ESSAYS IN POLITICAL ECONOMY

Ekim Cem Muyan

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

____

Camilo Garcia Jimeno, Professor of Economics

Graduate Group Chairperson

____

Jesus Fernandez-Villaverde, Professor of Economics

Dissertation Committee

Camilo Garcia-Jimeno, Professor of Economics

Antonio Merlo, Professor of Economics

Kenneth I. Wolpin, Professor of Economics

To my parents and Apoorva for their unwavering support and understanding.

ACKNOWLEDGEMENTS

I am grateful to my advisors, Professor Antonio Merlo, Professor Kenneth I. Woplin and Professor Camilo Garcia Jimeno. They all provided invaluable guidance and constantly pushed me to do better. Prof. Merlo always encouraged me to follow my own interests while being mindful about the direction I am headed and illuminated the road ahead. Prof. Wolpin was an intellectual inspiration and reminded me the beauty behind the work we do at the hardest moments of my studies. Prof. Garcia-Jimeno was as much a friend as an advisor and a true joy to work with. I am very fortunate to have met them and more so that they were on my thesis committee.

I would also like to thank my friends and classmates in the department. They are whom I learnt and had fun with even while being agonized through the tough days at the office. I especially would like to thank Harun Alp, Diego Amador, Ian Appel, Sina Teoman Ates, Lorenzo Braccini, Murat Alp Celik, Nicolas Grau, Daniel Neuhann, Gokhan Oz, Devin Reilly and Francisco Silva.

I would also like to thank my mother, Gul Buyukbese, my father, Mete Muyan, and my uncle, Mesut Muyan, for their unconditional support and love. Even though they were physically far away during my graduate studies, they were actually closer to me than ever before. Finally, and most importantly, I thank Apoorva Jadhav for being the *home* I always longed for. Her love, understanding, support and belief in me were cardinal pillars of this dissertation.

Ekim Cem Muyan

Philadelphia, PA

March 23, 2016

ABSTRACT

ESSAYS IN POLITICAL ECONOMY

Ekim Cem Muyan

This dissertation consists of two chapters on topics in political economy. In the first chapter, I study alternative decentralization institutions. I argue that decentralization institutions differ from each other in terms of the degree of power they grant to local authorities in decision-making. Successfully designing decentralization institutions depends on understanding the local and central authorities' preferences over the types of public investments and how alternative decentralization institutions aggregate them. Focusing on these key components, I build and estimate a dynamic committee decision-making model to study how public investment choices vary with the degree of power granted to local governments. I characterize alternative decentralization institutions as voting mechanisms the committee can employ. I implement my model using a novel dataset from a unique institution in Colombia. I find that the local governments are more likely to invest on targeted transfers than is the central government. Counterfactual exercises show that a complete decentralized system would significantly increase the number and size of the targeted transfer spending.

In the second chapter, which is co-authored with Devin J. Reilly, I develop a model of campaign strategies, namely the choice to campaign negatively or positively. In particular, I construct a model of political campaigns, based off of Skaperdas and Grofman (1995), in which candidates allocate their budget between positive and negative campaigning. Elections vary according to politician- and district-specific characteristics, as well as the unobservable (to the econometrician) measure of voter types. I calibrate the model to match stylized facts on campaign tone that we document using a wide array of sources, including data on advertising tone from Wisconsin Advertising Project, campaign contributions from the Database on Ideology, Money in Politics, and Elections, and election results. The calibrated model implies that, overall, campaign spending is not particularly effective at in-

creasing votes – a 10% increase in the average candidate's budget, corresponding to about $240,000, raises his or her expected vote share by about 0.4 percentage points. The model also implies that negativity is marginally more useful for candidates who are trailing than those leading, though not by a wide margin.

Contents

List of Tables

List of Figures

Chapter 1 : Public Investment under Alternative Decentralization Institutions

Decentralization is an important mechanism used by governments to improve public service delivery. Especially during the last 30 years, many developing countries have adopted some form of political decentralization by granting more power to local governments in public investment decisions. These attempts have differed widely in terms of the decentralization mechanisms employed by each country, and they subsequently have generated a mixed set of results. Some countries have improved local public good delivery while others have not. Given the trend in increasing decentralization around the world and the centrality of these institutions to development outcomes, it is critical to quantitatively analyze and understand the relationship between specific decentralization mechanisms and economic outcomes.

In essence, decentralization institutions are mechanisms that aggregate the preferences of different layers of government. A political system becomes more decentralized as the weight assigned to the preferences of the local government in this aggregation mechanism increases. As long as there is some disagreement between the local and the central governments, this mechanism plays a crucial role in determining the types of public expenditure in the short-term and, hence, development in the long term. The empirical literature to date has focused on the effects of a specific decentralization institution on various outcomes. In this paper, I take the heterogeneity of decentralization institutions seriously and study their complex effects on public investment decisions. Specifically, I ask how public investment choices vary according to degree of power granted to different layers of government. I show that the answer to this question is the key to successfully designing decentralization institutions.

Answering this question is a challenging task that requires several layers of analysis. First, I build a model that can flexibly represent different degrees of decentralization. To do this, I treat and model decentralization institutions as voting mechanisms that aggregate the preferences of local and central governments. A voting mechanism can be mapped into a particular decentralization institution depending on the weight assigned to the preferences of each layer of government. The model, therefore, establishes a systematic method to analyze the effects of decentralization institutions on public investment choices. Second, I provide an empirical method to identify and estimate the preferences of different layers of government exploiting a unique institution in Colombia. My model, together with empirically uncovering the preferences of different layers of government, enable me to predict the outcome of any possible decentralization attempt in Colombia.

