a man.

The sudy by (Carrillo Barroso and Tamayo Sáez, 2011) provides relevant information to test this hypothesis. These authors use data from the Role of Government survey, by International Social Survey Programme to study what shapes preferences for public spending. In this survey, the subjects are asked, for different categories of government spending, whether they would like to see more or less government spending. In particular, they study whether women have different public spending preferences.During 1985-2009, women in Spain had consistently higher preferences for spending in three areas: unemployment insurance, pensions and health.

Starting in 2006, public spending by local municipalities was published based on its function, including the three aforementioned areas. This aids in testing whether, once the reform was introduced, public spending on unemployment insurance, pensions or health increased. Not all municipalities report all types of spending through all years. Therefore, the subsample is an unbalanced panel of candidates in these municipalities. In order to retain as much statistical power as possible, spending is calculated as the average yearly spending during the legislature following the municipal election (in logs). If the reform had an impact on public spending in these areas, we would expect to see positive effects on the coefficients of candidates in municipalities where gender quotas were in effect.

The results are shown in Table 1.5. No statistically significant effect of the reform on the type of government spending is observed in any component of public spending. This is consistent with the interpretation that the reform was successful exclusively on putting women in the ballot, but not placing them in influential positions. It does not support the theory that women's bargaining power increased due to the reform.

1.5 Extensions

1.5.1 Spillovers

So far, it appears this policy was only successful in the short run and women's empowerment was hindered by party's strategic behavior. However, the long term impact of gender quotas may have been understated if gender quotas also affected parties and candidates in municipalities where they were not in effect: it is possible that the imposition of gender quotas caused a shift in gender norms in untreated municipalities.

If municipalities under the population threshold changed their gender ratios by being exposed to nearby gender quotas, it would help explain the overall growth path in women's outcomes, and imply that the overall improvement was partially due to the new law. In case of spillovers, political parties introduce more women in their political lists to mirror nearby municipalities. If that is the case, the regression discontinuity estimates would understate the impact of the reform, and the conclusion that the reform was not very successful, or that it did not improve women's position in the long run would be wrong. The reform would lead, at least partially, untreated municipalities to catch up to those affected by gender quotas. This hypothesis is tested next.

Testing for spillovers is not easy, but the strategy followed here sheds some light on whether the reform affected neighboring municipalities under plausible assumptions. Through an algorithm and Google Maps coordinates, the data set is matched to its geographical location, and we are able to determine which municipalities share common borders. Based on this information, the number of bordering treated municipalities is determined for each municipality with population over 250.

Figure 1·4 shows the average number of bordering neighbors that are subject to gender quotas for each municipality in the sample in 207 and 2011. In 2003 the reform has not yet taken place. In 2007, the mode is zero, since the reform only affects counties with more than 5,000 people, and these are uncommon. However, the second most common outcome is to have one neighboring county that is subject to the reform, and it is almost as common as having none. In 2011, the distribution of number of treated neighbors shifts to the right, because the threshold has been decreased to 3,000, and generally more municipalities are treated. However, it remains true for most counties have none or one neighbor who is subject to gender quotas.

Based on this, measures are constructed to capture the extent to which the neighboring municipalities are treated: the proportion of neighboring municipalities subject to gender quotas. Given the independent variable, the following panel data regression is run:

$$\% \text{ Women Candidates}_{it} = \beta_0 + \beta_{1t}\mathbb{1}\left\{\% \text{ Treated Neighbors}\right\} x\text{Year} + \beta_3 Pop_{it} + \delta_i + \gamma_t + \varepsilon_{it}$$

The study focuses on the subsample of untreated municipalities, to test whether their growth is due to their proximity to places where the reform took place. The parameter of interest is β_{1t}, whether the percentage of women candidates in untreated municipalities changes differently when these places are in geographical contact with municipalities where the gender quotas are in effect for each election year. The dependent variable is the proportion of female candidates.

We should be concerned with endogeneity in this regression. It could be that geographical areas where neighbors are treated more often are different along some unobserved characteristics from those where share of treated neighbors is smaller. One could imagine that those areas with high share of treated neighbors are on average bigger, more economically developed and have different attitudes towards women. To deal with this concern: controls directly for population and province fixed effects are included. There are 52 provinces in Spain and this should help control for location specific trends. Moreover, the temporal structure of the data is used, and municipality fixed effects are included. Therefore, identification of β_{1t} comes from municipalities that switch their independent variable: either they change from having no treated neighbors to some treated neighbors, or their share of treated neighbors changes over time.

If in either of these regressions $\beta_{1t} > 0$, this would be compelling evidence of spillovers: when a municipality has more neighbors affected by gender quotas, its own gender ratio in the ballot increases. Note that, however, it would also be plausible to find that $beta_1 < 0$. If there is competition for female candidates, and they are a scarce resource, it would be possible that treated municipalities would "poach" these candidates from untreated municipalities. These would lead to a decrease in gender ratios in the lists presented in elections where gender quotas are not binding.

The results from these regressions are shown in Table 1.6. For this table, the percentage of women candidates or the increase in percentage of female candidates from 2003 to each election year are regressed on a measure of how often your neighbors are subject to gender quotas. Columns one and three do not include municipality fixed effects, but columns two and four (the preferred specification) do. In Table 1.6, the independent variable is the percentage of treated neighbors. Here, there is no positive statistically significant impact of having more treated neighbors on the percentage of women candidates a non treated municipality. However, the coefficient in 2007 is significant but negative. This is interesting, and it could be explained by the "poaching" theory taking place in 2007 but not later on, when political parties had more time to react and recruit new female candidates.

Finally, there is no evidence that the convergence of untreated municipalities was due to the reform itself. If anything, the reform may have just moved women candidates from municipalities non subject to gender quotas to those that were.

1.5.2 Effects by 2003 Complier Status

We learn that party behavior is the main driver behind the lack of success of the gender quota legislation. One fair is whether this effect is homogeneous across all parties, or certain parties, more "woman friendly" behave differently. It is difficult to measure "woman friendliness", particularly when the sample of political parties is so big and heterogeneous: recall thousands of different parties run every four years.

One way to try to answer this is by looking at how parties behaved in the election before the reform was enacted, in 2003. In particular, we can look at how many women were allocated as candidates: were these parties already in compliance with the 40% quota? If they were, we can follow them over time, and see how they react once the reform is enacted, and compare their behavior to those parties where the proportion of women was below the minimum imposed by the law four years later. These are referred as 2003 compliers and 2003 non compliers respectively.

Note, in the process of dividing candidates between compliers and non compliers, the sample must be restricted to candidates belonging to political parties that can be linked over time. This reduces the size of the sample, particularly because most of the observations of candidates from small parties are excluded. Therefore, when analyzing the results of this exercise, it is necessary to consider that we are looking at a particular subsample of candidates.

With that caveat in mind, I explore whether there are significant differences in how these two type of parties reacted to the reform. First, Panel A of Table 1.8, shows that the overall percentage of women candidates increased for both types of parties. This is not trivial, since compliers need not have an effect, but they do: in 2007, they increase the percentage of woman candidates by 3%, and by 5% in 2011, but there is no statistical impact in 2015. However, as expected, the effect is much larger for non compliers (10% in 2007, 9% in 2011, 8% in 2015). One key difference here is that the effect is permanent for non compliers, and temporary for compliers.

Panel B of Table 1.8 shows how this increase in female candidates translates into elected representatives. However, this panel shows that the despite the fact that women increased their numbers more in non complier parties, the effect on election rates was, at least initially, larger for women in complier parties. However, this was only a temporary increase, and by 2015, there was no longer an impact on election rates in complier parties, whereas there was a 5% effect on non compliers.

The overall evolution of these two groups can be seen in Figure 1·11 and Figure 1·12. Both figures arrive at a similar conclusion: the percentage of female candidates, or elected representatives in municipalities that never had access to the reform displays a gap between those in complier and non complier parties. We observe almost full instantaneous convergence in terms of candidacy rates once the reform is introduced: in 2011 for candidates in municipalities with population between 3,000 and 5,000, and in 2007 for bigger municipalities. In terms of election rates though, while there is a pattern of convergence, there is still a gap between compliers and non compliers non compliers. The gap is smaller for municipalities subject to the reform.

Panels C and D of Table 1.8 consider the type of positions women were put in: positions where they were expected to win or not. For 2003 compliers, the reform did not have any impact on putting women into realistic positions, it had a momentary effect of allocating them into irrelevant positions. For non compliers, the same pattern is observed as for the overall population: an increase in both types of positions, but a disproportionate increase in the proportion of women put into irrelevant positions, as opposed to positions where the party expected to the candidate to be elected.

1.6 Conclusion

The present study explores a reform in which a gender floor was introduced in Spanish municipal elections. This reform states that every list presented to municipal elections starting in 2007 must have a balanced gender representation. It defines a balanced gender representation as each gender composing at least 40% of the list. However, not all municipalities are subject to these changes: in 2007 the gender quotas are only imposed in municipalities with population greater than 5,000. In 2011, the threshold was lowered to 3,000 for all mainland municipalities. This is used to identify the causal impact of the reform at two points of the population distribution, and to explore both the short term and long term impact of gender quotas.

Therefore, a data set that merges election results from four election cycles (2003, 2007, 2011 and 2015) with individual candidate data for those years was collected. With this data set, a regression discontinuity design was used to uncover the causal impact of the reform on women's political position in Spain. This study found that the reform did increase women's political participation, measured as the percentage of female candidates. However, this increase did not empower women. Parties allocated female candidates to the lowest possible positions in their list such that they would still comply with the law. Moreover, parties disproportionately allocated women to positions where the candidate was not expected to win in the upcoming election. Because of political parties' behavior, the reform had quite a small impact in the proportion of women that were elected, and no effect on the probability of a woman heading the list. Ultimately, the reform did not shift public spending toward areas preferred by women.

The positive impacts associated with gender quotas legislation are stronger the year that these quotas become binding. However, they tend to diminish after two election cycles and are smaller and sometimes not significant by 2015. This is partly due to the convergence of untreated municipalities, whose outcomes are improving over time, even in the absence of the reform. Geographical data were used in order to test the hypothesis that it was the reform itself that affected the growth and convergence of untreated municipalities. If this were the case, the reform would have been more successful than what it appeared to be. A measure of how many neighbors are treated for each municipality was constructed. It was found that the proportion of women outcomes did not increase when municipalities were surrounded by more treated neighbors. If anything, in 2007, when parties had less time to react to the reform, the number of women candidates decreases in untreated municipalities. This could be due to treated municipalities "borrowing" female candidates from adjacent counties where the reform is not in place.

Tables and Figures

Table 1.1: Descriptive Statistics Women's Political Representation

	2003	2007	2011	2015
% Women Candidates	0.33	0.40	0.42	0.44
% Women Elected	0.14	0.15	0.15	0.15
P(Woman |Position = 1)	0.04	0.04	0.04	0.04
P(Woman |Realistic Position)		0.33	0.35	0.38
P(Woman |Irrelevant Candidates)		0.41	0.43	0.44
Returning Candidate	0	0.20	0.21	0.21
New Candidate	1	0.80	0.79	0.79
P(Woman |Position <4)	0.26	0.31	0.33	0.36
Position	7.06	7.22	7.46	7.62
Number of Candidates	205, 491	227, 261	235, 456	239, 390
Average Yearly Spending Pensions (log)		8.60	8.67	7.55
Average Yearly Spending Health (log)		8.92	8.65	7.54
Average Yearly Spending UI (log)		8.63	8.43	7.32
Population	3, 707	3, 942	4, 138	4, 190
Council Size	9.83	9.96	10.07	10.08
Turnout	0.66	0.64	0.61	0.58
Number Municipalities	5, 418	5, 368	5, 330	5, 245

Note: The level of observation is the individual candidate. All rows denote averages. The sample is composed by candidates in municipalities with population between 1,000 and 10,000. Realistic positions are defined as positions in the list of the party that obtained a seat in the previous election. Irrelevant positions are defined as positions in the list of the party that did not obtain a seat in the previous election.

Table 1.2: RD Estimates of the Impact of the Reform on the Proportion of Female Candidates

	Dependent variable:			
	% Female Candidates			
	2003	2007	2011	2015
	(1)	(2)	(3)	(4)
Placebo Estimate. No Reform	-0.009	-0.008		
at Either Side of Threshold	(0.011)	(0.008)		
Cutoff	5000	3000		
Main Effect. Reform		0.080^{***}	0.078^{***}	0.060^{***}
only on Right Side of Threshold		(0.011)	(0.009)	(0.009)
Cutoff		5000	3000	3000
Dynamic Impact. Reform Intensity			0.008	0.008
higher on Right Side of Threshold			(0.012)	(0.012)
Cutoff			5000	5000
Population > 3000 Treated	No	No	Yes	Yes
Population > 5000 Treated	No	Yes	Yes	Yes
$E(Y \mid Pop = 3000)$	0.31	0.37	0.39	0.41
Observations	93,736	101,133	99,323	96,630

Note: The sample is composed by all candidates in municipalities with population between 1,000 and 10,000. Displayed estimates are local polynomial Regression Discontinuity (RD) point estimators with a triangular kernel. Standard errors are calculated using near neighbor technique. Bandwidth is set to 2000 in each side. The dependent variable is the number of female candidates in all lists divided by the total number of candidates. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

Table 1.3: RD Estimates Women Realistic vs Irrelevant Positions

	Dependent variable:					
	% Female Candidates					
	Realistic Positions			Irrelevant Positions		
	2007	2011	2015	2007	2011	2015
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Estimate. No Reform	−0.024			−0.001		
at Either Side of Threshold	(0.019)			(0.013)		
Cutoff	3000			3000		
Main Effect. Reform	0.056^{**}	0.053^{***}	0.032	0.092^{***}	0.101^{***}	0.081^{***}
only on Right Side of Threshold	(0.028)	(0.020)	(0.021)	(0.018)	(0.013)	(0.015)
Cutoff	5000	3000	3000	5000	3000	3000
Dynamic Impact. Reform Intensity		0.009	0.050^{*}		−0.002	−0.007
higher on Right Side of Threshold		(0.028)	(0.028)		(0.018)	(0.019)
Cutoff		5000	5000		5000	5000
Population > 3000 Treated	No	Yes	Yes	No	Yes	Yes
Population > 5000 Treated	Yes	Yes	Yes	Yes	Yes	Yes
$E(Y\|3000)$	0.42	0.4	0.43	0.31	0.35	0.39
Observations	19,205	19,658	18,583	41,425	43,238	36,140

Note: The sample is composed by all candidates in municipalities with population between 1,000 and 10,000. Displayed estimates are local polynomial Regression Discontinuity (RD) point estimators with a triangular kernel. Standard errors are calculated using near neighbor technique. Bandwidth is set to 2000 in each side. The dependent variable is the number of female candidates in all lists divided by the total number of candidates. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

Table 1.4: RD Estimates on Women as Head of List and Women's Election Rates

Panel A:	*Dependent variable:*			
	P(Woman|Position = 1)			
	2003	2007	2011	2015
	(1)	(2)	(3)	(4)
Placebo Estimate. No Reform	−0.016	0.008		
at Either Side of Threshold	(0.006)	(0.005)		
Cutoff	5000	3000		
Main Effect. Reform		−0.002	−0.003	−0.007
only on Right Side of Threshold		(0.006)	(0.005)	(0.005)
Cutoff		5000	3000	3000
Dynamic Impact. Reform Intensity			−0.001	0.002
higher on Right Side of Threshold			(0.006)	(0.006)
Cutoff			5000	5000
$E(Y\|Pop = 3000)$	0.04	0.03	0.04	0.05
Observations	36,971	46,668	48,877	48,951

Panel B:	*Dependent variable:*			
	P(Woman|Elected)			
	2003	2007	2011	2015
	(1)	(2)	(3)	(4)
Placebo Estimate. No Reform	−0.001	−0.025		
at Either Side of Threshold	(0.020)	(0.016)		
Cutoff	5000	3000		
Main Effect. Reform		0.052**	0.042**	0.021
only on Right Side of Threshold		(0.017)	(0.011)	(0.017)
Cutoff		5000	3000	3000
Dynamic Impact. Reform Intensity			−0.000	0.029
higher on Right Side of Threshold			(0.023)	(0.023)
Cutoff			5000	5000
$E(Y\|Pop = 3000)$	0.29	0.34	0.35	0.4
Observations	25,903	26,776	26,789	26,242

Note: The sample is composed by all candidates in municipalities with population between 1,000 and 10,000. Displayed estimates are local polynomial Regression Discontinuity (RD) point estimators with a triangular kernel. Standard errors are calculated using near neighbor technique. Bandwidth is set to 2000 in each side. Dependent variable is the proportion of female candidates in the first position of a list for the first panel, and the proportion of women among elected candidates in the second. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

Table 1.5: RD Estimates Spending

	Impact on Spending:		
Panel A:	Unemployment Insurance		
	2007	2011	2015
	(1)	(2)	(3)
Placebo Estimate. No Reform	0.238		
at Either Side of Threshold	(0.181)		
Main Effect. Reform	0.116	−0.125	−0..060
only on Right Side of Threshold	(0.300)	(0.245)	(0.490)
Dynamic Impact. Reform Intensity		-0.527^{*}	−0.468
higher on Right Side of Threshold		(0.281)	(0.394)
Observations	83,757	94,735	61,253
Panel B:	Pensions		
Placebo Estimate. No Reform	−0.029		
at Either Side of Threshold	(0.201)		
Main Effect. Reform	−0.493	−0.372	−0.119
only on Right Side of Threshold	(0.333)	(0.249)	(0.741)
Dynamic Impact. Reform Intensity		-1.087^{***}	−0.215
higher on Right Side of Threshold		(0.271)	(0.908)
Observations	80,038	71,033	21,755
Panel C:	Health		
Placebo Estimate. No Reform	0.247		
at Either Side of Threshold	(0.176)		
Main Effect. Reform	−0.064	0.107	−0.232
only on Right Side of Threshold	(0.246)	(0.178)	(0.230)
Dynamic Impact. Reform Intensity		−0.137	0.253
higher on Right Side of Threshold		(0.230)	(0.284)
Observations	86,236	85,914	64,461

Note: Sample is composed by all candidates in municipalities with population between 1,000 and 10,000 for which public spending data is available. Estimates are local polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

Table 1.6: Spillover Effects: Share of Treated Neighbors

	Dependent variable:			
	% Women Candidates		$\Delta_{t,2003}$ % Women Candidates	
	(1)	(2)	(3)	(4)
% Treated Neighbors x 2007	0.019	-0.061^{*}	0.010	-0.062^{*}
	(0.012)	(0.037)	(0.013)	(0.037)
% Treated Neighbors x 2011	0.026	0.020	0.013	0.021
	(0.020)	(0.023)	(0.020)	(0.024)
% Treated Neighbors x 2015	0.024	0.014	0.010	0.012
	(0.015)	(0.017)	(0.014)	(0.017)
Province FE	Yes	Yes	Yes	Yes
Municipality FE	No	Yes	No	Yes
Observations	11,694	11,694	11,357	11,357
Adjusted R^2	0.180	0.529	0.039	0.607

Note: The sample is composed by all municipalities non treated municipalities from 2007 to 2015. All regressions include population controls. Standard errors are clustered at the municipality level. The main independent variable is the share of bordering municipalities that are subject to gender quotas in that election year. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

Table 1.7: RD Estimates on Candidates and Election Rates by Compliance Status

	Dependent variable:					
	% Female Candidates					
	2003 Compliers			2003 Non Compliers		
	2007	2011	2015	2007	2011	2015
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Estimate. No Reform	−0.011			−0.015		
at Either Side of Threshold	(0.023)			(0.012)		
Cutoff	3000			3000		
Main Effect. Reform	0.024	0.059**	0.017	0.107***	0.087***	0.076***
only on Right Side of Threshold	(0.031)	(0.025)	(0.017)	(0.010)	(0.014)	(0.014)
Cutoff	5000	3000	3000	5000	3000	3000
Dynamic Impact. Reform Intensity		0.019	0.007		0.004	0.019
higher on Right Side of Threshold		(0.033)	(0.033)		(0.019)	(0.019)
Cutoff		5000	5000		5000	5000
$E(Y\|3000)$	0.42	0.39	0.47	0.32	0.34	0.39
Observations	4,871	4,327	4,461	14,671	14,165	13,324
	P(Woman|Elected)					
Placebo Estimate. No Reform	−0.002			−0.024		
at Either Side of Threshold	(0.041)			(0.021)		
Cutoff	3000			3000		
Main Effect. Reform	0.067	0.068*	−0.021	0.63**	0.051**	0.040*
only on Right Side of Threshold	(0.042)	(0.044)	(0.031)	(0.017)	(0.024)	(0.024)
Cutoff	5000	3000	3000	5000	3000	3000
Dynamic Impact. Reform Intensity		0.045	0.028		0.007	0.052
higher on Right Side of Threshold		(0.060)	(0.060)		(0.033)	(0.034)
Cutoff		5000	5000		5000	5000
$E(Y\|3000)$	0.42	0.39	0.47	0.32	0.34	0.39
Observations	4,871	4,327	4,461	14,671	14,165	13,324

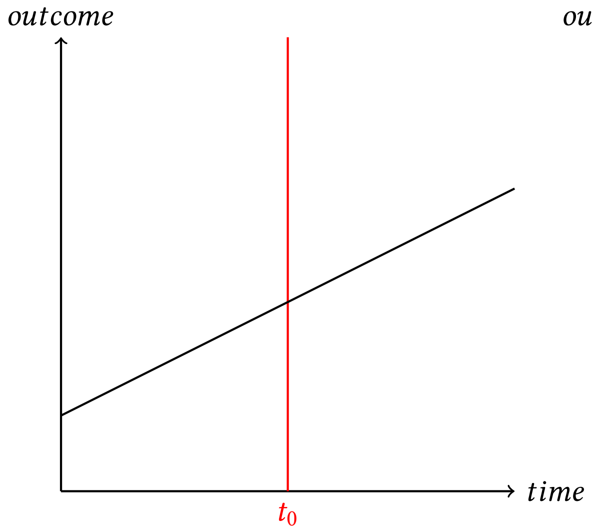
Note: The sample is composed by all candidates in municipalities with population between 1,000 and 10,000 for which political party can be linked over time. Displayed estimates are local polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