The unique institution I examine in Colombia is the distribution system of royalties collected from its rich supply of natural resources. In 2012, these royalties amounted to 1.4% of the Colombian GDP. Given that public spending that year was 14% of the GDP, royalties constitute a substantial share of the governmental expenditure. In accordance with the ambitious decentralization plan laid out in the 1991 Constitution, the Colombian government distributes about 30% of these royalties to local municipalities, as mandated by a fixed formula. The projects that municipal mayors can implement using the allocated budget can be broadly grouped into public goods (education, health, transportation infrastructure, etc.) or targeted transfers (sports venues, dwellings, agricultural support, etc.). However, mayors are not granted complete control over the allocated resources. Each project proposed by the mayor is subject to the approval of a local committee. Comprising the committee are representatives of each layer of the Colombian government: a representative of the central government, the mayor of the municipality, and the governor of the municipality's state. All are voting members. Proposed projects are implemented if the committee approves the project. Interestingly, there was an unanticipated change in the voting rule employed by these committees. When the system was first introduced in 2012, the central government could veto any project proposed by the mayor. However, in September 2013, the Supreme Court found this unconstitutional. Before the Supreme Court decision, the only decisive coalition that could approve a project consisted of the center and the mayor. After September 2013, the mayor and the governor became another decisive coalition. To empirically study the question at hand, I collected primary data on the approved projects in these committees from the 2013-2014 budget cycle, which covers the period immediately before and after the Supreme Court decision.

An interesting feature of this setting is the absence of the traditional trade-off of decentralization, as first outlined by Oates. This trade-off revolves around the superior information of the local government regarding the local needs and the ability of the central government to organize the spillovers from public investments. In the Colombian context, the municipal committees work on very small projects that have, at most, negligible spillovers. In addition, the structure of the committees enables information sharing among members where the local information is easily communicated to the center. These characteristics imply that variations in votes are generated by differences in preferences regarding the composition of public investments; they are not a product of informational differences or spillovers. Thus, the setting provides a unique opportunity to observe changes in public investment outcomes that are generated solely by the preferences of different layers of government.

I begin my analysis by developing a dynamic committee decision-making model. The model has finite periods where a committee consisting of three players decides how to invest a given initial budget. These players represent different layers of government (mayor, governor and center). In each period, a project is drawn randomly and it is characterized by its type (whether it is a public good or a targeted transfer spending), cost and player-specific preference shocks. After observing the drawn project, each player votes on the proposed project. These votes depend on the trade-off a player faces between the instantaneous utility the project provides and its opportunity cost, which is measured in terms of forgone future projects and savings. These votes are then aggregated according to a specific mechanism. This aggregation mechanism reflects how potential degrees of decentralization enter the model. For example, the aggregation mechanism that accounts only for the vote of the mayor (the center) represents complete decentralization (centralization). The aggregation mechanisms between these two extremes, on the other hand, reflect intermediate degrees of decentralization. If the proposed project is accepted according to the aggregation mechanism, it is implemented using the budget, and the committee starts the new period. I show that the sequential equilibrium of the game is unique when players vote sincerely. The equilibrium composition of public investment depends on the preferences of the players and the aggregation mechanism in use. In fact, it is easy to show that the set of admissible projects under alternative voting mechanisms varies according to the degree to which the preferences of different levels of government are aligned. If these levels have similar preferences regarding the types of public investment, the specific aggregation mechanism used does not alter the composition of public investment significantly. However, when preferences across levels of government are misaligned, the particular mechanism used plays an important role in determining the outcome.

The fundamental challenge of identifying preferences stems from the fact that the rejected projects and the individual votes for the accepted projects are unobserved in the data. Nevertheless, I show that identification is possible due to the variation in the accepted projects before and after the Supreme Court decision and several plausible exclusion-restrictions. The former uses the first order implication of the model: each municipality is affected differently by the Supreme Court decision. For example, where the preferences of the center and the mayor are aligned, the mayor is able to pass the projects she wants to implement with the help of the center during any period, and she does not need the vote of the governor. However, when the preferences of the mayor and the center differ, a large shift in the set of admissible projects is possible when the veto power is revoked. The magnitude of this change depends on the extent to which the preferences of the mayor and the governor are aligned. When those preferences are aligned, projects that were vetoed by the center can now be approved. This variation in municipal responses to the Supreme Court decision helps me identify the preferences of different layers of government. I also impose plausible exclusion restrictions on the preferences of the players and on the distributions of project draws. For the preference parameters, I assume that each player's utility from a specific type of project is affected by different municipal characteristics. These include political alignment variables for the targeted transfer projects. For example, the center's utility from a targeted transfer project depends on whether the mayor belongs to the party of the president. In the case of public good projects, each player values the public stock capital of the municipality. Moreover, the center's preferences with regards to public goods projects depend on the municipal debt. This is a source of concern for the central government because municipality debts caused a deep fiscal crisis during the early 2000s. Governors' preferences regarding public good projects, on the other hand, depend on the distance of the municipality from the state capital. Finally, the preferences of the mayor regarding public good projects depend on the recent investments she has made on these goods from the municipal budget.

In a first glance at the data, I compare the share of expenditures devoted to public goods before and after the Supreme Court decision. Comparison of the means shows that following the Supreme Court decision, the share of expenditures in public goods decreased 6 percentage points. Next, I show that the main predictions of the theory are supported by the data. My model implies that in committees with a mayor whose preferences resemble those of the governor should be strongly affected by the institutional change. To validate this prediction, I look at municipalities in which the mayor comes from the same party as the governor. In these municipalities, the share invested in public goods decreased by 19 percentage points following the Supreme Court decision. The other immediate implication of the theory is that the shift in the composition of expenditure observed in the data should be smaller when a municipality has a mayor whose preferences are aligned with those of the center. As suggested by the theory, in municipalities in which the mayor belongs to president's party, the drop in the share of expenditure on public goods is statistically insignificant.

These mean comparisons, however, can be misleading for several reasons. Most importantly, the committees in the veto and the no-veto periods are not facing the same problems. This is due to the fact that the Supreme Court decision is taken in the midst of a budget cycle. Thus, the committees before and after the Supreme Court decision face different budgets and dynamic concerns, which implies that the outcomes observed in different periods are incomparable through this technique. For example, committees could be more prone to spend on targeted transfers when little is left in the bank or when the budget cycle is close to an end. This would imply that the share of expenditure spent on public goods would drop even without a change in the institutions. In this case, the share of expenditure spent on public goods would drop even in the absence of institutional change. In fact, my structural estimates show that the impact of the institutions is smaller than the mean comparison approach suggests.

I estimate the model using maximum likelihood. Structural estimates of preference parameters suggest that, on average, mayors are more (less) likely to vote in favor of targeted transfer (public goods) projects compared to the central government. To quantify this, I look at a hypothetical municipality that has median characteristics and a median budget voting on mean cost project draws. During the periods with central veto power (before the Supreme Court decision), the mayor accepts 55% of the public good projects while the center accepts 82% of them. On the other hand, the same mayor accepts 51% of the targeted transfer projects while the center accepts only 25% of them. I also find that the governor's preferences in this municipality are closer to those of the central government. However, after the Supreme Court decision, the odds that a mean sized transfer project passes the committee increases twice more than the odds of a public goods project does. Moreover, the mayor and the center are more aligned when the municipality has low debt, low fiscal independence or when they belong to the same political party. These municipalities are the least affected by the institutional change. On the other hand, the preferences of the mayors and the governor are aligned when municipalities are close to the state capital or they belong to the same party. These municipalities are the ones that respond most to the Supreme Court decision. The model fits the data well, although it tends to under predict the size of the approved projects during the non-veto periods.

I perform several counterfactuals to assess public investment choices under alternative decentralization institutions. Because the mayor is the proposer of all the projects in the data, the estimated parameters for project arrivals are mayor-specific. This implies that my analysis is limited to the cases where the mayor proposes the projects. However, I provide complete centralization results under the restrictive assumption that the project arrivals do not change under this regime. Under full decentralization, where the allocated money is under the full control of the mayor, the average number of transfer projects per municipality increases by about 20%. There is also a 15% increase in the size of this type of projects while the size of the public good projects remains at a level similar to that observed in the data. Under full decentralization, 57% of expenditure is on public goods projects, compared to 68% in the data. I also compare two counterfactuals to assess the *true* effect of institutional change. First I look at the counterfactual where the veto power of the central government was never revoked. In this case, the share of expenditure spent on public goods is about 73%. For the counterfactual with the majority aggregation rule, the same statistic is 68%. Hence, my model predicts that the impact of the institutional change is about 5%. Thus, structural estimation shows that only 83% of the impact suggested by mean comparison is actually due to the shift in the institutions.

To the best of my knowledge, there are no studies that empirically analyze how changes in the degree of decentralization affect public investment choices. Doing this is my main contribution to the literature. The paper relates to several strands of literature –both empirical and theoretical. First is the empirical literature that assesses the impact of decentralization on economic outcomes. The paper closest to mine in this literature is Foster and Rosenzweig [2001], where the authors analyze the impact of Indian decentralization on the provision of different types of public investment. They show that decentralization has led to an increase in public investment that benefits poor citizens in India. However, their modeling choice and setting do not allow the analysis of alternative decentralization institutions. Just like Foster and Rosenzweig [2001], the rest of the empirical literature on decentralization studies the impact of a specific decentralization attempt without exploring the implications of different institutional designs[1]. The more recent empirical literature on decentralization focuses on local elite capture and corruption (for example, see Araujo et al. [2006], Galasso and Ravallion [2005], Bardhan and Mookherjee [2006], Bardhan and Mookherjee [2013], Besley et al. [2012]). Although I do not micro-found the reasons behind the divergence of local and central preferences, elite capture of local political institutions in Colombia could be a convincing channel to explain the results.

My empirical approach contributes to a recent trend that tries to uncover individual preferences from aggregate outcomes. For example, Merlo and de Paula [2010] non-parametrically estimate the distribution of voter preferences from aggregate observations on multiple election outcomes. My paper contributes to this literature by showing that it is possible to identify the specific preferences of individual players in smaller games under certain conditions. Finally, the estimation method I employ contributes to the literature on the estimation of dynamic discrete choice models initiated by Keane and Wolpin [1997]. I extend their method to accommodate the decisions of a committee rather than a single player.

I also contribute to the theoretical literature on decentralization. Following the traditional trade-off introduced by Oates [1972], a rich literature on the political economics of decentralization has emerged. For example, Besley and Coate [2003] model the underlying legislative processes in centralized and decentralized provisions of public goods. They find that what distinguishes the two systems is both the magnitude of spillovers and the di-

[1]For an in depth overview of this literature see Mansuri and Rao [2013]. For a review on the impact of decentralization on health and education outcomes around the world, see Channa and Faguet [2012]. For the special case of Colombia, see Faguet and Snchez [2014]. For a comparison of Colombia and Bolivia, see Faguet and Snchez [2008]. For Argentina, see Eskeland and Filmer [2002]; for Tanzania, see Maro [1990] and Ellis and Mdoe [2003]; for Nicaragua, see Larson [2002]; for Peru see Loayza et al. [2014]; for Uganda, see Akin et al. [2005]

vergence of preferences regarding the spending of localities. In another approach, Myerson [2015] builds an agency model to argue that the local governments can manage local agency problems better than the central government due to their political accountability to local residents. Similarly, Tommasi and Weinschelbaum [2007] use an agency model to study the effects of the size of constituencies on the accountability of governments. They find that the larger the jurisdiction of governments the lower the accountability of public officials. I contribute to this literature by employing a committee decision-making model to understand the heterogeneity in decentralization institutions. In contrast to the previous literature, my approach allows me to study different degrees of decentralization institutions though the aggregation mechanism employed in the committee. My model also contributes to the theoretical literature that studies the collective search problem of a committee. Compte and Jehiel [2010] and Albrecht et al. [2010] study this setting where a committee accepts or rejects randomly drawn projects. In their papers, the committees search for a single project. I extend this setting so that the committee decides on the dynamic allocation of a fixed budget across projects that arrive randomly.