Table 1.8: RD Estimates on Type of Position by Compliance Status

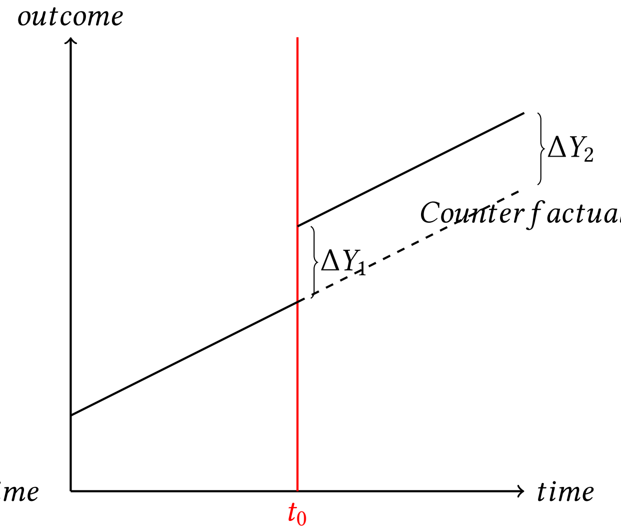
	Dependent variable:					
	P(Woman Realistic Position)					
	2003 Compliers			2003 Non Compliers		
	2007	2011	2015	2007	2011	2015
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Estimate. No Reform	−0.016			−0.017		
at Either Side of Threshold	(0.043)			(0.023)		
Cutoff	3000			3000		
Main Effect. Reform	0.064	0.067	−0.004	0.062*	0.043*	0.032
only on Right Side of Threshold	(0.047)	(0.047)	(0.051)	(0.034)	(0.026)	(0.028)
Cutoff	5000	3000	3000	5000	3000	3000
Dynamic Impact. Reform Intensity		0.021	0.020		−0.003	0.058
higher on Right Side of Threshold		(0.066)	(0.069)		(0.037)	(0.039)
Cutoff		5000	5000		5000	5000
$E(Y\|3000)$	0.42	0.4	0.43	0.31	0.35	0.39
Observations	4,818	4,117	3,913	14,387	12,994	12,856
	P(Woman Irrelevant Position)					
Placebo Estimate. No Reform	−0.017			−0.008		
at Either Side of Threshold	(0.020)			(0.013)		
Cutoff	3000			3000		
Main Effect. Reform	0.026	0.073***	0.031	0.115***	0.102***	0.097***
only on Right Side of Threshold	(0.017)	(0.019)	(0.020)	(0.013)	(0.011)	(0.011)
Cutoff	5000	3000	3000	5000	3000	3000
Dynamic Impact. Reform Intensity		0.001	−0.003		−0.010	−0.007
higher on Right Side of Threshold		(0.014)	(0.014)		(0.008)	(0.009)
Cutoff		5000	5000		5000	5000
$E(Y\|3000)$	0.44	0.43	0.47	0.38	0.39	0.41
Observations	10,169	8,571	7,291	31,256	25,428	22,131

Note: The sample is composed by all candidates in municipalities with population between 1,000 and 10,000 for which political party can be linked over time. Displayed estimates are local polynomial Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

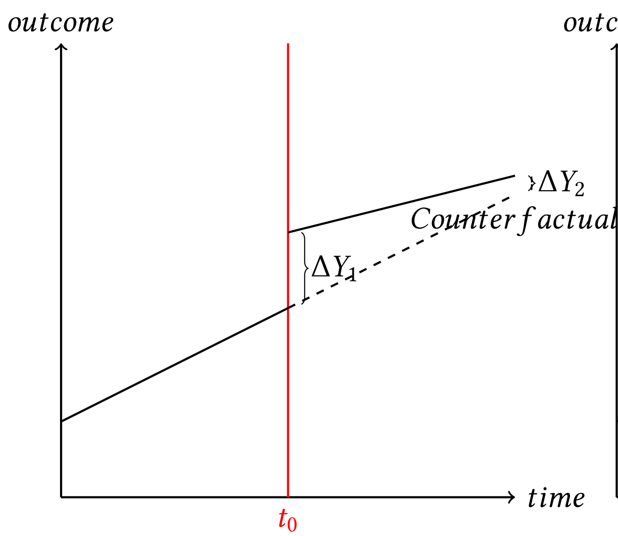
Figure 1·1: Possible Impact of Gender Quota Reform



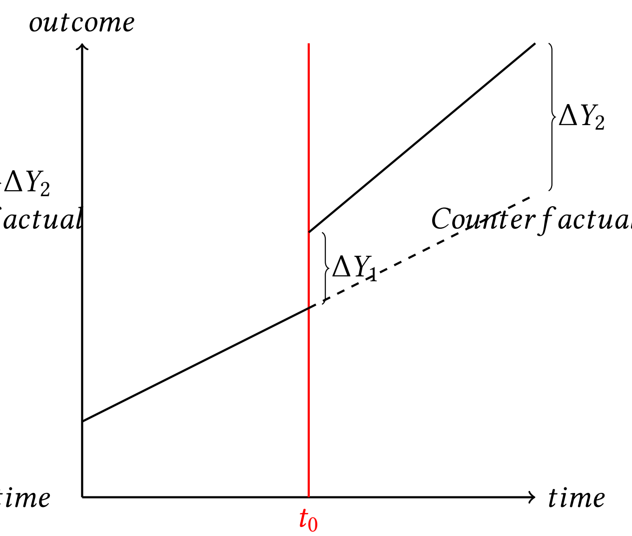
(a) No Impact



(b) One Time Permanent Impact



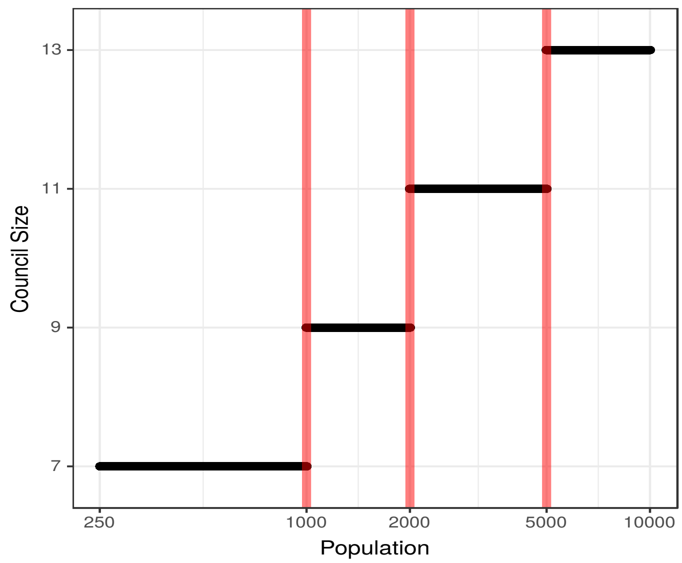
(c) Transitory Effect



(d) Poverty Trap

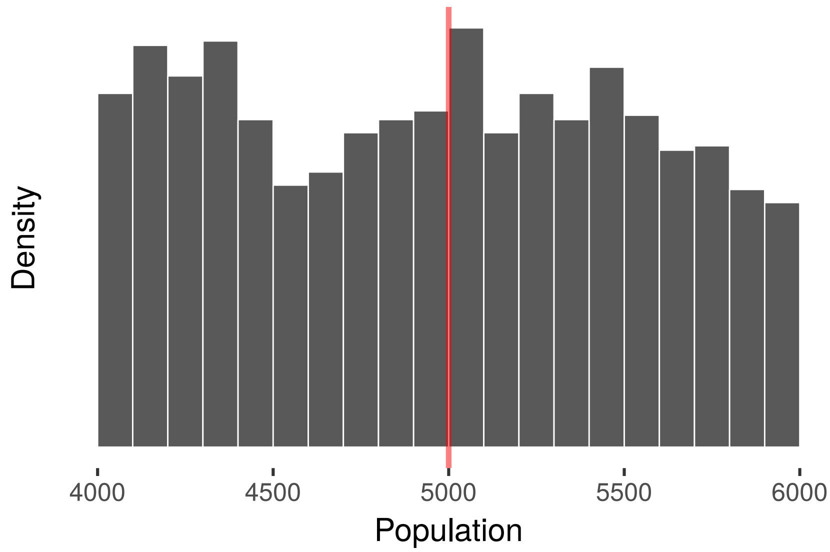
Note: The horizontal axis denotes time, with the vertical red line establishing the point at which a theoretical reform would take place (t_0). The vertical axis represents the outcome for which we are interested in calculating the causal effect of the gender quota reform. ΔY_1 denotes the short term impact of the reform, and ΔY_2 the long term impact.

Figure 1·2: Council Size by Population

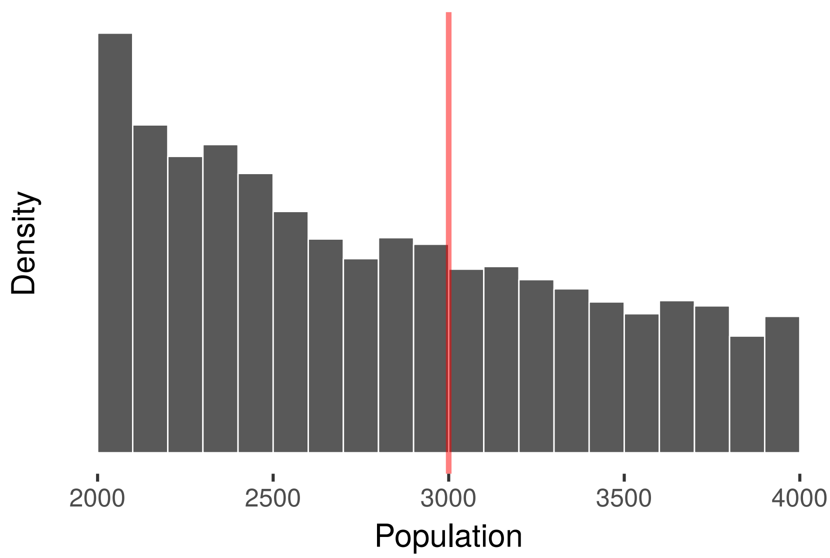


Note: Figure shows the number of politicians elected for each municipality (i.e. council size) depending on their population. The horizontal axis is restricted to municipalities with population higher than 250 and lower than 10,000. The unit of observation is the municipality times election year, for the period 2003-2015.

Figure 1·3: Density of Population Around Threshold



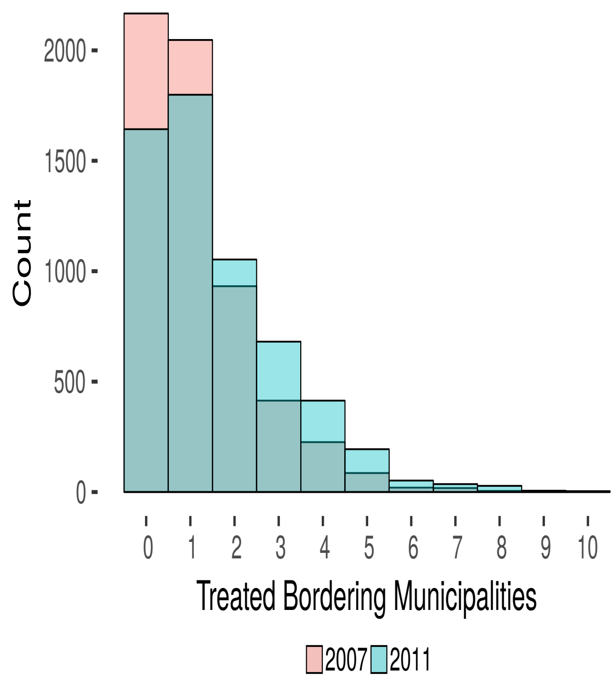
(a) 5,000 Threshold



(b) 3,000 Threshold

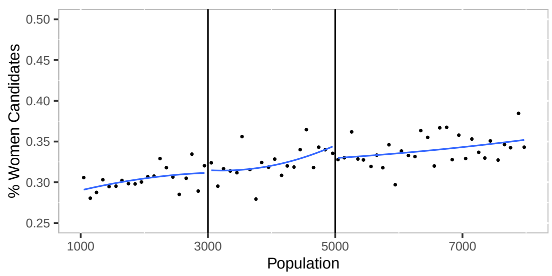
Note: Both figures are histograms showing population density around each of the gender quotas cutoffs. The unit of observation is the municipality times election year. The bin size of both histograms is set to 100. The vertical red lines denote the treatment cutoffs.

Figure 1·4: Number of Treated Neighboring Municipalities

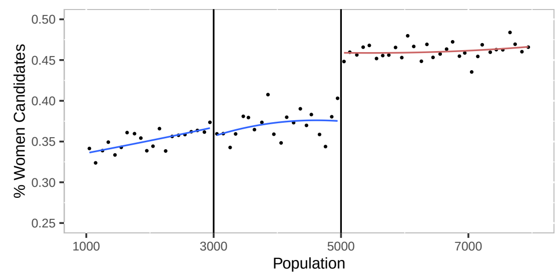


Note: Figure shows, for each municipality in my sample, the number of neighboring municipalities that are subject to gender quotas for both the 2007 and the 2011 election cycle. I don't observe municipalities under 250 population, but those are never treated by law, and are therefore considered as such.

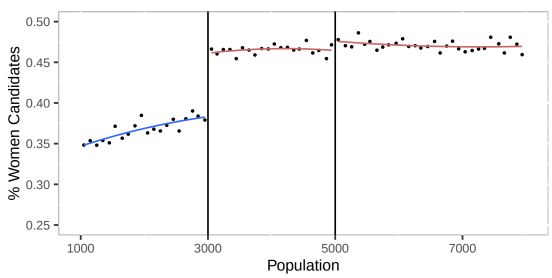
Figure 1·5: Proportion of Women Candidates by Year



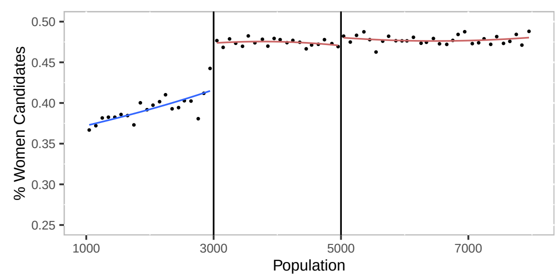
(a) 2003



(b) 2007



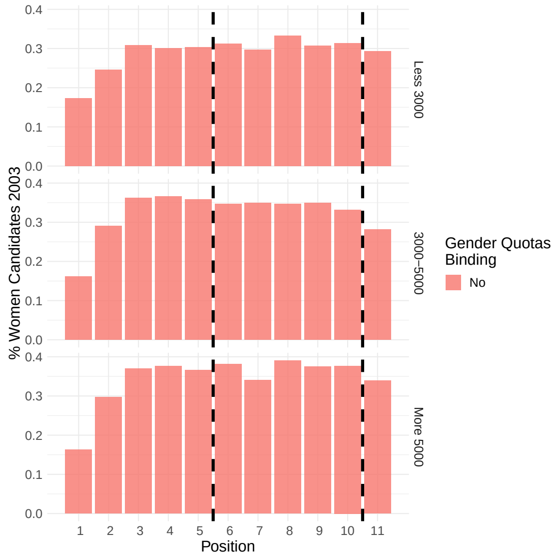
(c) 2011



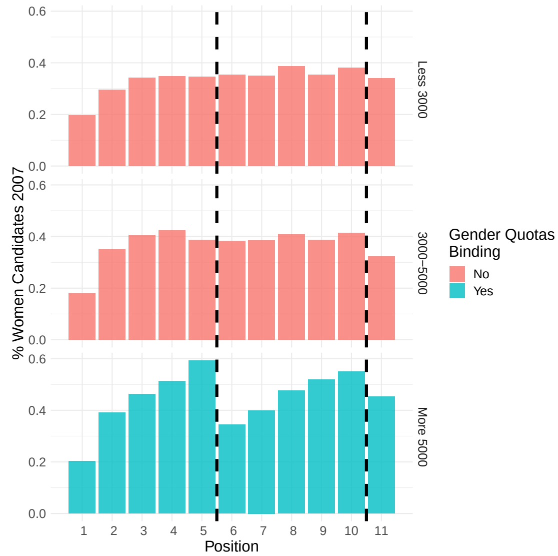
(d) 2015

Note: Unit of observation is the candidate. Sample is composed by all candidates in municipalities with population between 1,000 and 8,000. The vertical lines denote the population cutoffs that determine treatment status. Each dot represents an average of the outcome variable in a bin of size 100 of population. The blue and red lines denote local polynomial regression fitted lines with *loess* method. The color denotes the treatment status for those municipalities in that year: blue for non treated, red when gender quotas are binding.

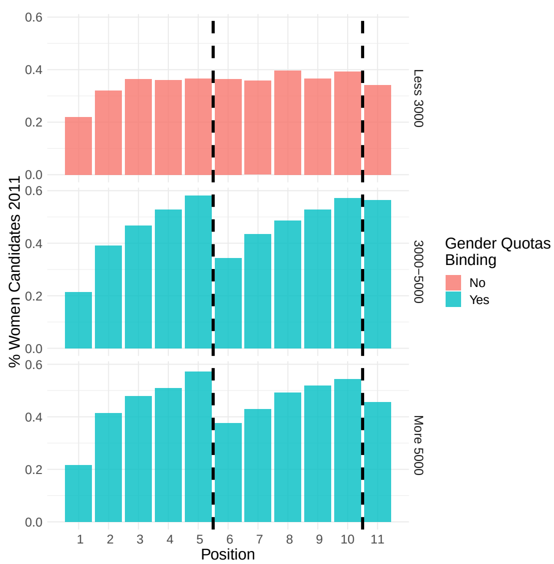
Figure 1·6: Proportion of Women Candidates by Position



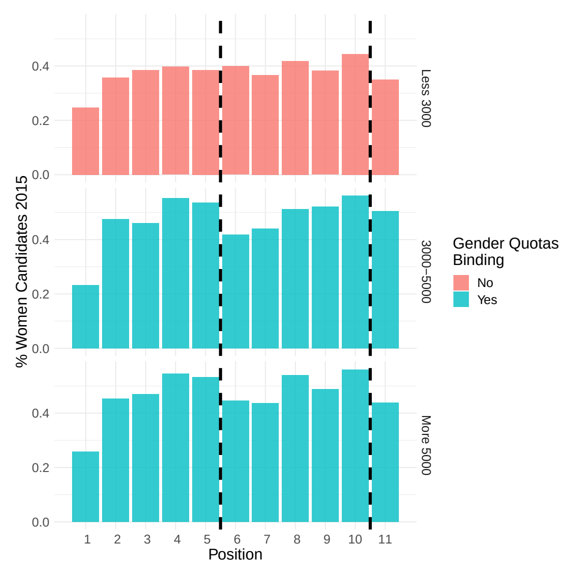
(a) 2003



(b) 2007



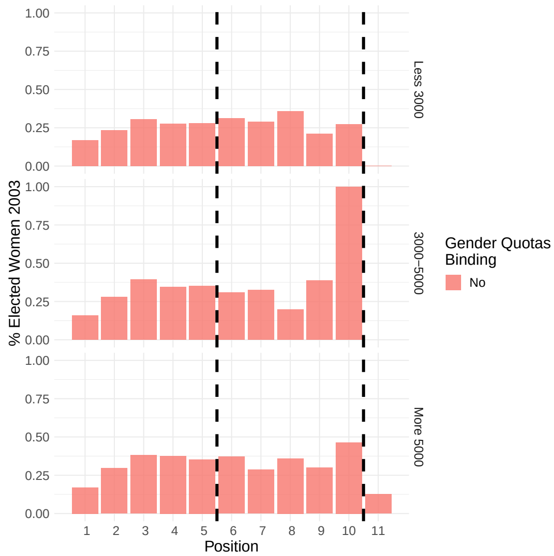
(c) 2011



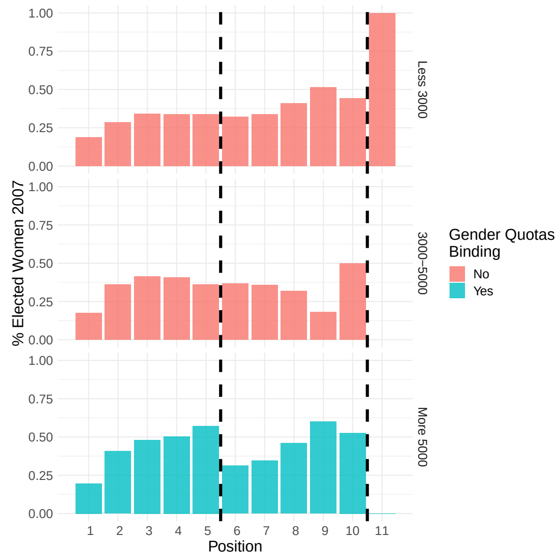
(d) 2015

Note: Unit of observation is the candidate. Sample is composed by all candidates in municipalities with population between 1,000 and 8,000. Each bar denotes the proportion of candidates that are women at each position in the list. The dashed vertical lines denotes the five position bracket cutoffs. The color of the bar denotes the treatment status for those municipalities in that year: red for non treated, blue when gender quotas are binding.

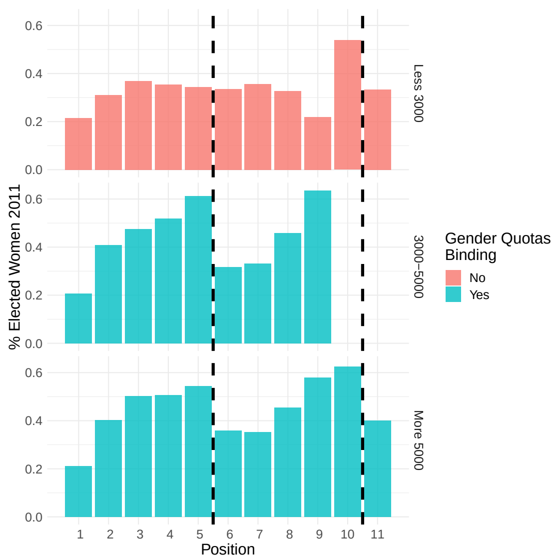
Figure 1·7: Proportion of Women Elected by Position



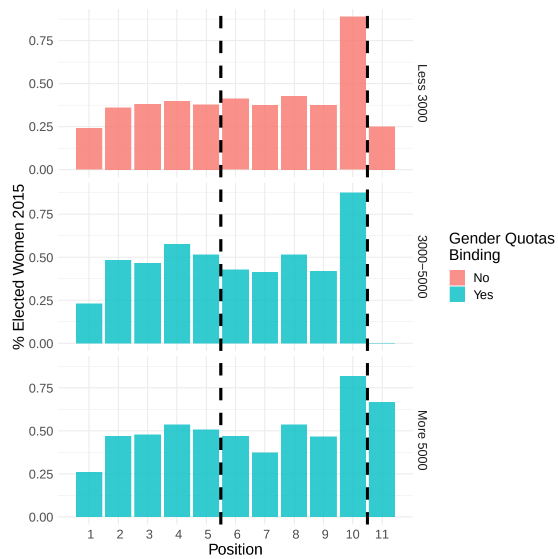
(a) 2003



(b) 2007



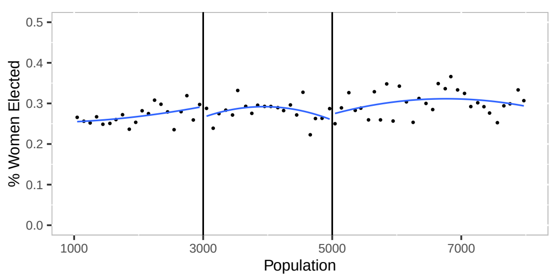
(c) 2011



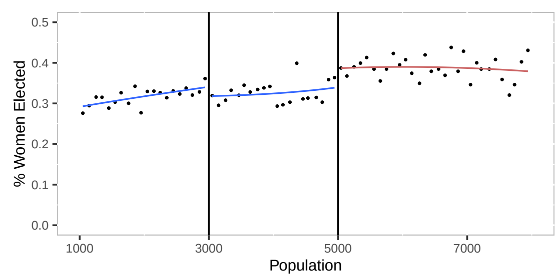
(d) 2015

Note: Unit of observation is the candidate. Sample is composed by all candidates in municipalities with population between 1,000 and 8,000. Each bar denotes the proportion of candidates that are women at each position in the list. The dashed vertical lines denotes the five position bracket cutoffs. The color of the bar denotes the treatment status for those municipalities in that year: red for non treated, blue when gender quotas are binding.

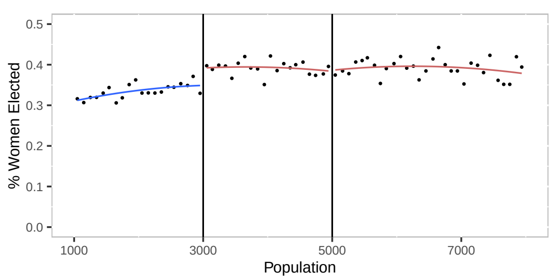
Figure 1·8: Proportion of Women Elected by Year



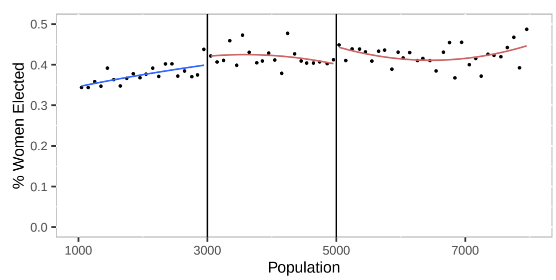
(a) 2003



(b) 2007



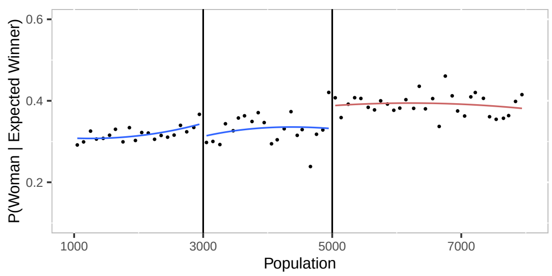
(c) 2011



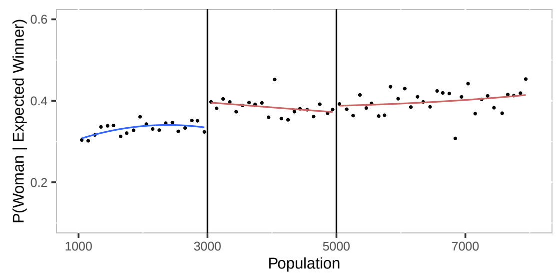
(d) 2015

Note: Unit of observation is the candidate. Sample is composed by all candidates in municipalities with population between 1,000 and 8,000. The vertical lines denote the population cutoffs that determine treatment status. Each dot represents an average of the outcome variable in a bin of size 100 of population. The blue and red lines denote local polynomial regression fitted lines with *loess* method. The color denotes the treatment status for those municipalities in that year: blue for non treated, red when gender quotas are binding.

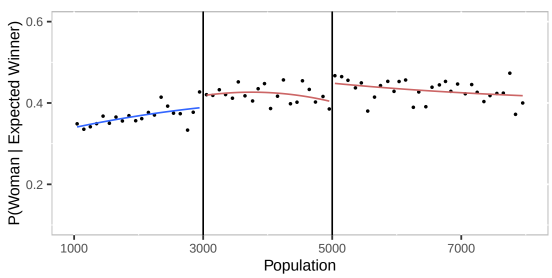
Figure 1·9: Proportion of Women in Realistic Positions



(a) 2007



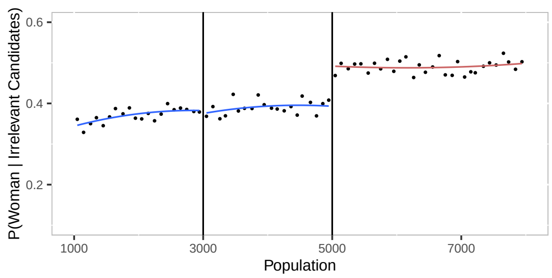
(b) 2011



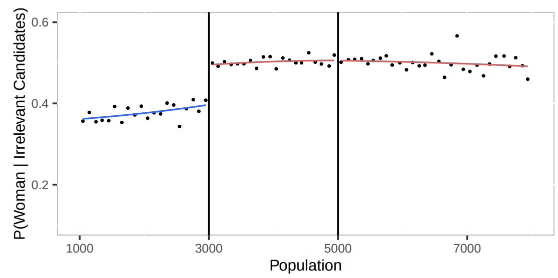
(c) 2015

Note: Unit of observation is the candidate. Sample is composed by all candidates in municipalities with population between 1,000 and 8,000. The vertical lines denote the population cutoffs that determine treatment status. Each dot represents an average of the outcome variable in a bin of size 100 of population. The blue and red lines denote local polynomial regression fitted lines with *loess* method. The color denotes the treatment status for those municipalities in that year: blue for non treated, red when gender quotas are binding.

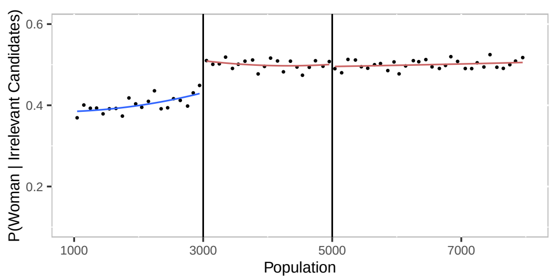
Figure 1·10: Proportion of Women in Irrelevant Positions



(a) 2007



(b) 2011



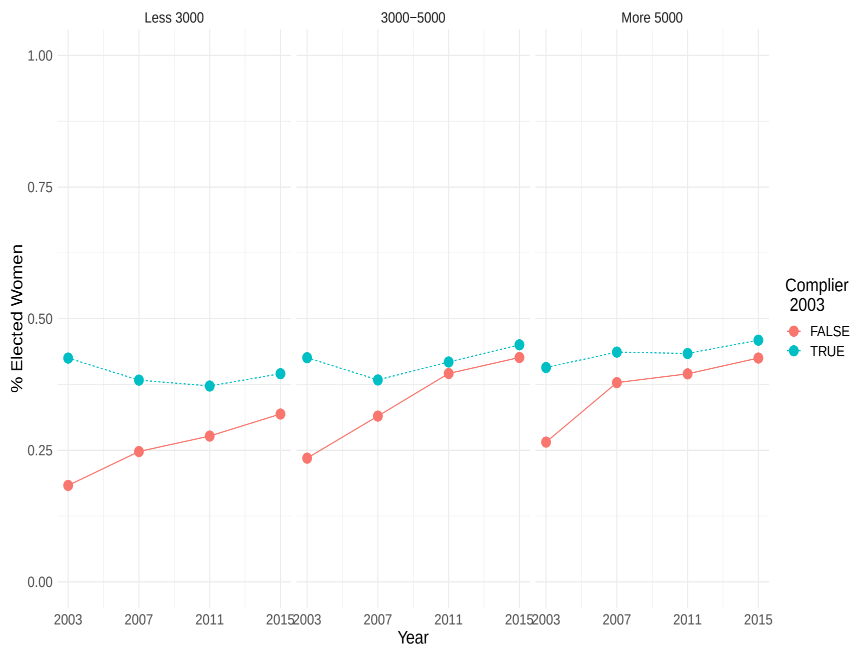
(c) 2015

Note: Unit of observation is the candidate. Sample is composed by all candidates in municipalities with population between 1,000 and 8,000. The vertical lines denote the population cutoffs that determine treatment status. Each dot represents an average of the outcome variable in a bin of size 100 of population. The blue and red lines denote local polynomial regression fitted lines with *loess* method. The color denotes the treatment status for those municipalities in that year: blue for non treated, red when gender quotas are binding.

Figure 1·11: Evolution of the Percentage of Women Candidates by 2003 Party Complier Status

Note: Unit of observation is the candidate. Sample is composed by all candidates in municipalities with population between 1,000 and 8,000 for which political party can be linked over time. Each graph corresponds to candidates in municipalities of a different population bracket. Municipalities with population lower than 3000 are never treated. Municipalities with population higher than 3000 and lower than 5000 are subject to gender quotas starting in 2011. Municipalities with population higher than 5000 are subject to gender quotas starting in 2007. Each dot represents the proportion of women candidates in that type of municipality and year. The color denotes whether candidates belong to a party that was already in compliance with the gender quotas in 2003 or not.

Figure 1·12: Evolution of the Percentage of Women Elected by 2003 Party Complier Status



Note: Unit of observation is the candidate. Sample is composed by all candidates in municipalities with population between 1,000 and 8,000 for which political party can be linked over time. Each graph corresponds to candidates in municipalities of a different population bracket. Municipalities with population lower than 3000 are never treated. Municipalities with population higher than 3000 and lower than 5000 are subject to gender quotas starting in 2011. Municipalities with population higher than 5000 are subject to gender quotas starting in 2007. Each dot represents the proportion of women that are elected in that type of municipality and year. The color denotes whether candidates belong to a party that was already in compliance with the gender quotas in 2003 or not.

Chapter 2

The More the Merrier: Evidence of Misreporting Population for Political Gain by Spanish Municipalities

2.1 Introduction

The relationship between fiscal and political institutions has long been of interest to economists (Alesina and Perotti (1999)). We have long theorized about the optimal government size, and the consequences of different degrees of fiscal federalism and political decentralization. One influential public choice theory, developed by Brennan and Buchanan (Brennan and Buchanan (1977), Brennan and Buchanan (1978), Brennan, Buchanan et al. (1980)) is the Leviathan model of government. According to this model, the goal of bureaucrats is to maximize the size of the government. There are two constraints to the Leviathan behavior of governments: constitutional constraints and intergovernmental competition. In this paper, I show, first, that Leviathan behavior is present in Spanish municipal elections, and second, that local governments circumvent constitutional constraints by manipulating population figures to maximize council size.

Elites and politicians using demographic manipulation to gain political influence is not new behavior. The landowners and peers who controlled *"rotten boroughs"* in 18th century England, for example, bribed the electors of these underpopulated constituencies in order to send their preferred candidate to the House of Commons. In the modern United States, district boundaries are manipulated (*Gerrymandered*) in order to obtain the desired demographic electoral composition. This paper provides evidence of direct manipulation of population reporting by Spanish officials to increase the size of the government.

Most of the economic literature that studies the relationship between political institutions and fiscal policy has focused on the impact of government size on public spending. In the United States[1], economists have focused on the effect of the mayor's council or council manager on fiscal policy[2]. Coate and Knight (2011) find that US municipalities governed by city councils have lower public spending. MacDonald (2008) tests for the effect of council size on public spending, finding none once fixed effects are included. Outside the US, Pettersson-Lidbom (2012) analyzes two natural experiments and finds that an increase in council size is associated with a decrease in government spending in Finland and Sweden.

This examines how municipalities circumvent constitutional and regulation constraints that dictate council size in local municipalities in Spain. Spain is particularly interesting due to its high degree of decentralization. According to Fernández-Caballero, Pedregal, and Pérez (2011), close to 50% of all government spending is managed by regional or local governments in Spain. Local elections in Spain have previously been studied in the literature, but with a slightly different focus. Costas-Pérez, Ollé, and Navarro (2011) account for the effect of corruption scandals on the probability of Spanish mayors' reelection. Of closer relevance to this project, Curto Grau, Solé Ollé, and Sorribas (2012) study the par-

[1] A comprehensive review of the literature can be found in Besley and Case (2003)

[2] Another political institution that economists have studied is tenure in office. (Coviello and Gagliarducci, 2010) exploit a natural experiment in Italy to study the relationship between tenure and economic outcomes. They find that longer tenure is associated with worse economic outcomes.

tisan targeting of government transfers. The authors find that transfers to a municipality are higher when the political sign of the local government is the same as the state or national government.

The main results of this paper are as follows: first, there is an unusual concentration of municipalities (bunching) with reported populations just above the threshold that increases the number of local representatives. Second, I present compelling evidence that elected officials manipulate population figures in order to maximize the size of the council in the upcoming election. In particular, municipalities are twice more likely to be found right after the threshold during election years compared to non election times. Also, it does not seem to be the case that what drives municipalities to bunch and misreport their population figures is the desire to obtain more responsibilities, or a higher degree of political decentralization. I also present evidence that transfers do not change discontinuously at the threshold that determines council size. Finally, I use luminosity data to construct a measure of misreporting and find that municipalities that have higher turnout tend to have higher levels of misreporting.

In its methodology, this paper is closer to the literature on bunching. This literature was developed initially by Chetty, Friedman, Olsen, and Pistaferri (2009) and Saez (2010), who study whether workers bunch at kink points of the income tax bracket, where there is a discontinuity in the marginal tax rate. There is less available literature on the relationship between bunching and local governments. Perhaps the best study to date in this area is Camacho and Conover (2011). The authors study how local governments in Colombia manipulate a poverty index in order to stay below a certain threshold that determines whether households in that municipality obtain state funds. Outside the bunching literature, the relationship between local government and corruption has also been explored: see Bergh, Fink, and Öhrvall (2012) in the context of Sweden, and Olken (2006) in Indonesia.

This paper also contributes to the literature of Regression Discontinuity Designs using population thresholds. Several papers (for instance (Bagues and Campa, 2017), (Campa, 2011), (Casas-Arce and Saiz, 2015)) use this cutoffs without taking into consideration that municipalities misreport their population figures. If misreporting is not random, as shown in this paper, and the characteristics of the bunching municipalities are correlated with the outcomes of interest in those papers, the estimated effects might be biased.

The rest of the paper is structured as follows. Section 2 discusses the Spanish institutional setting, the particular law analyzed and the identification strategy. Section 3 details the data used, and Section 4 shows the main results of the paper. Finally, Section 5 concludes.

2.2 Institutional Setting

Politically, Spain is organized into autonomous communities, provinces and municipalities. Municipalities constitute the lowest level of territorial division. There are 8,112 municipalities, all of which hold elections simultaneously every four years, usually the fourth Sunday of May, via a multi-party proportional list system. In every election, more than sixty-six thousand local representatives are elected. The mayor is then elected by the plenary assembly of these representatives. Spanish local governments are in charge of basic services, such as lighting public spaces, waste collection and management and water supply. Their revenue comes, mostly, from local taxes, income from self-owned assets, subsidies, and transfers from the central and autonomous government.

In this study, I am interested in the legislation that regulates how many politicians will be elected in the upcoming election. According to the law governing municipal elections, the number of representatives assigned to a municipality is determined by that municipality's population. Specifically the "Ley Orgánica 5/1985 del Régimen Electoral General" states the following: if the population in the municipality is of less than or equal to 100 people, the municipality is assigned three representatives. If the municipality has more than 100 and less than 250 inhabitants, five local representatives will be chosen, and so on. The full assignment rule can be seen in Table 2.1 and Figure 2·1. For municipalities whose population is higher than one hundred thousand, one representative is added for each additional hundred thousand people. If the result is an even number, another representative is assigned.

Given the findings of this paper, it is crucial to understand the process by which population is reported. Every 10 years, a national Census is carried out by the National Institute of Statistics (INE). Due to its low frequency, however, the Census is not enough to assign representatives. During the years in which a Census is not available, municipalities are in charge of creating the "Padrón Municipal" or Municipal register. According to the National Institute of Statistics: *"the Municipal register is the administrative register where municipality inhabitants are recorded. The respective town councils are responsible for the register's formation, maintenance, revision and custody. Its update is obtained from the revision of the municipal register on the 1st of January of each year, which is approved by the Government at the INE's request, after a favorable report by the Registration Board"*. Municipalities are in charge of the maintenance, creation and revision of the Padrón, while the National Institute of Statistics oversees that the reported numbers are correct. In 1996, there was a major modification of register regulations: *"a new continuous and computerized management system for municipal registers was established, based on the coordination of all of them by the National Statistical Institute"*. I will show that whether population is measured through a National Census or Municipal Padrón matters. I find that there is only manipulation when the local governments are the ones reporting population figures.

2.3 Data

The data set I compile includes information on Spanish municipalities' elections and their economic outcomes over time. As of 2014, Spain has a total of 8,122 municipalities. The first democratic municipal elections in Spain took place in 1979. Local representatives and mayors are chosen every four years.

The data set on local elections contains results from 1987 to 2011, a period of seven elections for each of the Spanish municipalities. This data is publicly available at the Ministerio del Interior (Spanish Ministry of Interior). The data set contains information on the number of votes for each party, as well as the total number of votes and the population of each municipality. Population figures are used to assign the number of seats or local representatives that each municipality will elect. The method for allocating seats is known as D'Hondt method. Given that the number of votes for each party and the total number of seats are known, I can back out the number of seats that each political party is assigned. In elections where the most voted party obtains more than 50% of the seats, the first candidate on that party's list becomes mayor. There is never a case in which the winning party obtains exactly 50% of the seats, because the number of seats is always odd. To identify the mayor's party, which is not trivial in cases where there is a coalition, since it is not necessarily the party with the most votes, I use a second data set that contains a list of mayors for each municipality and election. This data is also available at the Ministerio del Interior.

The average municipality has a population of around 5,000, although the median population is closer to six hundred. As we can infer from Table 2.2, Spain has gone through a process of urbanization in the last two decades, with people moving from small rural areas to cities. Consequently, the distribution of population has become more skewed, with large population centers growing even larger over time, and smaller municipalities becoming smaller.

In Table **??**, I include a summary of important political outcomes. Close to three parties obtain representation in the council on average. Turnout for local elections is high, around 60%. The winning party (in terms of votes) usually achieves two thirds of the seats, and the probability of a majority is 80%. This is due to the D'Hondt allocation method, which makes each consecutive seat "cheaper" to obtain. In a multi party system, it is difficult to construct a measure of the margin of victory. A proxy for that is the difference in vote shares between first and second most voted parties. On average, the share of votes by the winning party is 26% higher than that of the runner up. Finally, I include two dummy variables that measure how prevalent one and two party councils are. These are defined as councils where only one (two) party is represented. One party councils are present in 11% of the municipalities, whereas two party councils are more common, and compose 36% of the sample.

When investigating the motives for misreporting population, I consider the possibility that higher transfers is the reason for this behavior. To study this hypothesis, I combine the electoral results data set with annual data on each municipality's budget from 1985 to 2009. This second data set, provided by the Ministerio de Hacienda y Administraciones Públicas, contains information on the total public spending and revenue of each municipality. It also breaks down spending and revenue into subcategories. Public spending is thus divided into the following subcategories: spending in personnel, services, infrastructure, transfers, financial expenses and financial assets. Revenue subcategories include: direct and indirect taxes, fees, transfers received, sale of assets, financial revenue and public debt. Overall, I am able to link 21,461 elections with budget data.

A summary of public spending and public revenue is included in Table 2.4. The main spending categories are personnel, general services and infrastructure, together constituting more than 85% of total spending. Because of this, combined with the fact that these are the areas where municipalities enjoy a higher degree of discretion, I will be focusing on these three categories in the empirical analysis. The major revenue streams for these municipalities are direct taxes, fees, capital transfers and transfers from other public entities, such as Autonomous Communities and the Spanish central government.

Finally, in order to create an unbiased measure of misreporting, I match my data of municipalities' electoral results and public spending with the Defense Meteorological Satellite Program's luminosity data. This data, available worldwide for the period 1992 to 2013, contains measures of nighttime lights visible from space. Night lights are measured at a very small grid (30 arc seconds grid, or approximately 1 kilometer) and are bottom and top coded. My measure of luminosity for a municipality in a given year is the sum of the average yearly night light measure in each grid inside that municipality. Since I am trying to obtain an unbiased measure of population, the sum seems to be a better choice than the average.

In Figure 2·2, I illustrate the geographic distribution of night lights and their evolution over time with two heat maps. The first one shows the deciles of night lights in Spain in 1995, and the second one in 2011. Brighter areas represent areas with higher population or night light measures. As we can see, they generally track population, since the capital and the coast tend to be brighter, as well, as areas in the interior having lower population. Finally, the luminosity measure is getting brighter over time, consistent with a positive population growth.

2.4 Results

2.4.1 Evidence of Political Bunching

In this section, I present the results of this paper. Following the potential framework econometric literature, I want to define the treatment (D) as follows:

$$D_i = \begin{cases} 1 \text{ if Treated (Above Threshold)} \\ 0 \text{ if Control (Below Threshold)} \end{cases}$$

In this case, I have eight different discontinuities or thresholds ($RD_1, \ldots, RD_8$) at which point an extra representative is assigned. However, it seems reasonable to assume that the relevant threshold for each municipality is the closest one. If the cost of misreporting (or the probability of getting caught) increases in the distance between true population and reported population, then the only relevant threshold to a municipality is the one closest to them. Therefore, I define a new forcing variable, "distance to the threshold", Z, that measures the distance, in percentage terms, to the closest discontinuity (RD^*). The choice of percentage rather than absolute distance can be discussed. Suppose your municipality is 100 people short of the threshold: I believe it is plausible that the cost of misreporting and adding 100 people to your population registry is much lower for a highly populated municipality than a smaller one. On the other hand, the cost of adding 5% of your population could potentially be roughly similar independent of population levels, considered in terms of the probability of getting caught by the National Institute of Statistics. Following that reasoning, the new forcing variable Z defines the treatment (being above the closest discontinuity) as:

$$Z = \frac{\text{population} - RD*}{RD*}$$
$$RD^* = \underset{RD_1,\ldots,RD_8}{\text{argmin}} \left| \frac{\text{population} - RD_k}{RD_k} \right|$$

Given that forcing variable, the treatment is just a function of the distance between the reported population and the nearest discontinuity. That means that all municipalities with a positive (negative) Z are closer to a discontinuity "from the left" ("from the right"): for my analysis, they will count as those who obtained extra representatives to be elected in council.

$$D_i = \begin{cases} 1 \text{ if } Z > 0 \\ 0 \text{ if } Z \leq 0 \end{cases}$$

The main result of this paper is shown in Figure 2·3. In this graph, each point represents the number of municipalities in a small bandwidth of the forcing variable Z. The blue line shows a quadratic fit for the relationship between density and the distance to the threshold. As we can observe, right after $Z = 0$, there is an unusual spike in the number of municipalities.