The remainder of my paper proceeds as follows. Section 1.1 introduces details on the Colombian royalty distribution system. Section 1.2 provides a model of dynamic committee decision making and a discussion of the effects of a sudden change in the aggregation rule of preferences within these committees. Section 1.3 presents the data I collected from Colombian royalty distribution system and some reduced form analysis on the effects of the Supreme Court decision. Section 1.4 discusses the empirical specification and the identification of my model. I present in Section 1.5 my structural estimates and their implications and in Section 1.6 several counterfactuals that I then perform. Section 1.7 concludes.

1.1 Background and Setting

Colombia has undergone a massive decentralization process initialized by the 1991 constitution. The primary goals of the reform were to improve the access of people to social and public services, end poverty and diminish territorial inequalities. Since the 1991 reforms, the Colombian government has continuously experimented with different degrees of decentralization[2]. Thus, the scope of the Colombian political decentralization today is the outcome of an ongoing eclectic process. The fundamental reason behind this experimentation is the poor performance of the implemented reforms. For example, local fiscal decentralization led to a serious municipal debt crisis in the beginning of the 2000s. The central government responded to this crisis by heavily regulating the loaning capacity of the municipal governments. Similarly, central government realized that the decentralization of public service provision at the municipal level was ineffective due to the lack of local state capacity[3]. Most municipalities were either not able to generate substantial resources to fund large public investment projects or the quality of services were subpar due to a lack of human capital. This led the central government to diversify the levels of decentralization by sector. As a result, some sectors are completely decentralized at the local level whereas others are still at the responsibility of the central government. For example, water sanitation is completely decentralized at the local level. On the other hand, education is decentralized at the elementary school level but not at the high school level–where the state and central governments share responsibility.

In line with the 1991 constitution, use of the royalties from natural resources was also decentralized by the Colombian government. According to the Oil and Gas Journal (OGJ), Colombia had approximately 2.4 billion barrels of crude oil reserves as of January 1, 2014 and is also a significant producer of gold, nickel and coal. The government often gives the right to exploit these resources to private companies for a fee. The revenues from these fees, which are referred as the royalties, summed up to 4.6 billion dollars in 2012. These amount to 1.4% of the Colombian GDP that year whereas the total government expenditure was about 14% of the economy. Thus, the royalties are an important part of public expenditure and they have been steadily growing in the last 20 years.

[2]See Ramirez et al. [2014], Faguet and Snchez [2014], Bird [2012] for details on Colombian decentralization and its effects

[3]See Acemoglu et al. [2015] and Fiszbein [1997] for more on municipal state capacity in Colombia.

Before 2012, the royalties were mainly controlled by the municipalities that produced them. Hence, instead of redistributing the revenues through a central system, Colombian government allocated a large part of the profits toward the municipalities the natural resources were located at. This system had two fundamental problems. The distribution of royalties was very uneven- benefiting only a handful of states and localities: 70% of the royalties were received by 17% of the population between 1995 and 2011 (DNP [2013]). Moreover, this uneven distribution did not lead to significant increases in the well-being of the citizens living in the producer municipalities (DNP [2013]). The other complaint was the lack of any real checks and balances on the projects funded by the royalties. Even though the royalties were earmarked to be used for public goods and services, it was frequently reported that the mayors spent the money on unproductive projects contracted to companies they were connected to. Some mayors are still facing charges in Colombian courts regarding the projects financed by the royalties. During this era, although most of the producer municipalities grew above the national average, they were not able to significantly decrease poverty. Only two of the producer municipalities reduced poverty levels above the national average (DNP [2013]).

In June 2011, President Santos and the Colombian Congress signed a crucial reform regulating the use and distribution of royalties amid great opposition from the producing municipalities. The reform aimed at resolving the two problems mentioned above. It centralized the collection and distribution of royalties via a new institution called Sistema General de Regelias (SGR). This institution was established to be responsible of a progressive distribution of royalties. Even though the producer municipalities still receive a generous portion, localities without natural resources were guaranteed a substantial share of the centrally allocated royalties. Additionally, the reform aimed to solve the corruption problem by introducing a new system of decision-making. This consisted of the establishment of committees, called Organos Colegiados de Administracion y Decision (OCAD), for each entity that receives funds. Besides the portions kept by the central government, projects that use royalties have to go through the OCADs [4]. Each OCAD committee consists of three members representing central, state and local governments who vote on each proposed project towards the use of the allocated committee budget. This system gives all levels of government power in deciding the implemented projects. The argument of the central government for establishing these committees is to create a system of checks and balances to prevent corruption through the use of royalty income. Of course, the power granted to the central and state governments may cause vertical clientelism. That is to say, central and state governments can be more lenient toward municipalities that are politically close to them. This is a channel I explore in my empirical implementation.

There are OCADs at three different levels of administration. Almost every municipality and state has one OCAD. Moreover, there are 6 regions that consists of several states with one OCAD located at each. In total, there are around 1100 OCADs. State governors have access to the royalties allocated to the state and regional OCADs, in the sense that, only they can propose projects towards using budgets allocated to either of these types of committees. On the other hand, mayors are the project developers in the municipal OCADs. The regional OCADs have very large budgets and hence they work on costlier projects that span across states. Examples of these projects include interstate highways, energy pipe lines, large universities etc. State OCADs work on projects that benefit at least two municipalities at the same time. These could be highways or water pipes that connect municipalities. Municipal OCADs, on the other hand, work on smaller projects like building pavements to municipal roads, building new elementary school buildings etc. Each level functions in a very similar way, where the central government, state government and the local government each cast one vote. Moreover, each OCAD receives a fixed amount of budget for a two-year period. The allocated money is mandated by a formula created by the Congress that depends on the natural resource production, population and need of the locality. This is a particularly nice feature of OCADs: conditioning on the budget, the variation in the votes will be generated by preferences toward projects of different levels of

[4]This accounts for 60% of the total funds, around half of which goes to the municipalities.

government and not distributional concerns.