One naive explanation would be that, when reporting population, municipalities tend to choose round numbers. However, if that were the case, bunching would be seen at the left of the discontinuity, since the law stipulates that additional representatives are awarded when the population reaches 101, 251, 1,001 and so forth.

Stronger evidence that it is indeed the municipalities who are misreporting population figures is shown in Figure 2·4. This figure shows the distribution of municipalities for each election cycle. According to this Figure, there is no bunching in 1991. This fact actually strengthens the theory of political manipulation, since in 1991 there was a decennial national Census. This means that the National Institute of Statistics (INE) was the one reporting population figures. In another project, I take advantage of this fact to estimate the causal impact of government size on public spending and electoral outcomes. Every other year, it is up to the local municipalities to approve those numbers. As previously mentioned, a modification of the Padrón regulations was carried out in 1996. However, it does not appear that such modifications were successful in preventing municipalities from misreporting population: bunching is still evident in 1999 (Figure 2·4d), less so in 2003 (Figure 2·4e) but very clear again in 2007 (Figure 2·4f).

I reinforce the previous argument in Figure 2·5 and Table 2.5. The figure shows the amount on bunching on Census years (2011, 1991) versus the rest, where population was reported by the municipalities. As we can see, bunching is way more prevalent in non Census years. I can formalize this in a regression framewrok: in Table 2.5, I regress whether a municipality is found *right above their nearest threshold*, defined as within 5% above the threshold on whether the election happened on a year where a National Census was taking place. I control for population and distance to the discontinuity, as well as municipality and year fixed effects. Clustered standard errors at the municipality level are shown in parentheses. The results indicate that bunching is way less prevelant in years where a Census took place, confirming the visual evidence shown before.

2.4.2 Alternative Hypothesis

In this section, I will discuss some alternative hypotheses of what could motivate municipalities to misreport population figures above these particular thresholds. One possible explanation is that municipalities are indeed choosing to misreport population figures, but that they are doing so with no political gain in mind. This could be the case if they are trying to get over a particular population threshold in order to obtain any other privileges (grants, transfers from other government level, additional responsibilities...). There are a couple of things that can be done to alleviate those concerns.

First, tables 2.6 and 2.7 show the level of bunching in election versus non election years. If municipalities are misreporting population figures for motives other than electoral ones, there should be no differential spike in election years. To test this hypothesis, I focus on the sample of municipalities that are above the discontinuity for the years 2000 to 2014. The reason for choosing municipalities only above the threshold is comparability: all of these local governments have obtained extra council seats, so that eliminates one confounding variable. I use a linear probability model to estimate the probability of reporting population figures just above the threshold:

$$1\{\text{Right Above}\}_{it} = \gamma \text{Election Year}_t + \beta X_{it} + u \quad \text{if } \|Z_{it}\| < \epsilon \tag{2.1}$$

The dependent variable is a dummy that takes the value of one if the variable is right above the discontinuity, where bunching is observed. I determines that threshold graphically to be either 0.01 or 0.02%. Results are reported for the smallest interval, but robust to either. The main independent variable is a dummy for if a local election was held that year. The results show that the probability of any municipality reporting population right above the threshold is much higher during an election year that a non election year. In particular, the estimates suggest that the probability of a municipality reporting population levels right above the threshold jumps 50 to 150% in an election year. Results are robust to whether we restrict the sample to municipalities below within 5% or 1% of the threshold, or to the controls included (population or year fixed effects). This is further evidence of politically-motivated manipulation of the population figures.

Second, another possible motive for manipulation is tied to the fact that the level of decentralization changes at some of those thresholds. Whereas all municipalities have some basic responsibilities (street lightning, waste collection, water supply, road management), according to Spanish law, those responsibilities increase with the population of the municipality. Local councils in municipalities with more than 5,000 inhabitants must also provide public parks, public libraries and waste treatment; municipalities with more than 20,000 inhabitants must offer social services, fire fighters and a slaughterhouse. Finally, those with population levels over 50,000 are required to provide public transportation and environmental protection. It could be the case that local councils choose to misreport population in order to obtain additional responsibilities: if politicians are corrupt, and embezzlement, theft or nepotism are easier when the size of the local government increases, it could be optimal for them to misreport population. If that is the main reason for the observed bunching, however, we should only observe a spike in density right after those three thresholds, and not in the rest. The opposite could also be true: if municipalities' officials are time constrained and/or dislike work, they could instead desire less responsibilities. In that case, we would expect less bunching, or even bunching at the other side of the cutoff when responsibilities increase.

I test that hypothesis in Figure 2·6. In this figure, I compare the level of bunching when a) both the number of representatives and the level of decentralization increases and b) when only the size of the council increases. It can be observed that there is bunching in both situations. This seems to dismiss the hypothesis that municipalities are bunching in order to extract rents from more responsibilities.

Finally, I test the hypothesis that municipalities choose these population figures because they could potentially obtain more transfers from other levels of government by getting over certain population thresholds. One naive test would be to compare transfers below and above the cutoff using a Regression Discontinuity design. However, this would be a useless exercise. First, I have just shown that population is not exogenous: officials are misreporting population figures, so any comparison of municipalities above and below the threshold would compare different types of observations. Second of all, in election years, the *ceteris paribus* assumption does not hold. Even if municipalities where not misreporting population, council size increases at those thresholds. Then it would be impossible to know whether transfers change because population changes, or because the number of representatives in the local council rises.

However, I can test this hypothesis by focusing on the year 2001. The reason for this is twofold. First, 2001 is one of the years in which a National Census took place. This means that population reports should be accurate and unmanipulated. We can check this by using the McCrary test. Second, there were no municipal elections in 2001, but there were in 1999 and 2003. This implies that there is no change in council size from moving above or below the threshold. Therefore, it is plausible to compare municipalities above and below the threshold, as long as I control for whether they were above or below in 1999. If grants or other sources of public funds vary discontinuously at those thresholds, we can measure that with a Regression Discontinuity design.

I test the hypothesis of no manipulation by plotting the density of the forcing variable (distance to the discontinuity) in the year 2001 and performing the McCrary test. Figure **??** shows that we can't reject the null of no manipulation. Together with the knowledge that there was a National Census in 2001, there is reasonable evidence to trust the results regression 2.2. In this regression, the dependent variable is the amount of transfers per capita in a particular municipality. The parameter of interest is γ_1, whether transfers change significantly whether a municipality is above the threshold (in 2001). I control for whether the municipality was above the threshold in the previous election (1999), as well as flexibly for distance to the discontinuity, population and discontinuity dummies. Each column represents a different bandwidth of observations: 5, 10, 15, 20 and 25%.

$$\text{Transfers}_i = \gamma 1\{\text{Right Above}\}_i + \gamma 1\{\text{Right Above 1999}\}_i + f(Z_i) + X_i'\beta + u_i \quad \text{if } \|Z_i\| < \epsilon \tag{2.2}$$

The results in Table 2.8 show that transfers per capita do not change substantially at the threshold in 2001. This seems to indicate that perhaps capture of transfers are not the main driver behind the behavior of these municipalities. The effects for the 1999 dummy are negative and sometimes statistically significant. However, it should be remembered there is no causal interpretation here: municipalities do manipulate population in 1999.

To conclude, it seems that politicians are indeed manipulating population in order to increase the number of seats in local councils in the next election. This is consistent with models of corruption where the goal of the party is to maximize the number of representatives it obtains, not the number of votes.

2.4.3 Determinants of Misreporting

So far, we have established that local governments are over-reporting population figures. I have shown that this is probably due to political motives, given that over-reporting is greater during election years, and it is not due to a desire for higher decentralization, nor to capture transfers. In this final section, the aim is to understand when misreporting happens, and which municipalities are more likely to engage in this type of behavior.

In the bunching literature, determinants of different kinds of behavior that result in bunching observations bunching are usually explored. The methodology is simple, and it involves regressing a dummy for whether a particular observation is in the bunching area on a set of observables. Here, we can test whether bunching is more or less prevalent given a set of measures of the democratic health in that particular municipality. In particular, to measure the democratic institutions in a municipality I use the following variables: turnout, probability of majority, percentage of seats obtained by the winning party, the vote differential between winner and runner up, the share of protest votes and the number of parties that obtain representation in the council. I run the regression for the whole sample in the first column, and comparing those right above (the ones suspect of misreporting) to those in the next bin (with distance greater than 5% and less than 10%). Standard errors are clustered at the municipality level.

The results of this test are shown in Table **??**. The first main take away is that bunching is more likely to occur in municipalities with worse democratic outcomes: lower levels of turnout, and higher degree of protest votes, both signals of an unhappy electorate. One could be tempted to argue that the positive coefficient on majority means that bunching is more likely when there is a dominant party, but that is at least partially offset by the impact on the number of seats obtained by the first party, which is negative.

This approach, while offering some interesting hypothesis, does not work particularly well in this scenario. This is because the behavior of interest is misreporting, not bunching. When I created a dummy for whether a municipality is just above the threshold, that dummy encompasses two types of municipalities: those that happen to be just above the threshold because of their "natural" population, and those who misreport their population figures in order to increase council size in the next election. The inability to distinguish between those two types of municipalities makes the results shown in Table **??** less clear.

To address this problem, I propose obtaining an unbiased measure of population using luminosity data and census years population figures. The rationale for using census year population figures is that it seems like misreporting was not present in those years. Therefore, I can use 1991, 2001 and 2011 population reports to infer what the actual population was on years where misreporting takes place. In addition, I propose using yearly night lights measures to improve the accuracy of the imputation method. Luminosity data has long been used as a measure of population density, particularly in developing countries. Based on these two set of variables I will obtain an estimate of “true" (as in opposite of reported) population. I can then create measures of misreporting, and find the characteristics of the municipalities that are more likely to display misreporting behavior.

For this strategy to be valid, I will need to make four assumptions. First, only municipalities near the threshold misreport population. The assumption is that, for municipalities “far away" from the threshold, the reported population is the true one. If there are only gains to be had by going over the threshold, and it is costly to misreport, it is not profitable to misreport unless it is done in order to be over the discontinuity. This assumption is violated if municipalities have imprecise control over the reported population, or if there are other unobserved bunching points. I need this assumption to correctly estimate the relationship between night lights and population. Second, there was no misreporting in Census years. This allows me to use 1991, 2001 and 2011 population levels as a signal of other years' true population. I have already shown evidence that there bunching is way less severe at the threshold during these years, since National Censuses were carried out. Third, past luminosity data is a good predictor of population. Given this testable assumption, I will use average night light measures in a municipality in each of the previous three years as a predictor of population in election year. It is important to note that I will not use contemporaneous measures of night lights. I choose not to do this due to endogeneity issues: the measure of night lights could increase due to the behavior of elected officials or public spending decisions that might be correlated with possible manipulation. Finally, I need to assume that the relationship between night lights and population in Census years with current population is the same for those near the discontinuity and those that are not. This is a critical non testable identifying assumption. It would be violated if municipalities near the discontinuity are different in some dimension that makes them have different measures of night lights compared to equally sized municipalities far away from the discontinuity. If satisfied, this assumption allows me to impute predicted populations to municipalities suspect of “cheating", based on a relationship established for those those far away from the discontinuity.

If the previous assumptions are met, I can think of manipulation of population figures (or, for simplicity Z, since population uniquely determines Z) as a problem of non-classical measurement error, where the econometrician does not observe the true population (Z^*), only the reported one:

$$Z_{it} = \begin{cases} Z^*_{it} + u_{it} & \text{for those Right After the Discontinuity} \\ Z^*_{it} & \text{otherwise} \end{cases}$$

It is important to note that the error term u_{it} is not necessarily white noise. In particular, we expect u_{it} to be correlated with other municipality characteristics, observed and unobserved, such as public spending or corruption. This is not a problem for this procedure. The idea is to think of population as a function of 1991 population and previous measure of night lights, estimate that relationship and then use those estimates to impute an unbiased measure of “true population". Based on that measure, I can then construct the estimated true distance to the nearest threshold $\hat{Z}^*{}_{it}$.

$$\begin{aligned} \text{Population}_{it} = f_t(&\text{Population}_{i,1991}, \text{Population}_{i,2001}, \\ &\text{Population}_{i,2011}, \text{Night Lights}_{i,t-1}, \\ &\text{Night Lights}_{i,t-2}, \text{Night Lights}_{i,t-3}) \end{aligned}$$

In order to determine the best imputation procedure, I use different techniques from the machine learning literature. Independently of the actual technique used, the imputation is performed outside of the so called “excluded region". The excluded region is chosen visually and, consistent to the figures where bunching was estimated, it is determined to be all observations with distance to the nearest threshold smaller than 5%. Therefore, I restrict my sample to observations not on the excluded region. The rest of the observations are again split into two further subsamples: a training set with two thirds of the leftover observations, and a validation set with the other third. On that training set I estimate nine different imputation procedures. First, a simple OLS regression is used, since it’s been usually found that OLS is best at capturing linear relationships. I run an OLS model with the four variables of interest. Secondly, we repeat the first exercise including second and third degree polynomials of the four variables of interest to try to capture non linearities. Third, it has been shown that random forests are quite efficient at capturing non linearities, so I estimate a random forest where the number of trees is set to half of the size of the training set. Finally, in a spirit similar to ensemble learning techniques, I combine the random forest predictions and the OLS predictions. The rationale is that this method should excel at capturing both linear relationships and non linearities present in the data generating process. The four estimation procedures are then applied to the validation set. To compare among them, I calculate the mean square error of each method (the average of the square of the difference between population and estimated population).

The results of the imputation procedure are shown in Table **??**. According to that table, the random forest procedure yields the lower mean square error out of sample, so I use that method to impute population. The relationship between reported population and estimated population is shown in Figure 2·7. We can see that the estimation works best for lower levels of population (below 30,000). However, given that most of the discontinuities happen below those levels of population, this is not a concern for our analysis.

At this point, I “observe" both the reported population of the municipalities, as well as their “true" population. Based on those, I observe the percentage distance of the discontinuity, reported and estimated: Z_{it}, and $\hat{Z}^*_{it}$. Given these, I propose creating two measures that will try to capture misreporting. The first, and most natural one, is the difference between the reported population and the population that I estimate should be the truth. The second one, that I denote *Over-reporter* is a dummy variable that takes a value of one if the population reported by the municipality is higher than the estimated one.

Both variables try to capture the same concept: which municipalities are misreporting. Now I can regress these variables on a series of outcomes measuring the democratic health of the institutions. The explanatory variables that I choose are the following. Note, all of them are calculated for the previous election period, since that is the information that parties have when they decide to manipulate population reports. First: turnout. How involved are voters in the democratic process could affect whether parties feel comfortable enough to manipulate population figures or not. To calculate turnout, I divide total number of votes in the previous election period by 1991 population, since non census year population reports are unreliable. Second, a dummy for whether there was a majority in the previous election: if a single political party enjoys a majority, it could be easier to manipulate as well as reap the benefits of the extra council. The last three variables try to explore this mechanism as well: the percentage of seats obtained by the winning party, the difference in vote shares between the winning and the second party, and the number of parties that obtain representation.

The results are shown in Table 2.10. The most reliable results are shown in columns 2 and 4. In these two columns I restrict the sample to those observations in the bunching region (they are above the discontinuity, but their distance is lower than 5%). The intuition here is to compare municipalities in the suspect region by their predicted population. Those whose predicted population is lower than their reported one, how are they different than the rest? The results show that of the included variables, only two have a consistently significant impact on whether municipalities misreport population. The first is turnout in the previous election: the higher the turnout, the more higher the probability of misreporting by that municipality. This is counterintuitive, but it seems to be a strong effect. Second, the higher the number of parties obtaining seats in the previous election, the lower the probability of misreporting. This does indeed support our prior: the more parties, the higher accountability, and a lower potential benefit from misreporting.

2.5 Conclusion

This paper studies the behavioral response of local governments in Spain to the regulation that dictates the number of representatives to be elected in the next local election. Using data from all local elections held in Spain between 1987 and 2007, I find that municipalities bunch right after the discontinuities where additional representatives are awarded. This is consistent evidence with Leviathan behavior from Spanish local governments.

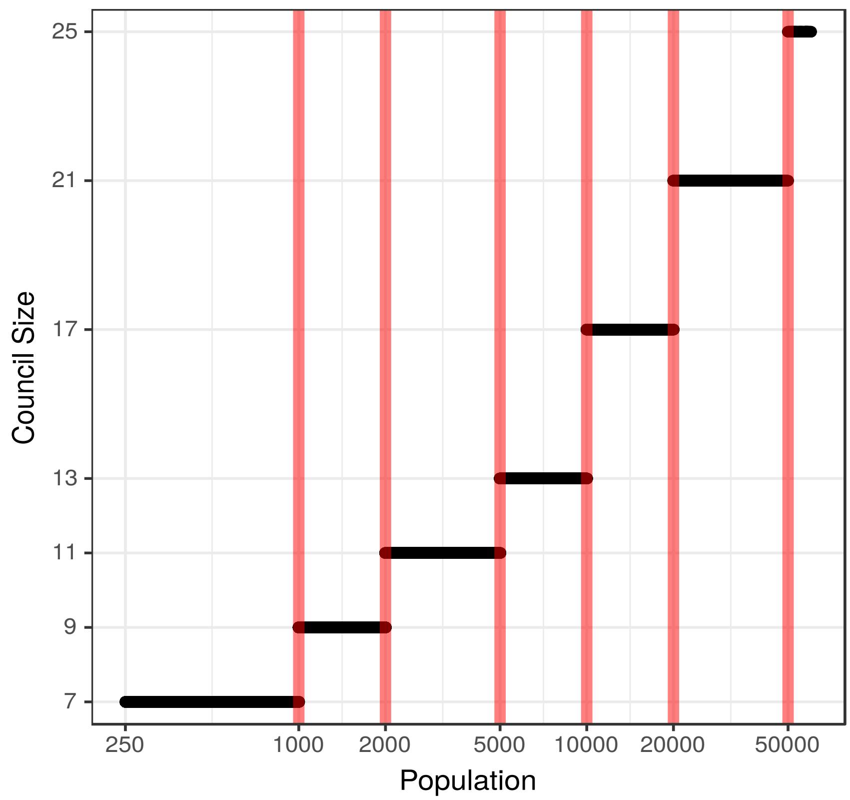
Bunching is present in every election year except for 1991. During 1991, a National Census was carried out. As a consequence, this is the only year in our sample in which the role of reporting population is taken away from local governments and is assigned to the National Institute of Statistics. It seems likely, then, that local governments and local politicians are actually misreporting population figures. My hypothesis is that they are doing so in order to maximize the number of representatives in local council, as predicted by Leviathan model of government.

To support the hypothesis that misreporting is carried out by political agents, I present several pieces of evidence: first, bunching is more likely to occur in election versus non election years. The probability of a municipality reporting population right after the threshold is 50-150% higher in an election than a non election year. Second, at some of the thresholds, the level of decentralization changes, i.e. municipalities have more responsibilities. However, bunching is still observed both at thresholds where the level of decentralization increases and remains constant, proving that this is not the main reason why municipalities misreport population. Finally, I use the year 2001, when another National Census took place and there was no election, to show that transfers do not change discontinuously at the threshold that determines council size. This is compelling evidence that misreporting is not taking place simply to capture higher transfers from other levels of government.

Finally, I try to better understand which conditions bread misreporting. I use machine learning techniques to create a measure of misreporting based on census population data and luminosity figures. Given this unbiased estimate of population, and the reported population, I construct two measures of misreporting. Based on both measures, I find that municipalities with higher voter participation in the previous election are more likely to misreport, and that municipalities where more parties were involved in the council are less likely to misreport.

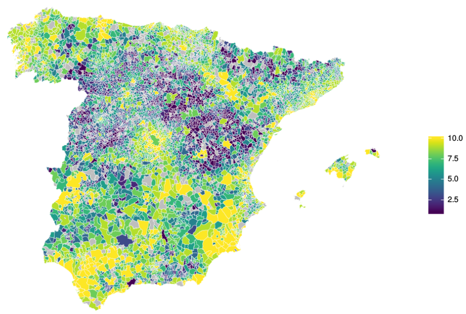
Figures and Tables

Figure 2·1: Assignment Rule of Representatives

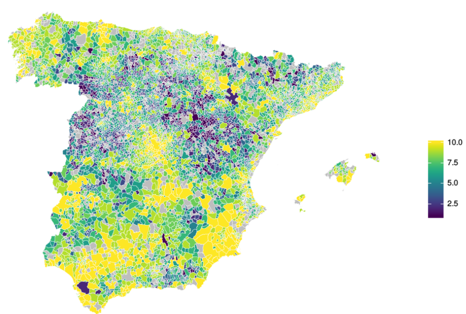


Note: Figure shows the number of politicians elected for each municipality (i.e. council size) as a function of the population. Population is displayed in log scale.

Figure 2·2: Evolution of Night Lights in Spain over Time



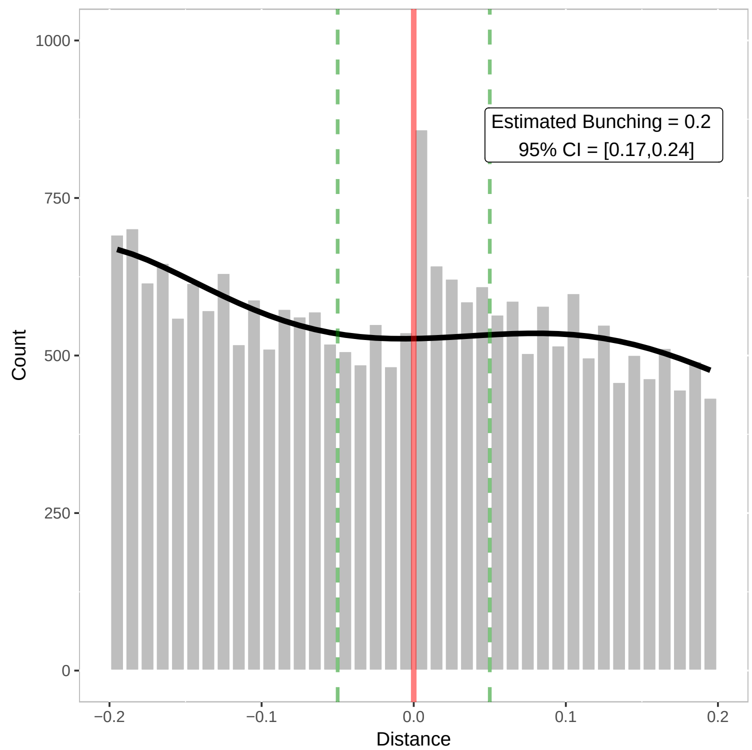
(a) Deciles of Night Lights in 1995



(b) Deciles of Night Lights in 2011

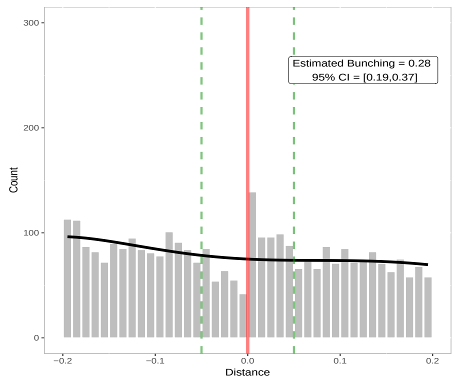
Note: The unit of observation is the municipality. Each map displays the decile of both night lights for the year 1995 and 2011. Brighter colors denote a higher decile of luminosity of population.

Figure 2·3: Count of Municipalities by Distance to the Discontinuity

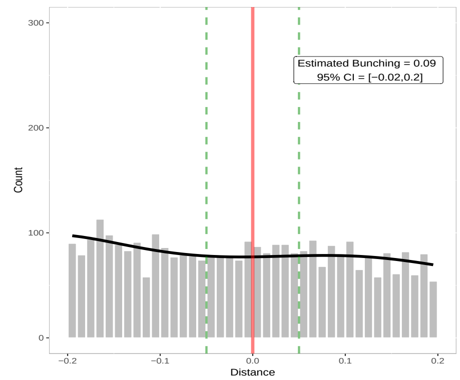


Note: The unit of observation is the municipality. Each dot displays the number of municipalities in a bin of 1% of distance to the nearest discontinuity. The estimated counterfactual distribution is calculated with a fifth degree polinomial of distance to the discontinuity. The dashed green lines denote the excluded region for the counterfactual discontinuity. Estimated bunching is calculated as the percentage of the overall mass in excess of the estimated mass in the excluded region above the discontinuity. The confidence interval is calculated via bootstrapping.

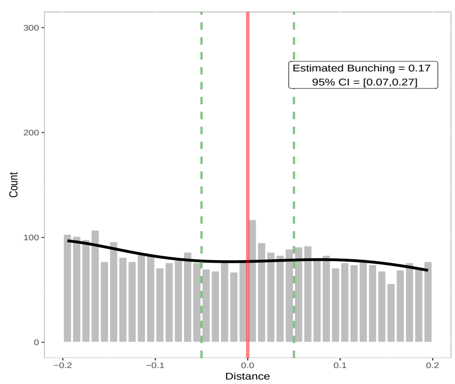
Figure 2·4: Bunching by election year



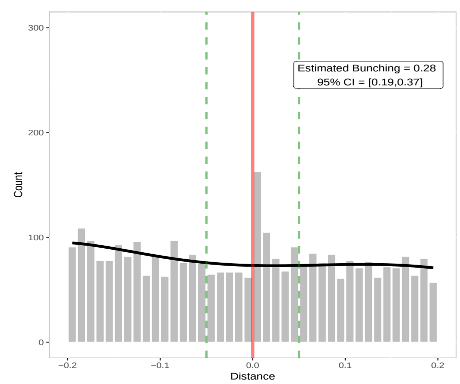
(a) 1987



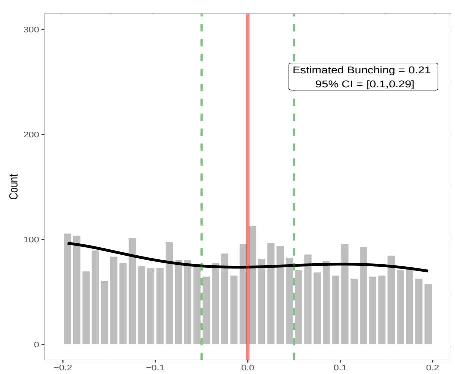
(b) 1991



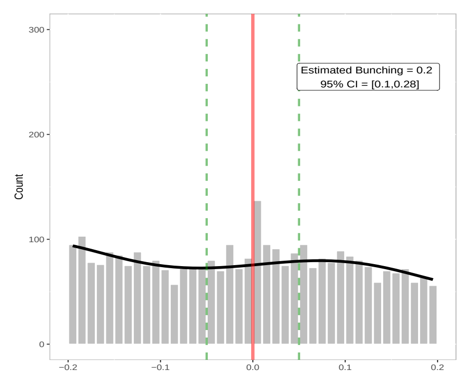
(c) 1995



(d) 1999



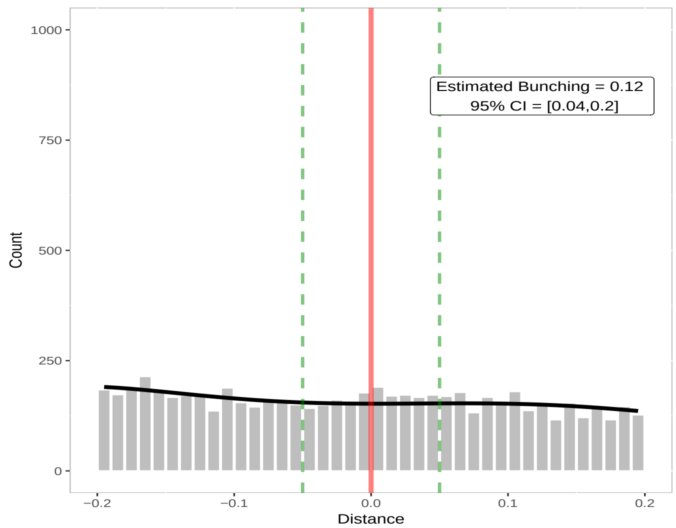
(e) 2003



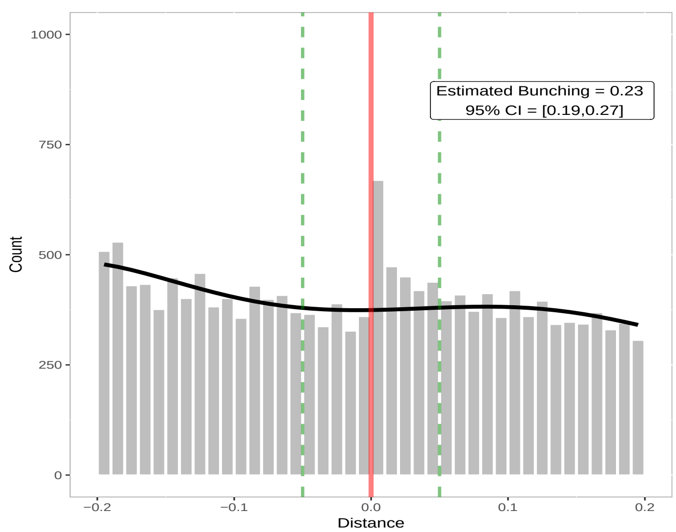
(f) 2007

Note: The unit of observation is the municipality. Each dot displays the number of municipalities in a bin of 1% of distance to the nearest discontinuity. The estimated counterfactual distribution is calculated with a fifth degree polinomial of distance to the discontinuity. The dashed green lines denote the excluded region for the counterfactual discontinuity. Estimated bunching is calculated as the percentage of the overall mass in excess of the estimated mass in the excluded region above the discontinuity. The confidence interval is calculated via bootstrapping.

Figure 2·5: Bunching by Census Status



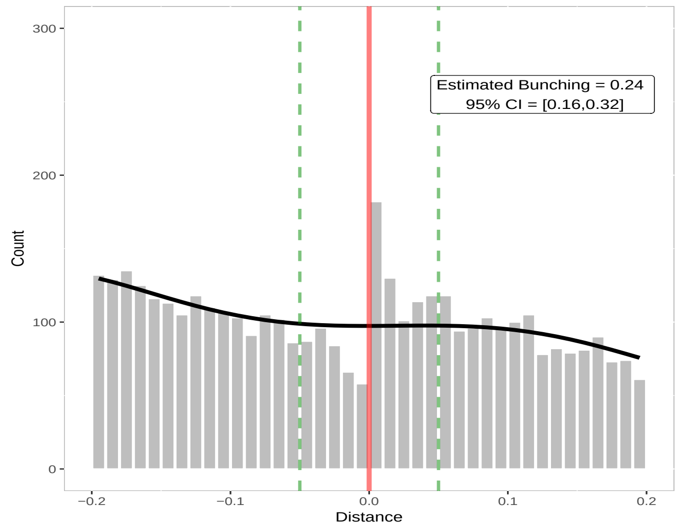
(a) Bunching in Census Years



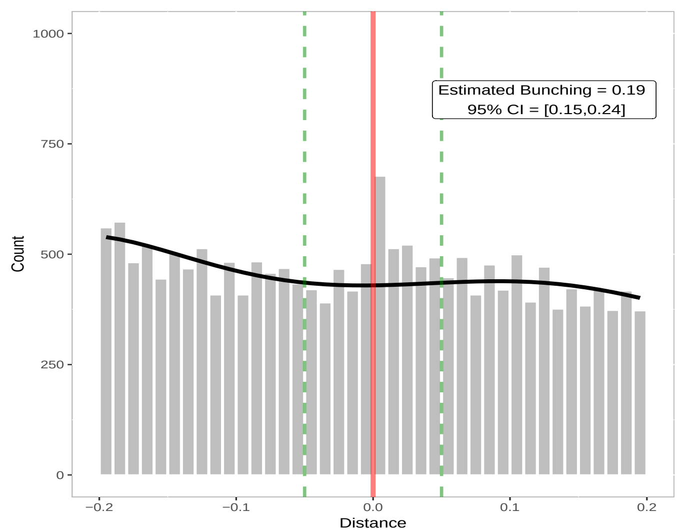
(b) Bunching in Non Census Years

Note: The unit of observation is the municipality. Each dot displays the number of municipalities in a bin of 1% of distance to the nearest discontinuity. The estimated counterfactual distribution is calculated with a fifth degree polynomial of distance to the discontinuity. The dashed green lines denote the excluded region for the counterfactual discontinuity. Estimated bunching is calculated as the percentage of the overall mass in excess of the estimated mass in the excluded region above the discontinuity. The confidence interval is calculated via bootstrapping.

Figure 2·6: Bunching by Decentralization Level



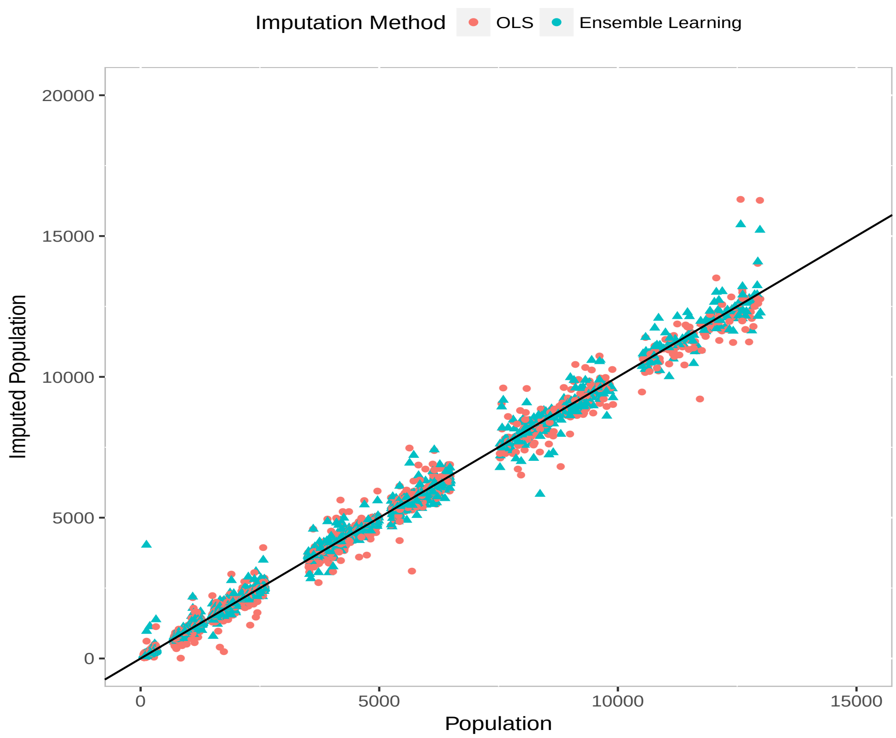
(a) Increase in Decentralization



(b) No Increase in Decentralization

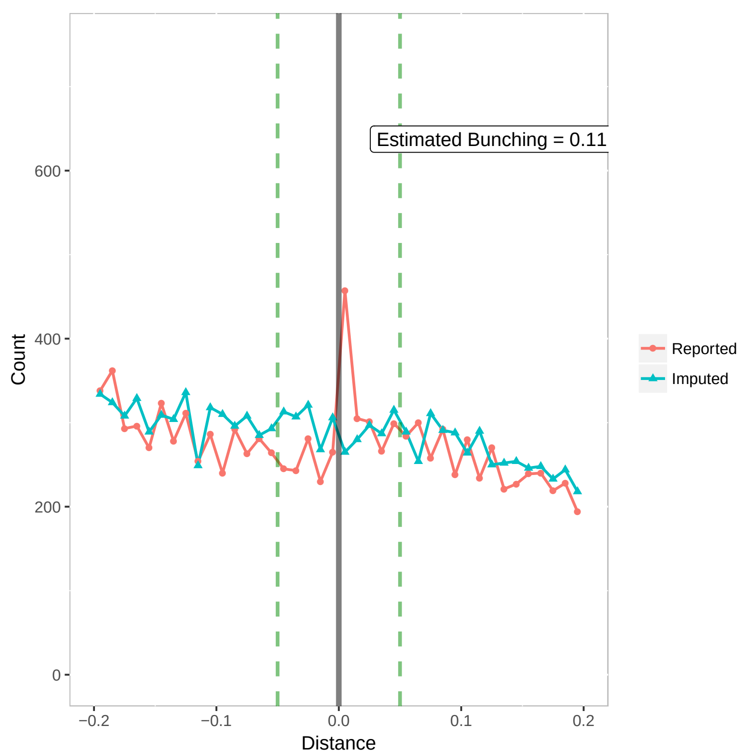
Note: The unit of observation is the municipality. Each dot displays the number of municipalities in a bin of 1% of distance to the nearest discontinuity. The estimated counterfactual distribution is calculated with a fifth degree polinomial of distance to the discontinuity. The dashed green lines denote the excluded region for the counterfactual discontinuity. Estimated bunching is calculated as the percentage of the overall mass in excess of the estimated mass in the excluded region above the discontinuity. The confidence interval is calculated via bootstrapping.

Figure 2·7: Imputation Results



Note: The graph is based on observations in the validation set. The validation set is composed by one third of all observations that are not in the excluded region. The horizontal axis represents population. The vertical one displays the imputed population. The 45 degree line that would represent perfect prediction is also displayed.

Figure 2·8: Imputed versus Reported Population



Note: The unit of observation is the municipality. Each dot displays the number of municipalities in a bin of 1% of distance to the nearest discontinuity. The dashed green lines denote the excluded regions for the imputation method. Estimated bunching is calculated as the percentage of the overall reported mass in excess of the imputed mass in the excluded region above the discontinuity.

Table 2.1: Council size rule

Population	Representatives
≤ 100	3
[101 , 250]	5
[251 , 1,000]	7
[1,001 , 2000]	9
[2,001 , 5,000]	11
[5,001 , 10,000]	13
[10,001 , 20,000]	17
[20,001 , 50,000]	21
[50,001 , 100,000]	25

Note: This table shows the council size in municipal elections as a function of the population in the municipality, in accordance to Law 5/1985.

Table 2.2: Evolution of Population in Spanish municipalities

	1987	**1991**	**1995**	**1999**	**2003**	**2007**	**2011**
Mean	4,778	4,951	4,992	4,964	5,176	5,526	5,810
Q1	230	225	209	198	190	184	178
Q2	628	628	607	587	577	576	587
Q3	2,203	2,236	2,190	2,190	2,212	2,360	2,493
99 perc	63,052	65,082	66,055	64,439	70,893	77,601	82,489

Note: The sample is composed by all Spanish municipalities that report electoral results for each election year. Q1, Q2 and Q3 denote each of the quartiles of the population distribution.

Table 2.3: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Number Parties	55,901	2.91	1.52	1	25
Pr Majority	55,901	0.81	0.39	0	1
Turnout	55,901	0.62	0.15	0.002	4.33
Seats by 1st Party	55,901	5.13	2.29	1	34
% Seats by 1st Party	55,901	0.66	0.19	0.18	1.00
% Difference in Votes	49,675	0.26	0.21	0.00	1.00
One Party Council	55,901	0.11	0.31	0	1
Two Party Council	55,901	0.36	0.48	0	1
Population	55,901	5,220.51	45,275.74	4	3,273,049

Note: The sample is composed by all Spanish municipalities that report electoral results for each election year for each election year in the period 1987-2011. Number of parties denotes the number of parties that obtained representation in the council. The difference in votes is defined as the share of votes casted for the most voted party minus the share of votes casted for the second most voted party. The variables one and two party council denote the proportion of municipalities in which only one or two parties obtained seats in the council.

Table 2.4: Summary of Public Spending and Revenue Subcategories

% of Total Spending	**Mean**	**SD**	**% of Total Revenue**	**Mean**	**SD**
Personnel	0.224	0.109	Direct Taxes	0.185	0.094
Services	0.277	0.101	Indirect Taxes	0.0312	0.038
Financial Expenses	0.044	0.046	Fees	0.148	0.087
Infrastructure	0.377	0.178	Revenue from Assets	0.061	0.094
Transfers to Public Entities	0.050	0.046	Government Transfers	0.261	0.125
Transfers to Private Entities	0.026	0.475	Debt	0.045	0.061
(N=21,461)			Capital Transfers	0.237	0.175

Note: The sample is composed by all Spanish municipalities for which public spending data is available in the period 1987-2009 and was matched to the electoral results data.

Table 2.5: Likelihood of Bunching by Census Year

	Dependent variable:			
	Just Above Threshold (dist < 0.05)			
	(1)	(2)	(3)	(4)
Census Year	-0.011^{***}	-0.007^{*}	-0.021^{***}	-0.021^{***}
	(0.003)	(0.004)	(0.005)	(0.008)
Log Population	0.002^{*}	0.021^{**}	0.002^{*}	0.020^{*}
	(0.001)	(0.011)	(0.001)	(0.011)
Distance	0.145^{***}	0.051^{***}	0.145^{***}	0.045^{***}
	(0.006)	(0.014)	(0.006)	(0.014)
Municipality FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Observations	33,131	33,131	33,131	33,131

Notes: Dependent variable is a dummy that takes a value of one if the municipality is right above the threshold, defined as above by less than 5% of the discontinuity. The sample is composed by all election years from 1987 to 2011. Census Year is a dummy variable that takes the value of one if the election year was also a Census Year (this happens both in 1991 and 2001). Robus standard errors in parenthesese clustered at the municipality level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 2.6: Bunching on Election Years

Dependent variable: Above Threshold	(1)	(2)	(3)
Election Year	0.020*	0.021**	0.062***
	(1.912)	(2.063)	(3.043)
Log population		0.020***	0.021***
		(8.699)	(8.638)
Bandwidth	< 1%	< 1%	< 1%
Year FE			✓
N	2900	2900	2900

Notes: Dependent variable is a dummy that takes a value of one if the municipality is right above the threshold, defined as above by less than 1% of the discontinuity. Election Year is a dummy variable that takes a value of one in years where municipal elections were held. Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 2.7: Bunching on Election Years (b)

Dependent variable: Above Threshold	(1)	(2)	(3)
Election Year	0.005**	0.005**	0.014***
	(2.048)	(2.075)	(2.983)
Log population		0.004***	0.004***
		(7.611)	(7.602)
Bandwidth	< 5%	< 5%	< 5%
Year FE			✓
N	12809	12809	12809

Notes: Dependent variable is a dummy that takes a value of one if the municipality is right above the threshold, defined as above by less than 5% of the discontinuity. Election Year is a dummy variable that takes a value of one in years where municipal elections were held. Robust standard errors in parentheses clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 2.8: Revenue from transfers (per capita)

Above Threshold	85.5	30.9	-42.3	17.9	37.1
	(0.53)	(0.26)	(-0.42)	(0.21)	(0.47)
Above Threshold 1999	-49.2	-76.8***	-52.4**	-42.9*	-40.0*
	(-1.34)	(-2.62)	(-1.98)	(-1.81)	(-1.72)
R^2	470.1	461.5	462.6	456.2	456.6
Y Mean	0.16	0.15	0.15	0.15	0.14
N	718	1331	1958	2538	3134

Notes: Each column corresponds to a different bandwidth: 5, 10, 15, 20 and 25%. Controls include: quadratic of distance to the discontinuity, interacted with a dummy for above the discontinuity, population and discontinuity fixed effects. t-stats in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Sample is for the year 2001.

Table 2.9: Comparison Imputation Methods

	MSE Tset	MSE Vset	N Tset	N Vset
OLS Population	370	335	10238	5119
OLS Polynomial Population	299	297	10238	5119
OLS Pop and NL	367	332	10238	5119
OLS Polynomial Pop and NL	294	294	10238	5119
RF Pop	452	594	10238	5119
RF Pop and NL	283	597	10238	5119
Ensemble Learning 1	266	289	10238	5119
Ensemble Learning 2	266	291	10238	5119
Ensemble Learning 3	230	405	10238	5119

Notes: Each row corresponds to a different imputation method. The imputation methods are estimated in the training set and the mean square error (MSE) is calculated in the validation set. Observations in the excluded region are not included in neither the training nor the validation set.