A particularly interesting feature of these committees is a sudden change in their governing rules. Before September 2013, the central government had the right to veto any project. Hence, a project needed a majority *with* the vote of the central government. With three voting members, the only way a project was approved in a municipal OCAD before September 2013 was with the votes of *both* the mayor and the central representative. The vote of the center was required because of its veto power and the vote of the mayor was guaranteed because she is the proposer. Similarly, in state and regional OCADs, the winning coalition was the governor *and* the central representative. However, the Colombian Supreme Court found the veto right unconstitutional on 11 September 2013 following an application of a lawyer. This implied that the votes of *any* majority became sufficient to implement a project. In the municipal OCADs, the implication of this decision was that the governor and the mayor could pass projects without requiring the vote of the central government.

In what follows, I assume that the Supreme Court decision provides an exogenous shift to the vote aggregation mechanism used in the OCAD committees. There are several indications toward this point. The first is the numerous interviews conducted at Bogota, where the bureacrats were either unaware of the case alltogether, or did not expect the decision to go against the center. In addition, in none of the nationally circulated newspapers, *El Tiempo*, *El Espectador* or *El Periodoco* had news about the Supreme Court case prior to the decision. Hence, the Supreme Court decision provides a plausibly exogenous shift to how OCADs aggregate preferences and hence provides me the power of identification that I will describe later.

I collected data on the projects approved by the municipal OCADs from 2012 to 2014. The reason I chose the municipal OCADs is twofold. First is the richness of the data. There is a lot of variation in municipal characteristics and around 5500 development projects, which helps me significantly in the empirical implementation. Second, municipal projects are rather small, which implies that the spillover of the projects to other municipalities is minimal. In fact, the larger projects with significant spillovers are discussed and implemented at the state and the regional OCADs. Hence, by design, municipal OCADs have negligible spillovers at most. This aspect, together with the structure of the committees that allows information sharing between the layers of governments, implies that the voting members are not concerned about spillovers or informational discrepancies, but solely about their preferences toward the types and costs of the projects in municipal OCADs. Hence, the setting is unique in its ability to emphasize the differences between preferences of layers of government *only* toward the development projects and nothing else .

Procedures of a municipal OCAD. In municipal OCADs, the local government is represented by the municipal mayor. She also is the only entity that can propose projects to utilize the centrally allocated budget to the committee. Moreover, the governor of the state municipality belongs to and a central representative from the Department of National Planning (DNP) are also present in the committees. Although there are no specific laws regarding the exact timing, OCADs meet every two months[5]. Each OCAD project has to satisfy several technical guidelines that require proposals to include thorough technical reports. This requirement is often too demanding for many municipal governments who lack any qualified human capital in their administrations. Thus, these municipalities tend to propose fewer projects, as they simply do not have the capacity to fulfill the technical requirements of proposing one. To alleviate this to a certain extent, the DNP assists municipalities in meeting these criteria. Hence, the central representative knows about the details of the projects well before it is submitted to a vote. These details are often shared with the governor before as well as the process to obtain the technical eligibility of a project is costly.

The budget allocated to the OCADs is not earmarked, however it is mainly aimed at providing funds for capital investments. Operating costs of the projects (like the wages for doctors, teachers etc) can only be covered for a limited period of time. Another important aspect of the OCADs is the leftover budgers. In case a municipality does not exhaust

[5]Law states that the OCADs cannot meet less than twice a year.

its budget from a particular cycle, the remaining funds roll over to the next one. Hence, mayors can strategically wait to accumulate funds for more ambitious projects. Moreover, municipalities can co-fund the projects from other resources if they see fit. For this, they have to guarantee that the fund already does exists, in that, they cannot rely on the expected future taxes or transfers.

The Supreme Court decision reflects on municipal OCADs in terms of who can jointly pass the projects. Before the decision, a successful project needed the votes of both the mayor and the center. Therefore, the vote (and the preferences) of the governor did not play any role in the decision-making process. After the Supreme Court decision, however, votes of the mayor and governor were also sufficient to pass a project. This implies that the types of project an OCAD passed before and after the decision will be different as long as the mayor and the governor want to implement projects that could be previously dismissed by the center.

1.2 Model

I model OCADs' decision making process as a collective dynamic choice problem. The choice of a dynamic game is obvious: each period is linked to another by the remaining budget. This implies that a decision today has an opportunity cost that will be faced tomorrow. Moreover, the setting of a committee gives a straightforward method to map aggregation rules into different levels of decentralization. There are two vote aggregation mechanisms that will be of particular interest in the empirical section. The first one necessitates the vote of the mayor (as she is the proposer) and the center (as it has the veto power) for a project to be successful. The other mechanism, which is employed following the Supreme Court decision, is the one without the veto right of the central government. Hence, in addition to the previous decisive coalition of the mayor and the center, the governor and the mayor can also approve projects. This implies that the system became *more* decentralizated after the Supreme Court.

In the model, there are three players and T periods. In each period, a project is randomly drawn by the mayor. The mayor then decides whether to bring the project to vote or not (which essentially is her vote). If a project is proposed, observing its characteristics, the governor and the central representative vote for it. The project is passed depending on the decision rule used. In the next period, a new project is drawn if the remaining budget is positive, or, if it is the last period, players get utility from the leftover budget. I first show that the equilibrium of this game is unique when committee members vote sincerely. Next, I discuss how an unexpected change in the aggregation mechanism will impact the equilibrium outcome of the game.