Table 2.10: Determinants of Misreporting

	Dependent variable:			
	Pop - $\hat{Pop}$		Over-reporter	
	All Sample	Just Above	All Sample	Just Above
	(1)	(2)	(3)	(4)
Turnout (t-1)	0.075***	0.039***	0.197***	0.089**
	(0.021)	(0.011)	(0.064)	(0.044)
Majority (t-1)	0.001	0.011	−0.021	0.023
	(0.003)	(0.007)	(0.014)	(0.041)
% Seats 1st Party (t-1)	−0.023	−0.026	0.029	0.011
	(0.019)	(0.056)	(0.074)	(0.242)
Diff Votes 1st 2nd Party (t-1)	0.003	0.002	−0.049	−0.009
	(0.014)	(0.030)	(0.044)	(0.142)
Number Parties (t-1)	−0.007***	−0.005**	−0.020***	−0.015
	(0.001)	(0.002)	(0.005)	(0.016)
Year FE	Yes	Yes	Yes	Yes
Discontinuity FE	Yes	Yes	Yes	Yes
Observations	15,641	1,608	15,641	1,608
R^2	0.222	0.300	0.222	0.300

Notes: In the first two columns the dependent variable is the difference between reported and predicted population. In the last two, the dependent variable is a dummy that takes a value of one if the reported population is above the predicted one. Odd columns run this regressions for all municipalities, even columns for municipalities above the nearest cutoff and with a distance of less than 5%. All regressions include controls for year and cutoff fixed effects. Standard errors are clustered at the municipality level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Chapter 3

The Long Term Impact of WWII Casualties on U.S. Political Preferences

3.1 Introduction

The quality of institutions plays a key role in economic development (Acemoglu, Johnson, and Robinson, 2001, 2002). However, what determines institutions, and how these are shaped remains an open question. In democratic countries, the main avenue for citizens to express their public preferences and affect institutions is through voting and electing public officials. In this paper, we study one of the main military conflicts in history, World War II, and focus on how this event may have shaped American institutions through the changes in political preferences of the electorate in the United States.

World War II is the armed conflict with the highest death toll in history, the estimates ranging between 60-75 million people. As such, it has shaped culture, institutions and policy world wide for the last 80 years. The United States suffered more than 400,000 military deaths during this conflict. In this paper, we study whether communities that suffered more fatalities changed their political preferences as a response to this “shock".

There are several mechanisms studied in the literature that could explain such an effect. First, communities that suffered heavier losses during the conflict may have learned about the true “costs of war". As such, they may have taken a more anti-war stance later on, resulting in higher support for isolationist candidates. Second, it could also be possible that these communities developed resentment towards the president and the political party that engaged the United States in this conflict. Both these avenues have been explored in the literature by a theory called the “proximate casualties theory" (Gartner, 2008; Kriner and Shen, 2014). This hypothesis states that popular support for American wars is undermined more from deaths of American personnel from nearby areas more than by overall fatalities. Previously, this was understood as a difference in exposure to information about war costs. This in turn leads to lower support for incumbent politicians, but most studies suggest that these effects dissipate either in a matter of weeks (Hayes and Myers, 2009) or in a matter of months (Gartner, Segura, and Barratt, 2004). Evidence of this theory after the Iraq war is explored by Karol and Miguel (2007), who find a negative relation between higher fatalities and Bush’s vote share even after controlling for observable characteristics. Althaus, Bramlett, and Gimpel (2012) find that local deaths have a much larger negative influence on public support for the Iraq War than cumulative national deaths, that the negative effect is larger among people who are inattentive to local news, and these local fatality effects decay substantially over time. Related to this literature, Loewen and Rubenson (2010) study whether voters punish incumbents for the costs of war in the context of Canada. In particular, they test whether in districts with higher war deaths there was less support for incumbents, and they don’t find evidence to back that hypothesis. They do find that support for the Conservative Party increases in districts with higher war deaths.

In order to study the relationship between fatalities and these attitudes, we collect individual level data about United States mobilization and fatality rates at the county level, and study how voting during presidential elections in counties that were heavily affected by the war evolved after the conflict ended. In this paper, we focus on the long term impact of these fatalities, and study whether the "proximate casualties hypothesis" shifted U.S. political preferences after counties suffered heavy losses from this conflict.

Some of our early findings seem to suggest that when we use fatalities to construct a proxy for the costs of military conflicts, fatalities during World War II may have had a deeper impact than what the proximate casualties theory would predict. Some of our specifications find that counties that were deeply affected by World War II, as measured in a higher rate of enlisted soldiers that died, show lower support for Republican presidential candidates in elections between the 1950s to the present day. We test the robustness of these results since this finding would contradict the "proximate casualties theory" in a couple of ways: first, Democrats were the party in power when the United States joined the conflict, and this would suggest a higher support for incumbents. Second, this effect seems long lasting and permanent, suggesting that the beliefs about the true costs of war could have been permanently updated. Upon further research, we find these early results to be unreliable: they are not robust to our measure of the costs of war nor to the choice of regression specification. Therefore, we fail to find enough evidence to conclude that political preferences permanently shifted in the United States as a result of fatalities suffered during World War II.

This paper would contribute to the literature that explores the deeper impacts military conflicts have had on the United States' economy and society. One key finding in this literature is the impact that mobilizing 16 million American to this conflict had on increasing female labor supply. Goldin (1991), Goldin and Olivetti (2013) Doepke, Hazan, and Maoz (2015) Bellou and Cardia (2016) Acemoglu, Autor, and Lyle (2004) Jaworski (2014) are some examples of this work. The hypothesis that this conflict also impacted the labor supply of African American workers has also been studied: Ferrara (2017) shows that the deaths of semi-skilled white workers lead to black workers upgrading their skills. Similar studies have analyzed the impact of more recent conflicts on labor supply, such as the Iraq and Afghanistan wars (Christensen, 2017). Some of these changes may have affected political preferences, and help explain the underlying relationship between fatalities and changes in voting behavior. In particular, some of the research relating World War II and labor supply hints at a change in culture. For instance, it has been argued that wives of children whose mothers worked as a result of the men mobilized during the conflict are more likely to work themselves (Fernández, Fogli, and Olivetti, 2004). It is worth noting that fatalities have also been documented to deeply affect marriage markets, as shown by (Abramitzky, Delavande, and Vasconcelos, 2011) and out of wedlock births (Bethmann and Kvasnicka, 2012). These are other avenues through which fatalities could affect long term voting behaviors.

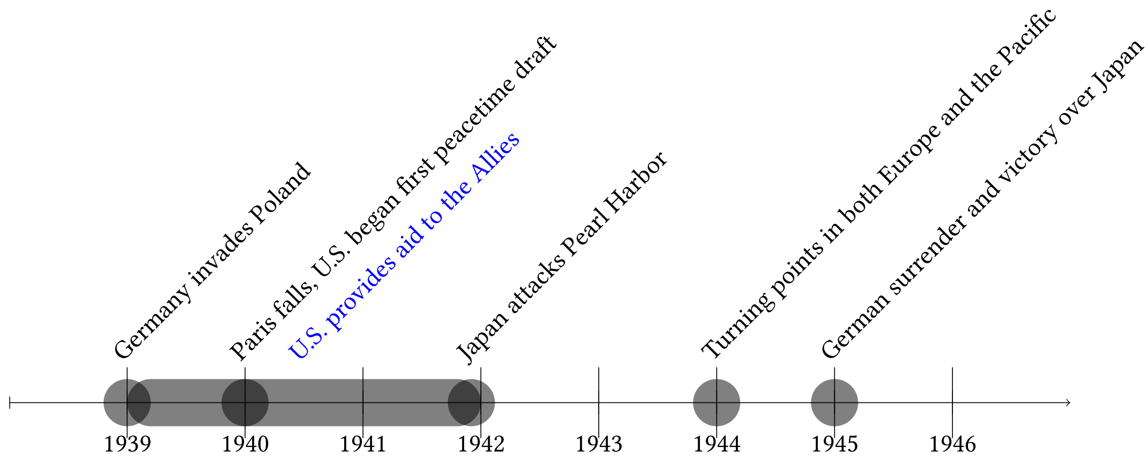
The rest of the paper is organized as follows: Section Two describes the historical background of this conflict and the following presidential campaigns. Section Three describes the data sources. Section Four explores the empirical strategy and results of this exercise, and Section Five concludes.

3.2 Historical Background

United States Involvement in World War II and the Draft

World War II was a war that involved nearly every country and major territory in the world to some extent in one of two opposing alliances, the Allies and the Axis. The war lasted nearly six years, from September 1st of 1939, when Germany invaded Poland, to 1945, when Germany signed its unconditional surrender on May 8th, and Japan on September 2nd, and the death toll of this conflict is the highest in history.

The United States involvement in this war is summarized by the time-line below. Initially reticent to be directly involved in the conflict, the president at the time, the Democrat Franklin D. Roosevelt declared its country's neutrality in 1937 through the Quarantine Speech. However, while still neutral, the United States provided aid to the Allies in the form of money, weapons and materials.



Formal entry of the United States into the war happened after the Japanese took unannounced military action against the United States by attacking the naval base of Pearl Harbor in Hawaii on December 7th, 1942. The following day, the United States declared formal war on Japan, and on December 11th, Germany and Italy declared war on the United States.

The United States deployed troops mostly in the Pacific area during 1942, in the war effort against Japan. U.S. forces in Europe were deployed more significantly during the period 1943-1945 working together with British Forces, particularly in Italy. At this time, the United States forces represented about a third of the Allied forces deployed.

By 1944, the tides of the conflict had turned in favor of the Allies both in Europe and the Pacific: the main invasion of France took place, and Japanese islands were captured by U.S. troops that year. Finally, Berlin fell to the Soviet army in 1945, and Germany surrendered.

Particularly relevant to this project is the process by which United States troops were enlisted during World War II. On September 16th, 1940, the United States enacted the first peace time draft through the Selective Training and Service Act. This law required all male citizens between the ages of 21 and 35 to register for the draft. The Selective Service System, an independent agency originally created during the first World War was reestablished and in charged with the responsibility of identifying possible enlistees. This act also limited military service to 12 months unless Congress deemed an extension necessary.

Once registered, draftees were given a number between 1 and 7,836 by one of the 6,443 local draft boards. A national lottery was held in Washington, where each possible number was place in a capsule, and all capsules were put in a bowl. The first time this lottery took place in 1940, it was Secretary of War Henry Stimson, while blindfolded, who selected numbers out of the bowl. Those registered men that held one of the numbers selected from the bowl were brought to the local draft boards to be considered for service.

Local draft boards then placed men into one of four categories. Class I were those ready for military service. Class II were those who obtained a work deferment. These were mostly farmers and workers in industries key to the war efforts. Class III were those male citizens with dependents. Finally, those placed Class IV were citizens with some type of physical, mental or moral incapacity. It is important to note that the Selective Training and Service Act prohibited racial discrimination. However, the Army's policy was to segregate training facilities and military units, therefore many black men were excluded from service, or even when they served, they were less likely to be in the front line of battle.

The Selective Training and Service Act was amended in several key ways during the course of the conflict. The length of military service was extended by 18 months in 1941. After the attack on Pearl Harbor: the age bandwidth was expanded to all male citizens between the ages of 18 and 44, the process became more centralized, reducing the power of local draft boards, and the length of service extended to the duration of the war plus six months. More importantly, in December 1942 voluntary enlistments for all men between the ages of 18 and 37 were suspended. We observe this pattern clearly in our data in Figure 3·1. In this Figure the number of enlistees are plotted for each year, divided by those with dependents and those without. For the years 1940, 1941 and 1942 a big proportion of enlistees had no dependents, however, for the reminder of the conflict, the proportion of enlistees with dependents is almost the same as those without, compelling evidence for this change in policy. This was the policy that determined enlistment in the United States Army for the duration of the conflict as well as the Navy and Marine Corps starting in 1943.

Short History of Post War United States Presidential Elections

This paper tests the hypothesis that fatalities during World War II affected U.S. political preferences. This hypothesis hinges on the fact that, on average, Republican candidates represent different preferred policies over foreign policy, interventionism and defense spending, areas we believe are more likely to be impacted by fatalities during World War II. Therefore, we proceed to discuss whether the Republican party actually represented substantially different policies in post war presidential elections.

Dwight Eisenhower was the presidential candidate in both 1952 and 1956. He won the presidential election in landslides, defeating Adlai Stevenson in both occasions, obtaining 55.2% and 57.4% of the popular vote, respectively. This marked the first time since 1932 that the United States had a Republican president. One of the main campaign issues in the 1952 election was the handling of the Korean war by president Truman, the Democratic incumbent, as well as the lack of foreign intervention in Latin America, failing to stop the expansion of Communism. In 1956, Stevenson ran a campaign in which he called for an increase in government spending on social programs and a decrease in military spending. This seems to support that the idea of Republicans taking a more interventionist approach than Democrats during this decade, and it being a salient point during both campaigns.

The Democrats took control of the oval office back in 1960 and 1964. In 1960, Richard Nixon, who served as Vice President of the United States on both of Eisenhower's terms, lost to John F. Kennedy. During this election, once again, the Cold War was the main issue to dominate the campaign. Kennedy challenged Eisenhower on the administration's defense spending policies, as well as promoting a push for the "Space Race" and an increased spending in research and technology to catch up with the Soviet Union, who had successfully launched the first man-made satellite to orbit Earth in 1957. In 1964, the election focused around civil rights issues and Barry Goldwater lost to incumbent Democratic President Lyndon B. Johnson, who had succeeded Kennedy after his assassination. Goldwater was labeled by Democrats as an extremist and prone to the use of the nuclear weapons, as evidenced by a famous television ad, that aired only once but had tremendous popular impact, the "Daisy Ad". In this ad, a little girl plucked daisy petals in a meadow. When she reached nine, her voice was supplanted by a nuclear countdown. When the countdown reached zero, the mushroom shape associated with a nuclear blast was shown. Afterwards, a second voice over asked viewers to vote for Johnson.

The Republicans won again in 1968 and 1972, when Richard Nixon defeated the Democrats Hubert Humphrey and George McGovern. While Nixon vowed to put an end the Vietnam war during the 1968 campaign, he proceeded to bomb Cambodia and order incursions in Laos. The peace accords were signed in 1973. Nixon is said to have appealed to the "silent majority" or Americans, those who had a disliked for the 1960s counter culture movement and the anti war protesters. At the same time, he instituted the "Nixon Doctrine", which meant American soldiers would not be sent to conflicts involving American allies unless a nuclear threat was involved. He also became the first American president to visit the Republic of China and he enacted a policy of détente with the Soviet Union, lowering the threat of an actual war.

After Nixon's resignation in 1973, Gerald Ford was sworn president on the United States. He ran for reelection in 1976, and lost to Democrat Jimmy Carter. During Gerald Ford's time in office, he continued the détente policy put forth by his predecessor. Perhaps most indicative of their attitudes towards war and war protesters, was the fact that Jimmy Carter promised a pardon for religious refugees and Vietnam War dodgers. In the eighties, the Republicans took back the Oval Office, as Ronald Reagan defeated incumbent Jimmy Carter and Walter Mondale in 1980 and 1984, with future president George Bush as his running mate. One of the main campaign promises that Reagan ran under was the promise of restoring America's military strength. He ran under a similar platform in 1984, whereas his opponent, Walter Mondale, proposed a nuclear freeze.

In 1988 and 1992, George H.W. Bush was the Republican presidential candidate. In 1988, he defeated Michael Dukakis, who was labeled by Republicans as an extreme liberal. Bush campaign focused on maintaining Reagan's policies, without radical changes. With Bush as president, the United States engaged, together with its allies, in the Gulf War between August of 1990 and February of 1991, as a response to Iraq's invasion of Kuwait. At the end of the conflict, Bush received the highest job approval rating any president had had since the 1930s. In 1992, he lost the reelection to Democrat Bill Clinton. Clinton was attacked for his lack of foreign policy experience, but the economy was in recession, and after the end of the Cold War and the dissolution of the Soviet Union, foreign policy was regarded of secondary importance during the campaign. In 1996, Bill Clinton secured reelection by defeating Republican candidate Senator Bob Dole obtaining 379 electoral votes, compared to the 159 obtained by Dole. Dole ran under a platform of lowering taxes and supply side economics, but the booming economy made his chances of victory slim.

George W. Bush was elected president in 2000, after defeating Al Gore, and reelected in 2004, when he faced John Kerry. While the first campaign was focused mostly on domestic issues, after the attacks of September 11, foreign policy was the main driver of the 2004 campaign. Kerry criticized Bush's strategy leading and during the Iraq War. Bush sought to present himself as the candidate that would be tough on terrorism, whereas Kerry's supporters were more critical of the United States' intervention in Iraq. In 2008 and 2012, Democrat Barack Obama defeated Republicans John McCain and Mitt Romney. Particularly relevant to this paper is the fact that, in 2008, campaigning focused heavily on the Iraq War. McCain supported the war, while Obama strongly opposed it. The 2008 was more heavily focused on domestic issues. Finally, in 2016, Republican Donald Trump won the general election over Democrat Hillary Clinton.

In conclusion, barring some exception, such as the 1960 campaign, it seems that Republican candidates have had a higher level of support for interventionist policies, as well as being more in favor of increasing defense spending and engaging in military actions.

3.3 Data and Descriptive Statistics

In order to assess the effect of World War II on United States political preferences, we combine three different data sets. The first one are fatalities and missing in action records from the National Archives for both the Army and the Navy at the county level. These documents that we digitize contain the name and rank of each person in the United States military that lost their life during this conflict or were missing in action, the county of origin and whether they served for the Army or the Navy. These documents contain the county for each state, and the name of the soldier. Next to the name there are two set of acronyms. The first documents the rank of the soldier (Sargent, Private). The second state what was his fate (killed in action, died of wounds, died non in battle, etc). Given these documents we manually create a data set at the county level with the number of fatalities in each United States county for both the Navy and the Army. Further data could be extracted from these documents, such as fatalities by rank, or type of fatalities, but given that these are scanned documents that are not digitized, that was not feasible at this time.

We combine fatalities data with World War II Army Enlistment Records documenting the period 1938 to 1946 from the National Archives. These data set contains individual level enlistment data, with county of origin, educational attainment and whether or not the enlisted individual had dependents. This data allows us to create a variable that measures the number of enlistees from each county. Using County Census Data from 1940 from ICPSR, we create fatality to enlistment, fatality to population, and enlistment to population variables. When we refer to population we use the closest census year data available, 1940. These will be the main variable we use to measure the shock that counties suffered following World War II. One shortcoming of our study is the non availability of Navy enlistment data. Therefore, for robustness, we will perform the same analysis using only Army fatalities.

Descriptive statistics for this data are shown in Table 3.1. The average county in the U.S. suffered 113 fatalities, mostly from the Army, 94 on average, with respect to 19 from the Navy. The population of the average U.S. county in 1940 was 40,608, and around 2,300 men were enlisted per county. This Table also shows the number of men age 15-44 in 1940. This variable represents the population most likely to be drafted during the conflict. Out of these men, 24% of them are drafted into the Army. The fatality to enlistment ratio is on average 6%, meaning that the average county loses 6% of their enlisted men during the conflict. If we focus exclusively on Army fatalities, the average drops to 5%. We also show the percentage of all 1940 men that lose their life due to the war.

One possible concern with the fatality rate is that maybe all counties suffered similarly from World War II conditional on enlistment rates. This would mean that we would not have enough variation to pick up on any changes in culture due to fatalities. We address this concern in Figure 3·2. This histogram shows the distribution of the fatality rate in U.S. counties. Most counties lost 5-6% of all enlisted men, but some lost over 10-12%.

We further study the geographic distribution of fatality rates and enlistment rates in a couple of maps. In these heat maps, brighter colors denote higher values. In Figure 3·3, percentiles of the fatality rate are displayed. Geographical patterns are clear, with the Midwest of the United States having higher fatality rates on average, whereas counties in Florida, Georgia and South Carolina seem to have lower fatality rates on average. This will prompt the use of state fixed effects in all of specifications, so that we can exploit variation across counties within a state.

In the second figure, we show the proportion of men that were enlisted to fight during the conflict (Figure 3·4). What we can observe in this graph is that although counties in the Midwest suffered higher fatalities rates, they had fewer men serving during World War II. This is consistent with the fact that there is a higher share of farmers on these parts of the United States, and farmers were given deferrals in order to keep the domestic food supply. Figure 3·5 shows the percentiles of the proportion of 1940 men that died in World War II. Here, a different trend appears compared to the fatality rate: counties in the West suffered heavier losses, whereas counties near the South still seem to have lower than average fatalities. These lower casualties in the South seem to concentrate around the area known as "the cotton belt", counties in the Southern region of the United States known for their cotton production between the 18th and 20th century. Therefore, these are counties with a higher proportion African American, consistent with the hypothesis that African American were less likely to serve and die during the conflict.

Finally, to attest the impact that fatalities had on culture, we use county level voting records in presidential elections. This data comes from Leip's Atlas of Presidential Elections, and we focus on elections from 1900 to 2016. This data includes turnout records, as well as the vote share for the Republican and the Democratic candidate in each county. Figure 3·6. This Figure shows the evolution of turnout, Democratic and Republican vote shares over time, for all elections between 1944 and 2016.

3.4 Empirical Strategy and Results

The empirical strategy we use in this paper is an event study analysis. We want to test whether political preferences changed as a response to fatalities suffered during World War II. Therefore, our baseline specification is the following:

$$Y_{ct} = \sum_{t \neq 1940} \beta_t \mathbb{1}\{\text{Year = t}\} x \text{Measure of Fatalities}_c + \sigma_{st} + \sigma_c + \varepsilon_{ct} \tag{3.1}$$

The unit of observation is county c in election year t. As our measure of political preferences we use the percentage of all votes in a county for the Republican candidate in each presidential election (in logs). The main reason for using presidential vote share is that commanders in chief have a lot of influence over foreign policy, so if we expect attitudes towards war to change after being affected by the fatalities suffered during World War II, presidential elections are a better measure than state, local or midterm elections. Furthermore, the American presidential election system has been dominated by only two parties since the very beginning, making possible the comparison of voting preferences over time.

Given the geographic trends that fatalities and enlistment showed previously, a full set of county (γ_c) fixed effects as well as state times election fixed effects ($\sum_{t \neq 1940} \gamma_t \mathbb{1}(Year = t) x \sigma_s$) are included in the main specifications. Standard errors are clustered at the county level, allowing error dependence within counties over time.

The main explanatory variable is some measure of the shock caused by the fatalities in these counties. Through fatalities, we are trying to measure how costly citizens in these counties understand war to be. There are plausible stories for two different measures.

The first is what we call the fatality over enlistment ratio, the proportion men died in county c among those that were enlisted (in log scale). The intuition behind this measure is that perhaps citizens estimate the cost of war by the survival rate of soldiers in their communities. Let's consider two towns that each send 100 men to serve during World War II. If town A loses half of these man, but town B loses only one tenth, it seems plausible that citizens in town A will update their beliefs about the costs of war differently than those in town B. This variable has a number of advantages over the proportion of the total population that died during the war. This alternative suffers from some known biases by not controlling for enlistment rates. Enlistment is not random: counties with a higher share of farmers are likely to mobilize less troops to the conflict. Counties with a higher share of black population are also less likely to be mobilized, since at the time, black men were less likely to serve. One concern to our specification is that, even when black men did serve, they were less likely to be put on the front lines due to unsegregated regiments, and they were more likely to serve as cooks, janitors, etc. These trends were shown to be present in the maps of fatalities.

By creating as our main explanatory variable the proportion of enlistees that died during the war, the main identification assumption is that a county having a high or a low fatality rate, given the number of troops sent to the conflict is as good as random. Although this is a non testable assumption, we can present some compelling evidence for our identification strategy in the form of pre-trends. We run the regression for election years between 1900 and 1940 as well as post 1940. If somehow our measure of fatalities during the war affects voting prior to the conflict, then that would be evidence that some non random characteristic in our variable is correlated with voting patterns. The results show that there are no pretends impact of fatalities on voting for the Republican candidate on presidential elections. Finally, all our regressions include a full set of state times year fixed, as well as counties fixed effects, and are weighted by 1940 population.

Therefore, Figure 3·7 shows the estimates from the following specification:

$$Y_{ct} = \sum_{t \neq 1940} \beta_t \mathbb{1}\{\text{Year} = \text{t}\} x \frac{\text{Fatalities}_c}{\text{Enlistees}_c} + \sigma_{st} + \sigma_c + \varepsilon_{ct} \tag{3.2}$$

The first panel in Figure 3·7 displays the estimates of our specification. Each dot in the Figure is the estimated effect of the fatality over enlistment ratio interacted with an election year dummy. The omitted category is the election year 1940. Prior to World War II, none of the coefficients are significantly different from zero. This is reassuring of our identification strategy. Post World War II, the estimates are negative, and they start to become significantly different from zero after 1950. Note, this is the first time that American Army General Dwight Eisenhower ran for president, and the first time post World War II that foreign policy was the main issue during the campaign. The estimates remain negative for the rest of the sample period, indicating that this effect is long lasting and does not dissipate over time. For the period 1950-2016, the impact of this effect is fairly large: the estimate is, on average, -0.5. Since all regressions are run in logarithmic scale, this implies that an increase of one percent in fatality rate is associated with a decrease of half a percentage point in the support of the Republican presidential candidate.

A first concern with this specification is the fact that we do not that we do not observe Navy enlistment rates. Therefore, we run the same specification using exclusively Army fatalities to compute the fatality rate. The results of this regression are shown in the second panel of Figure 3·7. The results show very similar estimates, with the exception that the estimated impact becomes non significant in the last three elections. Once again, there is no significant impact prior to 1944.

We run two different robustness tests on Figure 3·8. The first panel addresses the concern that fatalities may just be a function of the number of enlisted soldiers, and we are just picking up the effects of enlistment with out measures of fatalities. We address this by regressing the presidential vote share on the enlistment rate of each county (in logs):

$$Y_{ct} = \sum_{t \neq 1940} \beta_t \mathbb{1}\{\text{Year} = \text{t}\} x \frac{\text{Enlistees}_c}{\text{Population}_c} + \sigma_{st} + \sigma_c + \varepsilon_{ct} \tag{3.3}$$

The estimates are collected on the first panel Figure 3·8. This figure shows that the enlistment rate does not seem to have any impact on presidential vote shares over the period following World War II.

A second important consideration is the racial composition on counties prior to the War. As our maps showed, it seems that fatality over enlistment is higher in counties with a higher share of white men, and lower in those with a higher fraction of African American population. Therefore, the driver behind the estimates on Figure 3·7 could be the racial composition of counties. We address this hypothesis by controlling for the percentage of African American population interacted with election year:

$$\begin{aligned} Y_{ct} = \sum_{t \neq 1940} \beta_t \mathbb{1}\{\text{Year} = \text{t}\} x \frac{\text{Fatalities}_c}{\text{Enlistees}_c} + \\ \sum_{t \neq 1940} \alpha_t \mathbb{1}\{\text{Year} = \text{t}\} x \frac{\text{Black Pop}_c}{\text{Population}_c} + \sigma_{st} + \sigma_c + \varepsilon_{ct} \end{aligned} \tag{3.4}$$

As we can observe, even controlling for race, there is still a significantly negative impact of the fatality over enlistment ratio can on the Republican vote share. This seems compelling evidence of race not being the main driver of this negative effect.

However, a second plausible measure of the shock these counties suffered is the proportion of people in the county that died during the conflict. This is a better measure if we believe that the channel through which political preferences are affected is a more personal one: perhaps political preferences change when a citizen in these communities personally knows a soldier who lost their live in World War II, and make these citizens views of war, foreign policy and politics change. If that is the case, counties where a higher share of population lost their life have a higher proportion of people who knew a fallen soldier. We try to capture this by regressing the Republican vote share on the fatalities over population ratio, interacted with election year dummies:

$$\begin{aligned} Y_{ct} = \sum_{t \neq 1940} \beta_t \mathbb{1}\{\text{Year = t}\} x \frac{\text{Fatalities}_c}{\text{Population}_c} + \\ \sum_{t \neq 1940} \alpha_t \mathbb{1}\{\text{Year = t}\} x \frac{\text{Enlistees}_c}{\text{Population}_c} + \sigma_{st} + \sigma_c + \varepsilon_{ct} \end{aligned} \tag{3.5}$$

The spirit of this specification is similar to the first regression: we want to isolate the effect of fatalities, while comparing counties with similar enlistment rates. However, the results, shown by Figure 3·9 are quite different. There seems to be no significant impact of fatalities on voting preferences once the enlistment rate is accounted for. Therefore, it would appear that the results shown by Figure 3·7 are not particularly robust. The difference between the two panels of Figure 3·9 is whether controls for enlistment are included in the regression. However, the point estimates on the share of fatalities are extremely similar and do not seem to be affected by the inclusion of these controls.

Since the results appear to be not very robust, we show the results of several different specifications in Tables 3.2 and 3.3. In this tables, we collect the estimates from six plausible different specifications, running the following regression. This regression is a simpler version of the estimates shown in previous graphs, where the impact of the fatalities are imposed to be constant across 5 election periods (or 20 year periods), while still flexibly controlling for state times election year fixed effects. The difference across columns in Table 3.2 is the measure of fatalities. In the first column we use the number of fatalities divided by the number of enlistees. In the second column, we use fatalities over enlistment but control for the average education of enlistees, as well as the proportion of African Americans living in the county in 1940. These controls are interacted with period dummies. In the third column, the main explanatory variable is the Army fatalities over enlistment ratio. Columns four and five use the fatalities over population ratio as the measure of loss of life during World War II. The difference between these two columns is that the sixth one controls for enlistment interacted with 20 year period dummies, whereas the fifth one does not. Finally, the last column uses the enlistment rate as our main explanatory variable. Table 3.2 shows the results where both the Republican Vote Share and the measure of fatalities are calculated in logarithmic scale, whereas Table 3.3 does not use logarithmic scale.

$$\begin{aligned} Y_{ct} = \beta_{pre}\mathbb{1}\{\text{Year} < 1944\}x\text{Fatalities}_c + \\ \beta_{1944-1960}\mathbb{1}\{\text{Year} \in [1944, 1960]\}x\text{Fatalities}_c + \\ \beta_{1964-1980}\mathbb{1}\{\text{Year} \in [1964, 1980]\}x\text{Fatalities}_c + \\ \beta_{1984-2000}\mathbb{1}\{\text{Year} \in [1984, 2000]\}x\text{Fatalities}_c + \\ \beta_{2004-2016}\mathbb{1}\{\text{Year} \in [2004, 2016]\}x\text{Fatalities}_c + \\ +\sigma_{st} + \sigma_c + \varepsilon_{ct} \end{aligned} \tag{3.6}$$

Analysis of both Table 3.2 and 3.3 allows us to draw several conclusions. First, the use of logarithmic scale on both side of the regression equation does not seem to play a significant role in the results, the estimates are fairly similar across both tables. Second, as we have shown previously, the choice of controls, whereas it is the proportion of African American population, the education of enlistees or enlistment rates, does also not play a key role. However, the negative impact of fatalities on Republican vote share depends heavily on two factors: first, the choice of explanatory variable, and second, whether or not the regression is weighted by 1940 population. The significantly negative effect is present when the explanatory variable is the fatality over enlistment ratio, and when bigger counties are given a bigger weight in the regression analysis. In fact, when our measure of fatalities is the ratio of fatalities over population, our regressions fail the pre-trend analysis, and the effect on Republican vote is reversed and positive.

This leads us to conclude that there is not enough evidence in our analysis to determine that fatalities during World War II significantly impacted long term political preferences. Perhaps that was the case for bigger counties, but further research about the mechanism and measure of these shocks is required.

3.5 Conclusion

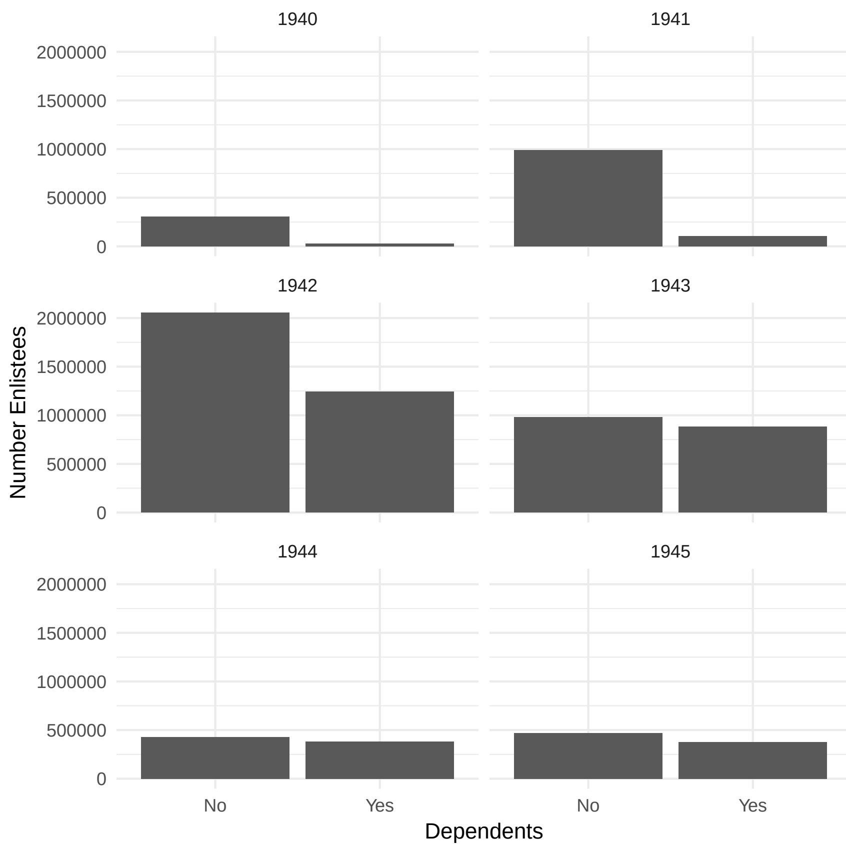
In this paper, we study the impact of World War II enlistment and fatalities on U.S. political preferences. Economists and political scientists have developed the "proximate casualties theory" which states that fatalities during military conflicts impact short term political preferences of voters, who tend to punish incumbents. The goal of this paper was to test whether counties that suffered heavier losses during World War II developed different political preferences, not only in the short run, but in the long run.

To do that, we created a new data set containing the fatalities that each county in the United States suffered during World War II. We combined this data with Army individual level enlistment data to test the impact of fatalities conditional on enlistment.

We use county level variation to analyze whether political preferences changed as a result of fatalities during the conflict in these communities. Although, we initially find some support for the hypothesis that fatalities may have decreased support for Republican presidential candidates throughout the 20th century, this result is heavily dependent upon the choice of measure of fatalities and the weighting of the regression analysis. Therefore, we fail to find sufficient evidence that political preferences shifted as a result of fatalities suffered during World War II, or the avenue through which these changes may have taken place.

Tables and Figures

Figure 3·1: Number of Enlisted Solider to the Army by Year and Dependent Status



Note: This figure contains the number of enlistees to the United Army per year, between 1940 and 1945. For each year, they are divided into whether or not they have any dependents.

Table 3.1: Descriptive Statistics 1940

	Mean	SD	Min	Max
Number of Fatalities (Army)	94	328	0	9, 736
Number of Fatalities (Navy)	19	77	0	2, 541
Total Number of Fatalities	113	399	1	11, 451
Population in 1940	40, 608	137, 920	564	4, 063, 342
Men 15-44 in 1940	9, 714	34, 401	139	1, 011, 040
Enlistment (Army)	2, 376	8, 441	15	212, 512
Fatality to Enlistment Ratio	0.06	0.03	0.001	0.43
Army Fatality to Enlistment Ratio	0.05	0.02	0	0.37
Enlistment to 1940 Men 15-44 Ratio	0.24	0.33	0.02	13.85
Fatalities to 1940 Number of Men Ratio	0.01	0.004	0.0003	0.06
Fatalities to 1940 Population Ratio	0.003	0.001	0.0001	0.02

Note: The sample is composed by all counties in the U.S. where data on fatality rates is available (2,875 out of 3,0007). The Fatality Rate is calculated as the total number of fatalities divided the number of enlisted soldiers. The enlistment rate is calculated as the number of enlisted men divided the number of men age 15-44 in 1940.

Figure 3·2: Histogram Fatality to Enlistment Ratio

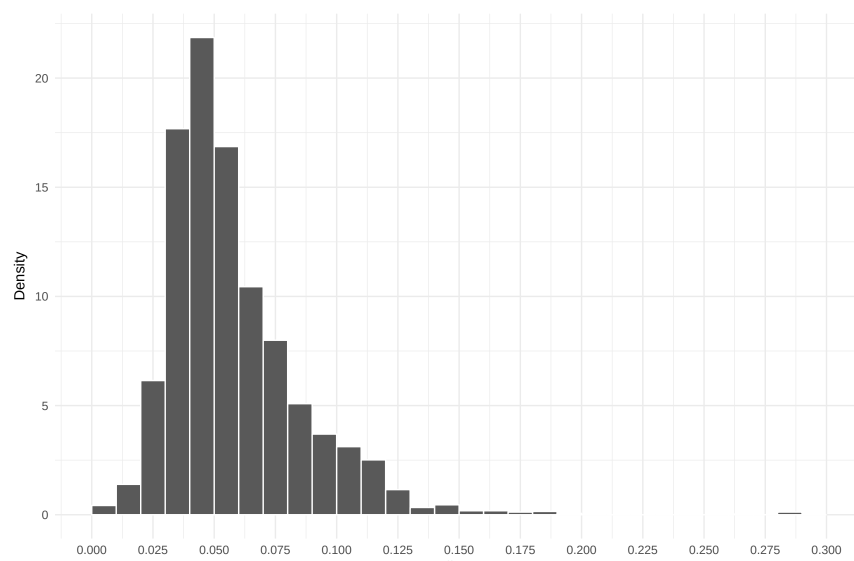
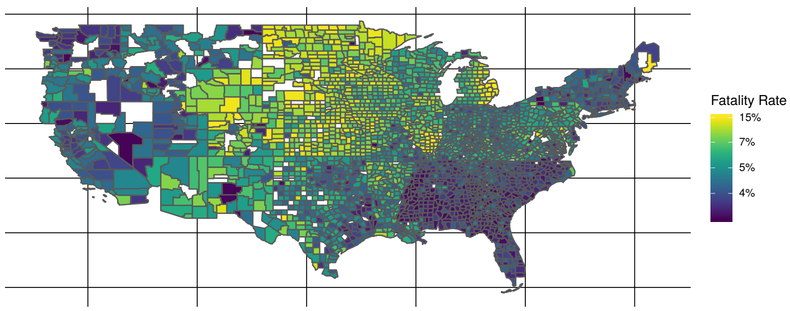
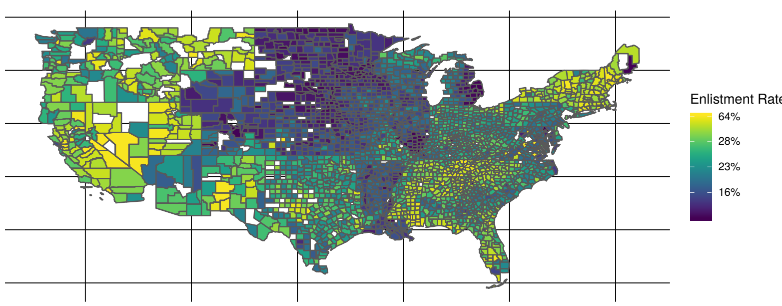


Figure 3·3: Map of the Fatality to Enlistment Ratio



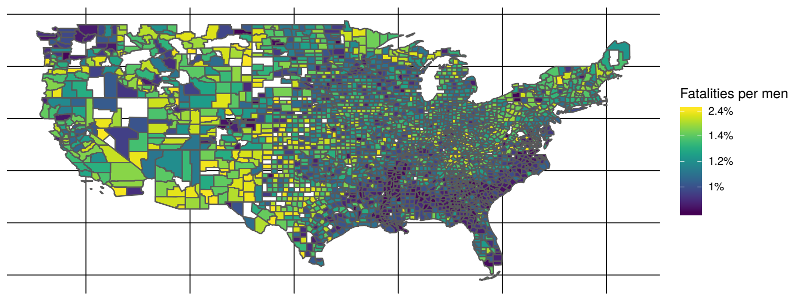
Note: The sample is composed by all counties in the U.S. where data on fatality rates is available (2,875 out of 3,0007). The variable displayed is the percentile of the total number of fatalities during World War II divided by the number of men enlisted to the army. Brightest colors represent counties that suffered a higher fatality rate.

Figure 3·4: Map of Enlistment to 1940 Men 15-44 Ratio



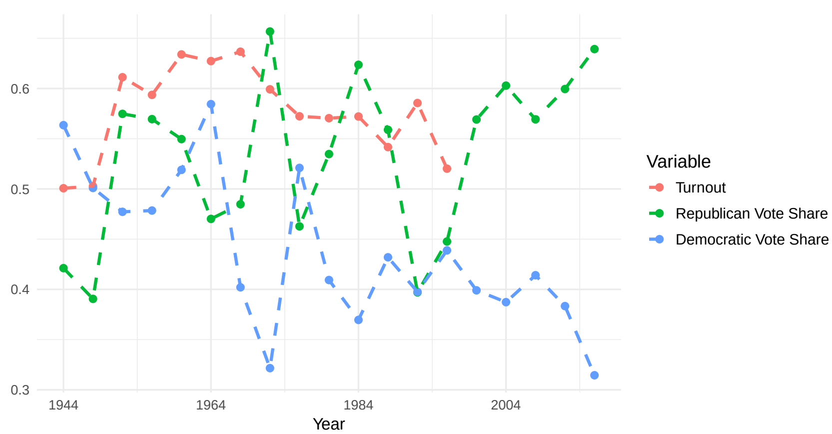
Note: Sample is composed by all counties with available data. The variable is the percentile of the total number of fatalities during World War II divided by the number of men living in that county according to the 1940 Census. War II divided by the number of men enlisted to the army.

Figure 3·5: Map of Fatalities to 1940 Men Ratio



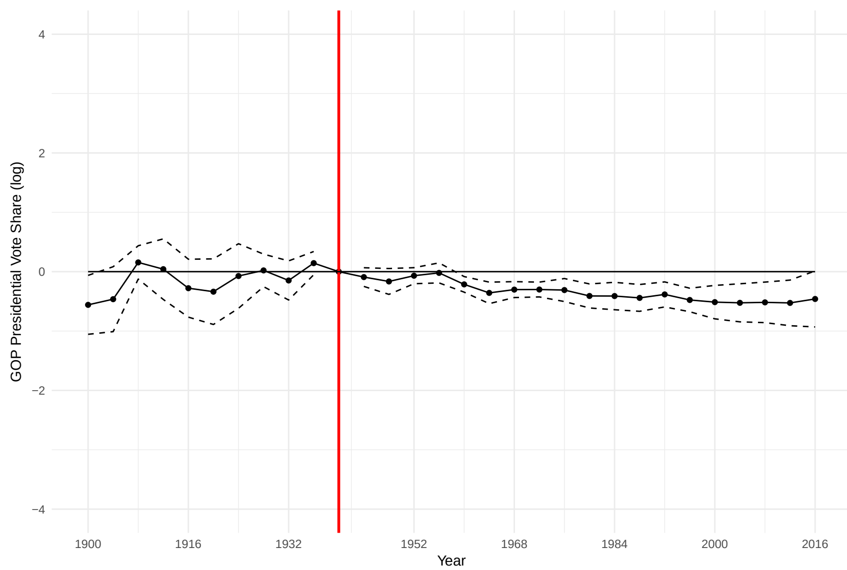
Note: Sample is composed by all counties with available data. The variable is the percentile of the total number of fatalities during World War II divided by the number of men living in that county according to the 1940 Census. War II divided by the number of men enlisted to the army.

Figure 3·6: Evolution of Voting Variables over Time

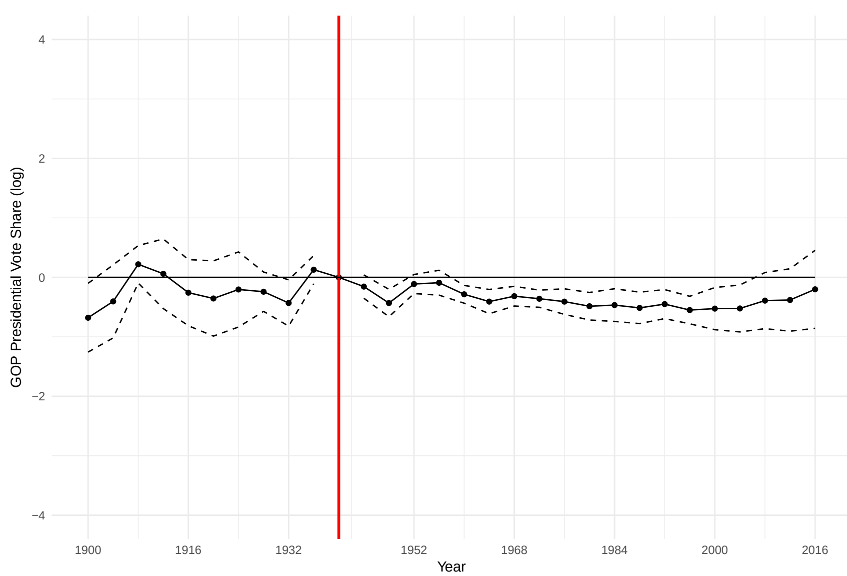


Note: This figure contains the population weighted average vote share for Republican and Democrat candidates, as well as the population weighted average turnout in presidential elections from 1944 to 2016. Turnout data is unavailable for election years 2000-2016.