1.2.1 Setting

Basic setting. The game has $T < \infty$ periods, $t = 1, 2, \cdots, T$. There is a committee that decides on development projects to be implemented from the fixed allocated budget B_0. This committee has three members, central government (center, c), municipal government (mayor, m) and state government (governor, g) the municipality belongs to. A project is characterized by its cost ϕ, it's type s –public goods or targeted transfers[6]– and a vector of preference shocks $\{\varepsilon_m, \varepsilon_c, \varepsilon_g\}$. At each period, a project is drawn from the distribution $F(X, B_{t,m})$ where X is a vector of relevant municipal characteristics and B_t is the amount of budget available to the committee in period t. Players then vote on the project, which is discussed in more detail below.

Players and strategies. The three players, the mayor, the governor and the central representative, collectively decide on whether to accept or reject a drawn project. The instantaneous utility that player i gets from project $\omega_i = (s, \phi, \varepsilon_i)$ is denoted by the continuous and bounded function $u_i(Y_{m,s,i}, s, \phi, \varepsilon_i) = u_i(Y_{m,s,i}, \omega_i)$ where $Y_{m,s,i}$ is the vector of relevant characteristics of the municipality m to player i when considering a project of type s. The action space of a player in a given period is accepting or rejecting the drawn project.

[6]This categorization of public spending is similar to that used in Persico and Lizzeri [2001] among others.

I assume that the strategies only depend on the period, remaining budget, municipal and project characteristics. Hence the strategy space is given by:

$$d_{i,t} : \{1, \cdots, T\} \times \mathbb{R}_+ \times Y_i \times X \times \mathbb{R}_+ \times \{1, 2\} \times \mathbb{R} \to \{0, 1\}$$

Here $d_i(t, B_t, Y_i, X, \phi, s, \varepsilon_i) = 0$ means that player i's action is to reject a proposal characterized by the vector $\omega_i = (Y, \phi, s, \varepsilon_i)$ at period t with budget B_t. $d_i(t, B_t, Y_i, X, \phi, s, \varepsilon_i) = 1$, on the other hand, means that player i accepts the same project. Notice that this restricts the strategy space not to depend on history[7]. Soon I will make further restrictions on the strategy space and impose that every player behaves *sincerely* and votes *as if* they are pivotal in making the decision.

Moreover, I assume an end-of-game saving function. Define $w : \mathbb{R}_+ \to \mathbb{R}$, a continuous, increasing function that captures the utility from savings at period T. As for the information structure, player i's information at period t is all the relevant distributions and the current draw (s, ϕ, ε).

Decision Rule. The decision rule $D : \{0, 1\}^3 \to \{0, 1\}$ aggregates all the actions of the players. That is, D maps the votes of each player to a final decision. For example, the usual q-majority rule would be

$$D(\{d_i\}_{i \in N}) = \begin{cases} 1 & \text{if } \sum_i d_i \geq q \\ 0 & \text{otherwise} \end{cases}$$

Similarly, suppose that player p is the proposer, and the decision rule is a q-majority. Then,

[7] That is, I do not consider the possibilities where history can matter in any form. This normally would restrict the possibilities of reciprocal strategies or punishment. Given the game is finite horizon, this is not be a big problem. I discuss the limitations imposed by the finite horizon assumption later.

$$D(\{d_i\}_{i \in N}) = \begin{cases} 1 & \text{if } \sum_i d_i \geq q \text{ and } d_p = 1 \\ 0 & \text{otherwise} \end{cases}$$

There are two decision rules that will be of interest in the empirical application. The first is where the municipality is the proposer and the center has the veto power. This is the decision rule in municipal OCADs prior to the revoking of the veto power. This decision rule is given by:

$$D^V(\{d_i\}_{i \in N}) = \begin{cases} 1 & \text{if } d_m = 1 \text{ and } d_c = 1 \\ 0 & \text{otherwise} \end{cases}$$

The second one is the majority rule with a proposer:

$$D^M(\{d_i\}_{i \in N}) = \begin{cases} 1 & \text{if } d_m = 1 \text{ and } \sum_i d_i \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

Problem of a player. The problem of player i, given the strategies of all other players, $d_{-i}(\cdot)$, is:

$$\max_{d_i(\cdot)} \quad \left(\sum_{t=1}^{T} \beta^{t-1} \mathbb{E}_0(u_i(Y_i, \omega_i)\ D(d_i, d_{-i})) \right) + \beta^T w \left(B_0 - \sum_{t=1}^{T} D_t(d_i, d_{-i}) \phi_{t,s} \right)$$
$$\text{s.to} \quad \sum_{t=1}^{T} \phi_t D_t \leq \tilde{B}$$

where the expectation is taken over future project draws and subsequent budgets. The second equation in the objective function is the utility obtained from saving. A solution to this problem is a function that maps each possible project and municipal characteristic under each possible budget and period into a vote, in favor or against.

1.2.2 Basic results

Notice that since the game has a finite horizon, one can solve it backward. To impose tractability on equilibria, I assume that players always vote *sincerely.* That is, they always vote in favor of a project if implementing it (the instantaneous utility) provides higher utility than rejecting. For example, suppose that a project is drawn with characteristics $\omega_i = (s, \phi, \varepsilon_i)$ in the last period. The utility of accepting this project is equivalent to:

$$u_i(Y_s, \omega_i) + \beta w(B_T - \phi) \tag{1}$$

where $u_i(Y_s, \omega_i)$ is the instantaneous utility driven by player i from project (s, ϕ, ε) and $w(B_T - \phi)$ is the utility obtained from savings after implementing the project. On the other hand, utility from rejecting the rejecting the project is:

$$\beta w(B_T) \tag{2}$$

which is the utility from saving the entire budget $B_{m,T}$ instead of spending on the drawn project ω. The *sincere voting* assumption implies that player i will vote in favor of a project characterized by $\omega = (s, \phi, \varepsilon)$ if and only if

$$u_i(Y_i, \omega_i) + \beta w(B_T - \phi) \geq \beta w(B_T) \tag{3}$$

which is equivalent to the flow utility of the project being larger than its *opportunity cost*: $u_i(Y_i, \omega i) \geq \beta \left[w(B_m) - w(B_m - \phi)\right]$. Here, the opportunity cost,

$$\beta \left[w(B_T) - w(B_T - \phi)\right]$$

is the forgone utility in savings when ϕ is invested on the projects. Notice that this assumption is not strong. Since there are only two choices and the game has finite periods, strategic voting is not a concern. The assumption simply avoids non-singleton best response for non-pivotal players[8].