Figure 3·7: Estimates of the Impact of the Fatality Rate on Republican Vote Share



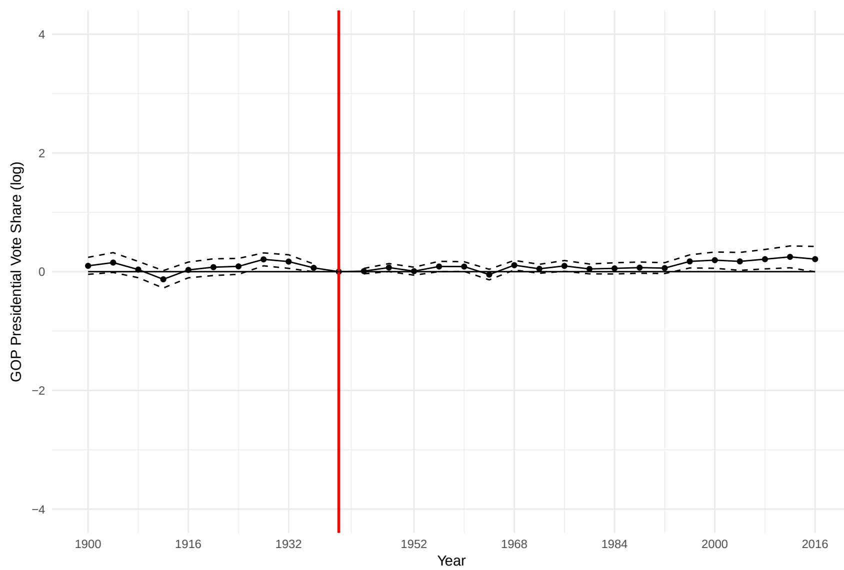
(a) Fatality Rate



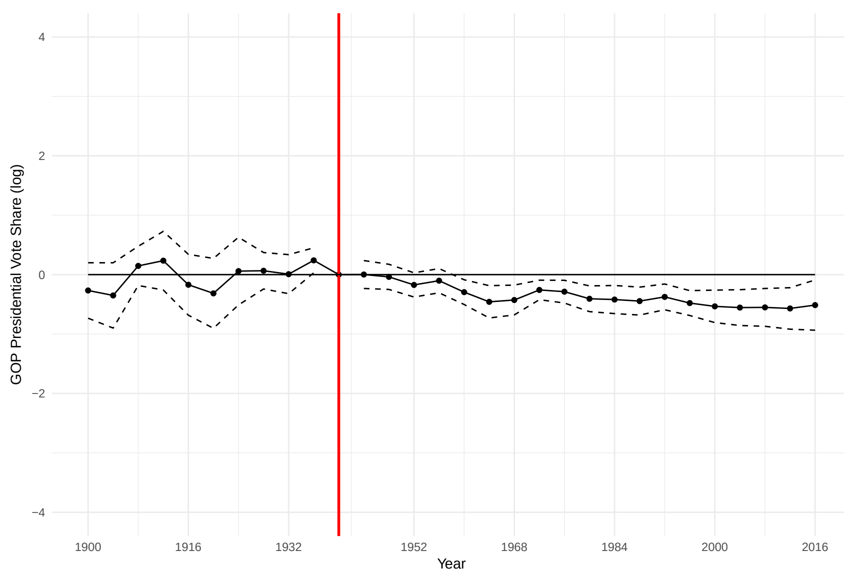
(b) Army Fatality Rate

Note: The sample is composed by all counties for which fatality data is available, and all election years between 1900 and 2016. The main explanatory variable is the number of total fatalities divided by the number of Army enlistees in the first panel and the number of total Army fatalities divided by the number of army enlistees in the second. Both the explanatory variables and the Republican vote share are calculated in logs. Estimates are obtained via Weighted OLS, the weights being 1940 population. State times year fixed effects and county fixed effects are included. The dotted line represents the 95% Confidence Interval for the point estimates based on clustered standard errors at the county level.

Figure 3·8: Robustness checks on the Estimates of the Fatality Rate on Republican Vote Share



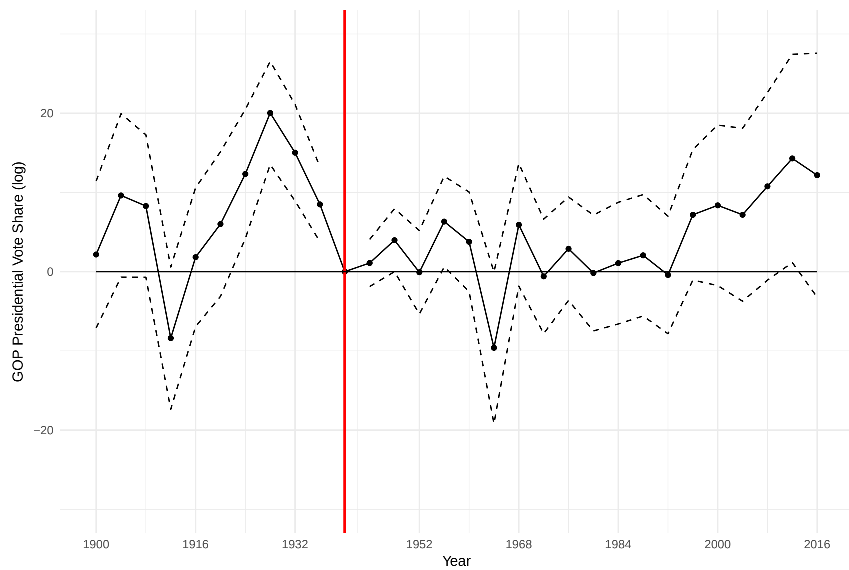
(a) Impact of the Enlistment Rate on Republican Vote Share



(b) Impact of the Fatality Rate controlling for % Blacks

Note: The sample is composed by all counties for which fatality data is available, and all election years between 1900 and 2016. The main explanatory variable is the number of total fatalities divided by the number of Army enlistees in the first panel and the number of total Army fatalities divided by the number of army enlistees in the second. Both the explanatory variables and the Republican vote share are calculated in logs. Estimates are obtained via Weighted OLS, the weights being 1940 population. State times year fixed effects and county fixed effects are included. The dotted line represents the 95% Confidence Interval for the point estimates based on clustered standard errors at the county level.

Figure 3·9: Estimates of the Impact of the Fatalities over Population on Republican Vote Share



(a) Without Controlling for Enlistment

(b) Controlling for Enlistment

Note: The sample is composed by all counties for which fatality data is available, and all election years between 1900 and 2016. The main explanatory variable is the number of total fatalities divided by the number of Army enlistees in the first panel and the number of total Army fatalities divided by the number of army enlistees in the second. Both the explanatory variables and the Republican vote share are calculated in logs. Estimates are obtained via Weighted OLS, the weights being 1940 population. State times year fixed effects and county fixed effects are included. The dotted line represents the 95% Confidence Interval for the point estimates based on clustered standard errors at the county level.

Table 3.2: Robustness of the Estimates of Fatalities on Log of Republican Vote Share

	Independent Variable (in logs):					
Panel A:	(1)	(2)	(3)	(4)	(5)	(6)
OLS Weighted	$\frac{Fatalities}{Enlistment}$	$\frac{Fatalities}{Enlistment}$	$\frac{ArmyFatalities}{Enlistment}$	$\frac{Fatalities}{Population}$	$\frac{Fatalities}{Population}$	$\frac{Enlistment}{Men15-44}$
Pre War	-0.09	-0.05	-0.14	7.46***	7.51**	-0.00
	(0.17)	(0.18)	(0.20)	(3.12)	(3.22)	(0.02)
1944-1960	-0.10*	-0.10*	-0.18***	2.59	3.03	-0.02
	(0.06)	(0.06)	(0.07)	(2.25)	(2.01)	(0.08)
1964-1980	-0.34***	-0.32**	-0.39***	-0.94	-0.39	-0.02
	(0.07)	(0.07)	(0.08)	(3.46)	(3.39)	(0.08)
1984-2000	-0.45***	-0.43***	-0.50***	3.25	3.65	0.04
	(0.11)	(0.12)	(0.13)	(3.93)	(3.90)	(0.10)
2000-2016	-0.51***	-0.46***	-0.37	10.78*	11.17*	0.14
	(0.19)	(0.17)	(0.26)	(6.52)	(6.45)	(0.13)
Panel B:						
OLS Not Weighted	$\frac{Fatalities}{Enlistment}$	$\frac{Fatalities}{Enlistment}$	$\frac{ArmyFatalities}{Enlistment}$	$\frac{Fatalities}{Population}$	$\frac{Fatalities}{Population}$	$\frac{Enlistment}{Men15-44}$
Pre War	-0.09	-0.07	-0.10	5.42***	5.79***	-0.01**
	(0.07)	(0.07)	(0.07)	(1.56)	(1.55)	(0.01)
1944-1960	-0.04	-0.01	-0.06	2.70**	3.48***	-0.04
	(0.05)	(0.06)	(0.06)	(1.19)	(1.33)	(0.03)
1964-1980	-0.15**	-0.05	-0.18***	2.15	3.12*	-0.05
	(0.06)	(0.07)	(0.07)	(1.58)	(1.65)	(0.03)
1984-2000	-0.08	-0.04	-0.13	4.24**	5.05***	-0.03
	(0.08)	(0.08)	(0.09)	(1.58)	(1.89)	(0.03)
2000-2016	0.16	0.15	0.20*	8.67***	9.42***	-0.02
	(0.10)	(0.10)	(0.11)	(2.39)	(2.48)	(0.02)
σ_{st}, σ_c	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls		Black, Education		No	Enlistment	
N	88671	88671	88671	88671	88671	88671

Note: The sample is composed by all counties for which fatality data is available, and all election years between 1900 and 2016. The dependent variable is the log of Republican presidential vote share. The omitted category is the year 1940. Each column denotes a different explanatory variable. Estimates are obtained via OLS Weighted by 1940 Population in Panel A, and regular OLS in Panel B. State times year fixed effects and county fixed effects are included. Clustered standard errors at the county level are shown. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

Table 3.3: Robustness of the Estimates of Fatalities on Republican Vote Share

	Independent Variable:					
Panel A:	(1)	(2)	(3)	(4)	(5)	(6)
OLS Weighted	$\frac{Fatalities}{Enlistment}$	$\frac{Fatalities}{Enlistment}$	$\frac{ArmyFatalities}{Enlistment}$	$\frac{Fatalities}{Population}$	$\frac{Fatalities}{Population}$	$\frac{Enlistment}{Men15-44}$
Pre War	-0.11	-0.04	-0.20	12.80***	12.68***	0.02
	(0.23)	(0.23)	(0.27)	(4.73)	(4.65)	(0.02)
1944-1960	-0.13*	-0.14*	-0.22***	6.49**	5.95*	0.02*
	(0.07)	(0.07)	(0.08)	(3.06)	(3.10)	(0.01)
1964-1980	-0.40***	-0.40***	-0.47***	3.47	2.74	0.02
	(0.08)	(0.08)	(0.11)	(4.82)	(4.85)	(0.01)
1984-2000	-0.52***	-0.51***	-0.59***	8.76*	7.50	0.04**
	(0.13)	(0.13)	(0.16)	(5.25)	(5.30)	(0.02)
2000-2016	-0.56**	-0.51***	-0.39	18.82**	16.89**	0.07*
	(0.23)	(0.19)	(0.32)	(8.65)	(8.58)	(0.03)
Panel B:						
OLS Not Weighted	$\frac{Fatalities}{Enlistment}$	$\frac{Fatalities}{Enlistment}$	$\frac{ArmyFatalities}{Enlistment}$	$\frac{Fatalities}{Population}$	$\frac{Fatalities}{Population}$	$\frac{Enlistment}{Men15-44}$
Pre War	-0.11	-0.10	-0.13	9.26***	9.26***	0.00
	(0.09)	(0.09)	(0.10)	(2.26)	(2.26)	(0.00)
1944-1960	-0.01	-0.01	-0.02	5.67***	5.50***	0.01
	(0.07)	(0.08)	(0.08)	(1.82)	(1.83)	(0.01)
1964-1980	-0.13	-0.06	-0.17*	5.98***	5.90***	0.00
	(0.08)	(0.09)	(0.10)	(2.22)	(2.23)	(0.00)
1984-2000	-0.03	-0.04	-0.09	9.61***	9.47***	0.01***
	(0.10)	(0.10)	(0.11)	(2.56)	(2.56)	(0.00)
2000-2016	0.33**	0.24*	0.41***	17.17***	17.02***	0.01
	(0.14)	(0.14)	(0.15)	(3.57)	(3.57)	(0.01)
σ_{st}, σ_c	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls		Black, Education		No	Enlistment	
N	88671	88671	88671	88671	88671	88671

Note: The sample is composed by all counties for which fatality data is available, and all election years between 1900 and 2016. The dependent variable is the Republican presidential vote share. The omitted category is the year 1940. Each column denotes a different explanatory variable. Estimates are obtained via OLS Weighted by 1940 Population in Panel A, and regular OLS in Panel B. State times year fixed effects and county fixed effects are included. Clustered standard errors at the county level are shown. Significance at the 10% level is represented by *, at the 5% level by ** and at the 1% level by ***.

Bibliography

Abramitzky, R., A. Delavande, and L. Vasconcelos (2011): "Marrying up: the role of sex ratio in assortative matching," *American Economic Journal: Applied Economics*, 3, 124–57.

Acemoglu, D., D. H. Autor, and D. Lyle (2004): "Women, war, and wages: The effect of female labor supply on the wage structure at midcentury," *Journal of political Economy*, 112, 497–551.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001): "The colonial origins of comparative development: An empirical investigation," *American economic review*, 91, 1369–1401.

——— (2002): "Reversal of fortune: Geography and institutions in the making of the modern world income distribution," *The Quarterly journal of economics*, 117, 1231–1294.

Alesina, A. F. and R. Perotti (1999): "Budget deficits and budget institutions," in *Fiscal institutions and fiscal performance*, University of Chicago Press, 13–36.

Althaus, S. L., B. H. Bramlett, and J. G. Gimpel (2012): "When war hits home: The geography of military losses and support for war in time and space," *Journal of Conflict Resolution*, 56, 382–412.

Bagues, M. and P. Campa (2017): "Can gender quotas empower women:? Evidence from a Regression Discontinuity Design," Tech. rep., mimeo.

Bagues, M. F. and B. Esteve-Volart (2010): "Can gender parity break the glass ceiling? Evidence from a repeated randomized experiment," *The Review of Economic Studies*, 77, 1301–1328.

Balinski, M. L. and H. P. Young (1978): "The Jefferson method of apportionment," *Siam Review*, 20, 278–284.

Baltrunaite, A., P. Bello, A. Casarico, and P. Profeta (2014): "Gender quotas and the quality of politicians," *Journal of Public Economics*, 118, 62–74.

Bazzi, S., M. Fiszbein, and M. Gebresilasse (2017): "Frontier Culture: The Roots and Persistence of" Rugged Individualism" in the United States," .

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2009): "Powerful women: does exposure reduce bias?" *The Quarterly Journal of Economics*, 124, 1497–1540.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012): "Female leadership raises aspirations and educational attainment for girls: A policy experiment in India," *science*, 335, 582–586.

Bellou, A. and E. Cardia (2016): "Occupations after WWII: The legacy of Rosie the Riveter," *Explorations in Economic History*, 62, 124–142.

Bergh, A., G. Fink, and R. Öhrvall (2012): "Public Sector Size and Corruption: Evidence from 290 Swedish Municipalities," .

Besley, T. and A. Case (2003): "Political institutions and policy choices: evidence from the United States," *Journal of Economic Literature*, 41, 7–73.

Besley, T., O. Folke, T. Persson, and J. Rickne (2017): "Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden," *American Economic Review*, 107, 2204–2242.

Bethmann, D. and M. Kvasnicka (2012): "World War II, missing men and out of wedlock childbearing," *The Economic Journal*, 123, 162–194.

Brennan, G. and J. M. Buchanan (1977): "Towards a tax constitution for Leviathan," *Journal of Public Economics*, 8, 255–273.

——— (1978): "Tax instruments as constraints on the disposition of public revenues," *Journal of Public Economics*, 9, 301–318.

Brennan, G., J. M. Buchanan, et al. (1980): *The power to tax: Analytic foundations of a fiscal constitution*, Cambridge University Press.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014): "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 82, 2295–2326.

Camacho, A. and E. Conover (2011): "Manipulation of social program eligibility," *American Economic Journal: Economic Policy*, 41–65.

Campa, P. (2011): "Gender quotas, female politicians and public expenditures: quasi-experimental evidence," .

Carrillo Barroso, E. and M. Tamayo Sáez (2011): "La formación de las preferencias de gasto público: Un análisis comparado por políticas públicas," *Frontera norte*, 23, 193–229.

Carvajal, J. (2017): "The More the Merrier: Evidence of Misreporting Population for Political Gain by Spanish Municipalities," *Working Paper.*

Casas-Arce, P. and A. Saiz (2015): "Women and power: unpopular, unwilling, or held back?" *Journal of political Economy*, 123, 641–669.

Chattopadhyay, R. and E. Duflo (2004): "Women as policy makers: Evidence from a randomized policy experiment in India," *Econometrica*, 72, 1409–1443.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2009): "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records," .

Christensen, G. (2017): "Occupational Fatalities and the Labor Supply: Evidence from the Wars in Iraq and Afghanistan," *Journal of Economic Behavior & Organization*, 139, 182–195.

Coate, S. and B. Knight (2011): "Government form and public spending: Theory and evidence from US municipalities," *American Economic Journal: Economic Policy*, 3, 82–112.

Costas-Pérez, E., A. S. Ollé, and P. S. Navarro (2011): "Corruption scandals, press reporting, and accountability: Evidence from Spanish mayors," *Documents de treball IEB*, 1.

Coviello, D. and S. Gagliarducci (2010): "Tenure in office and public procurement," .

Curto Grau, M., A. Solé Ollé, and P. Sorribas (2012): "Partisan targeting of intergovernmental transfers & state interference in local elections: evidence from Spain," *Documents de treball (Facultat d'Economia i Empresa. Espai de Recerca en Economia), 2012, E12/288*.

Dahlerup, D., Z. Hilal, N. Kalandadze, and R. Kandawasvika-Nhundu (2013): *Atlas of electoral gender quotas*, IDEA.

De Paola, M., V. Scoppa, and R. Lombardo (2010): "Can gender quotas break down negative stereotypes? Evidence from changes in electoral rules," *Journal of Public Economics*, 94, 344–353.

Doepke, M., M. Hazan, and Y. D. Maoz (2015): "The baby boom and World War II: A macroeconomic analysis," *The Review of Economic Studies*, 82, 1031–1073.

Duflo, E. (2012): "Women empowerment and economic development," *Journal of Economic Literature*, 50, 1051–1079.

Edlund, L. and R. Pande (2002): "Why have women become left-wing? The political gender gap and the decline in marriage," *The Quarterly Journal of Economics*, 117, 917–961.

Esteve-Volart, B. and M. Bagues (2012): "Are women pawns in the political game? Evidence from elections to the Spanish Senate," *Journal of Public Economics*, 96, 387–399.

Fernández, R., A. Fogli, and C. Olivetti (2004): "Mothers and sons: Preference formation and female labor force dynamics," *The Quarterly Journal of Economics*, 119, 1249–1299.

Fernández-Caballero, L., D. J. Pedregal, and J. J. Pérez (2011): "Monitoring subcentral government spending in Spain," *Banco de Espana Working Paper*.

Ferrara, A. (2017): "Economic and Social Integration of Minorities: The Effect of WWII on Racial Segregation," .

Ferreira, F. and J. Gyourko (2014): "Does gender matter for political leadership? The case of US mayors," *Journal of Public Economics*, 112, 24–39.

Gagliarducci, S. and M. D. Paserman (2011): "Gender interactions within hierarchies: evidence from the political arena," *The Review of Economic Studies*, 79, 1021–1052.

Gartner, S. S. (2008): "The multiple effects of casualties on public support for war: An experimental approach," *American Political Science Review*, 102, 95–106.

Gartner, S. S., G. M. Segura, and B. A. Barratt (2004): "War casualties, policy positions, and the fate of legislators," *Political Research Quarterly*, 57, 467–477.

Goldin, C. and C. Olivetti (2013): "Shocking labor supply: A reassessment of the role of World War II on women's labor supply," *American Economic Review*, 103, 257–62.

Goldin, C. D. (1991): "The role of World War II in the rise of women's employment," *The American Economic Review*, 741–756.

Hayes, A. F. and T. A. Myers (2009): "Testing the "proximate casualties hypothesis": Local troop loss, attention to news, and support for military intervention," *Mass Communication and Society*, 12, 379–402.

Jaworski, T. (2014): ""You're in the Army Now:" The Impact of World War II on Women's Education, Work, and Family," *The Journal of Economic History*, 74, 169–195.

Karol, D. and E. Miguel (2007): "The electoral cost of war: Iraq casualties and the 2004 US presidential election," *The Journal of Politics*, 69, 633–648.

Kriner, D. and F. Shen (2014): "Responding to war on Capitol Hill: Battlefield casualties, congressional response, and public support for the war in Iraq," *American Journal of Political Science*, 58, 157–174.

Levitt, S. D. (1996): "How do senators vote? Disentangling the role of voter preferences, party affiliation, and senator ideology," *The American Economic Review*, 425–441.

Loewen, P. J. and D. Rubenson (2010): "Canadian war deaths in Afghanistan: Costly policies and support for incumbents," *Working Paper. Toronto, Canada: University of Toronto.*

Lott, Jr, J. R. and L. W. Kenny (1999): "Did women's suffrage change the size and scope of government?" *Journal of political Economy*, 107, 1163–1198.

Miller, G. (2008): "Women's suffrage, political responsiveness, and child survival in American history," *The Quarterly Journal of Economics*, 123, 1287–1327.

O'Brien, D. and J. Rickne (2016): "Gender Quotas and Women's Political Leadership," *American Political Science Review*, 110, 112–126.

Olken, B. A. (2006): "Corruption and the costs of redistribution: Micro evidence from Indonesia," *Journal of public economics*, 90, 853–870.

Pettersson-Lidbom, P. (2012): "Does the size of the legislature affect the size of government? Evidence from two natural experiments," *Journal of Public Economics*, 96, 269–278.

Saez, E. (2010): "Do taxpayers bunch at kink points?" *American Economic Journal: Economic Policy*, 180–212.

Sullivan, J. L., A. Fried, and M. G. Dietz (1992): "Patriotism, politics, and the presidential election of 1988," *American Journal of Political Science*, 200–234.

Svaleryd, H. (2009): "Women's representation and public spending," *European Journal of Political Economy*, 25, 186–198.

Weeks, A. C. and L. Baldez (2015): "Quotas and qualifications: the impact of gender quota laws on the qualifications of legislators in the Italian parliament," *European Political Science Review*, 7, 119–144.

CURRICULUM VITAE

[redacted]