To see the implications of the sincere voting assumption at any other t, let $\omega_i = (s, \phi, \varepsilon_i)$ and define $V : \mathbb{R}_+ \to \mathbb{R}$ as

$$V(B,t) = \int_{D(d_i(t,\omega_i,B))=1} \big(u(Y,\omega) + \beta V(B-\phi, t+1)\big) dF(\omega|B,X) + \int_{D(d_i(t,\omega,B))=0} \beta V(B,t+1) dF(\omega|B,X)$$

The existence of V is guaranteed by the fact that T is finite, u is bounded and players vote sincerely. Then, for any t, accepting a project characterized by ω provides the utility:

$$u_i(Y,\omega_i) + \beta V_i(B_t - \phi_t) \tag{4}$$

On the other hand, rejecting *any* project will give player i the utility:

$$\beta V_i(B_t) \tag{5}$$

This implies that player i votes in favor of project ω if

$$u_i(Y,\omega) + \beta V_i(B_t - \phi_t) \geq \beta V_i(B_t) \tag{6}$$

Rearranging this inequality leads to a very intuitive form:

[8] Notice that this assumption could be critical had the history mattered. In this alternative model where players can build reputations and punish out-of-equilibrium , sincere voting assumption would cancel many interesting channels.

$$\underbrace{u_i(Y, \omega_i)}_{\text{Instantenous Utility}} \geq \underbrace{\beta\left[V(B, t+1) - V(B-\phi, t+1)\right]}_{\text{Opportunity Cost}} \tag{7}$$

Hence, sincere voting assumption implies that a player votes in favor of a project whenever the instantaneous utility from that project, $u_i(Y, \omega_i)$ provides a higher utility than does the opportunity cost of that project, $\beta\left[V(B, t+1) - V(B-\phi, t+1)\right]$. The opportunity cost of a project is the value of forgone projects due to spending ϕ from the budget. Letting $O(B, \phi, t) = [V(B, t+1) - V(B-\phi, t+1)]$, it is easy to show the following lemma.

Lemma 1.2.1. *$O(B, \phi, t)$ is increasing in ϕ.*

The proof is straight-forward as $V(B-\phi, t+1)$ is obviously decreasing in ϕ. Next, notice that the sincere voting assumption implies that:

$$d_i^*(t, B, \omega_i | d_{-i}) = \begin{cases} 1 & \text{if } u_i(\omega_i) \geq \beta[O(B, \phi, t)] \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Lemma 1.2.2. *Under the sincere voting assumption, the sequential equilibrium of the game is unique where the strategies are given by equation 8 and beliefs are consistent with $F(\omega | X, B)$.*

The idea behind the uniqueness is very simple. Notice that when voting, players can perfectly anticipate the future decisions of each player due to the sincere voting assumption and the fact that they know the distributions of future projects. Therefore, at each possible draw, players have one sincere vote given their beliefs, which constitutes the equilibrium of the game and is given by equation 8. No player has a profitable deviation as the off-equilibrium action always provides at most as much utility as the equilibrium action, given the beliefs.

From an empirical stand-point, the probability that player i will accept a project ω_i with budget B in a municipal OCAD with characteristics Y is

$$Pr(d_i(t,B,\omega_i)=1) = Pr\left(\varepsilon_i : u_i(Y,\omega_i) \geq \beta\left[O(B,\phi,t)\right]\right) \tag{9}$$

Now recall that the decision rules of interest in the empirical implementation are D^M and D^V. If ε_i are independent, the probability of accepting an offer under the majority decision rule with mayor proposing is given by

$$\begin{aligned} Pr(D^M(t,B,\omega_i)=1) = \quad & Pr(d_m(t,B,\omega_i m=1) Pr(d_c(t,B,\omega_c)=1) \\ & + Pr(d_m(t,B,\omega_m)=1) Pr(d_g(t,B,\omega_g)=1) \\ & - Pr(d_m(t,B,\omega_m)=1) Pr(d_c(t,B,\omega_c)=1) Pr(d_g(t,B,\omega_g)=1) \end{aligned} \tag{10}$$

which is equal to the probability that the mayor and at least one other player accepts the project. Similarly, for the decision rule that grants the center the veto right, we have:

$$Pr(D^V(t,B,\omega_i)=1) = Pr(d_m(t,B,\omega_m)=1) Pr(d_c(t,B,\omega_c)=1) \tag{11}$$

which is the probability that *both* the mayor and the center accept the project characterized by ω.

There are couple of things to note. First, under D^V, the preferences of the governor *does not* matter. Since any successful project requires the vote of both the mayor (because of the proposer role) and the center (because of the veto power); a majority is automatically obtained. Second, note that, *facing the same decision*, it is more likely that a project will be approved under the majority with a proposer decision rule, as long as $Pr(d_c(\cdot)=1) < 1$ and $Pr(d_g(\cdot)=1) > 1$. This implies that we are more likely to observe *any* project under the majority rule. This is summarized in the following lemma.

Lemma 1.2.3. *The same project ω, drawn at the same time period, is more likely to be accepted under D^M than D^V as long as $Pr(d_c(t, B, \omega_c) = 1) = 1) < 1$ and $Pr(d_g(t, B, \omega_g) = 1) = 1) > 1$.*

To further analyze the game, I will assume out the preference shocks and convert the game into a deterministic one. Moreover, I will impose concavity on the utilities. This will allow me to show that different decision rules translate into different sets of accepted projects.

Assumption 1. *Assume that $\varepsilon_{i,t} = 0$ for all t and i. Moreover assume that u and w are concave functions.*

Absent ε, each project is identical up to its type and cost. I will analyze the set of admissible sizes of the projects, given a type. Thus, the main conflict between the players is about *whether* to spend the money and *when* to spend it. With assumption 1, the set of projects player i will accept, given (B, t, s, Y, X), is a deterministic object. Dependence on s and Y is redundant, so I will drop these for the moment. Define the admissible set for player i at period t as the set of proposals that will be accepted by player i:

$$A_i(B, t) = \{\phi \in [0, B] \mid d_i^*(B, \phi, t) = 1]\}$$

Lemma 1.2.4. *Given assumption 1, $A_i(B, \phi, t)$ is an interval.*

Sketch of the proof. If $A_i(B, t) = [0, B]$ for all B, t, we are done. Otherwise, note that concavity assumptions in 1 guarantees the concavity of u and convexity of $O(\cdot)$ in ϕ. Next, suppose that $A_i(B, t) \subset [0, B]$. Take $\phi_1 < \phi_2$ such that $\phi_1, \phi_2 \in A(B, t)$ (if we cannot find such ϕ_1, ϕ_2, the statement is immediately correct). Using the concavity of u and convexity of O, one can show that $\phi_\lambda = \lambda\phi_1 + (1 - \lambda)\phi_2 \in A(B, T)$.

Let $dA_i(B, t) = ||A_i(B, t)||$. This measures the size of the admissible set for player i. The smaller dA_i, the more *conservative* player i is about spending the budget. Moreover, define $A^D(B, t)$ as the set of projects that will be admitted given the decision rule D. For example, if D is the unanimity rule, then $A^{UNAN}(B,t) = \bigcap_i A_i(B,t)$. The admissible sets for the decision rules that I will use in the empirical section, on the other hand, become:

$$A^M(B,t) = A_m(B,t) \cap (A_c(B,t) \cup A_g(B,t)) \quad (12)$$

$$A^V(B,t) = A_m(B,t) \cap A_c(B,t) \quad (13)$$

Define a q-decision rule, where q is the number of votes necessary to accept a project. The following lemma states that as we increase the required number of votes to pass a project, the set of admissible projects shrinks. This is intuitive. Each additional vote required by the decision rule imposes the need to match the admissible set of a new player.

Lemma 1.2.5. *$dA^q(B,t)$ is decreasing in q.*

Proof. This follows immediately from the observation that $A_i(B,t) \cup A_j(B,t) \subset A_i(B,t)$.

This result immediately implies that the most *conservative* aggregation rule is unanimity. Let $A_c(B,t)$ be the player with the smallest admissible set, such that $A_{min}(B,t) = \min_i A_i(B,t)$. Then, the most the unanimity rule shall obey the admissible set of this player and admit potentially a smaller set than hers. Following corollary makes this argument formal.

Corollary 1.2.1. $A^{UNAN}(B,t) \subseteq A_{min}(B,t)$. *That is, the admissible set under unanimity rule is a subset of that of the most conservative player.*

While the previous results tie the admissible sets to the decision rule used, the next result lays out the relationship of the admissible sets across time under the same decision rule.

Lemma 1.2.6. *If a project of type s and size ϕ is accepted at period t, it will also be accepted in period $t+1$. That is, $A(B,t,s) \subseteq A(B,t+1,s)$.*

Sketch of the proof. The proof follows from induction. To see this, note that if a project is accepted at $T-1$, we have

$$
\begin{aligned}
u(Y,s,\phi) &\geq \int_{D(d_i(T-1,\omega_i,B))=1} \big(u(Y,\omega)+\beta w(B-\phi)\big)dF(\omega|B,X) \\
&+ \int_{D(d_i(t,\omega,B))=0} \beta w(B)dF(\omega|B,X) \\
&\geq \beta[w(B)-w(B-\phi)]
\end{aligned}
$$

The second inequality follows from the fact that the opportunity cost of a project at t is always higher than the opportunity cost at $t+1$ as there are more *opportunities* for new projects. Inductively continuing this argument delivers the result.

The intuition behind this result is simple. The opportunity cost of implementing a project decreases as the game nears to an end. This is because the probability of drawing more favorable projects than the one that has already been drawn decreases. Thus, the opportunity cost of accepting a project also decreases.

With the following assumption, we can impose a tighter structure on the model and drive more results.

Assumption 2. *$u(Y,0,s,\varepsilon) \geq 0$. That is, $u(Y,\omega_i) \geq 0$ when $\phi = 0$.*

Given this assumption, we have that $A_i(B,t) = [0,a_i]$ where a_i is the largest project size player i would approve. Rank the players $1,2,3$ so that $dA_1 \geq dA_2 \geq dA_3$. We can then prove the following lemma.

Corollary 1.2.2. *Under assumption 2, $A^{UNAN} = A_3$ and $A^{MAJ} = A_2$.*

Assumption 2 provides the stronger result that unanimity decision rule admits the same projects as the decision of the most conservative player. In the majority rule, the decision rule is the same as that of the median player. Further note that under this assumption, $A^V(B,t) = [0,\min\{a_c,a_m\}]$ and $A^V(B,t) = [0,\min\{\max\{a_c,a_g\},a_m\}]$.

1.2.3 The (unanticipated) institutional change

In the empirical application, identification is partially provided by a sudden shift in how the OCAD committees aggregate votes. This is the change induced by suddenly changing the aggregation rule from D^V to D^M. I assume that this change is unanticipated. Suppose the SC decision is made at period V. Then, up until V, players will play *as if* D^V will be the aggregation rule till the end of the game. Hence, they compute their continuation values V under this assumption. When the decision is made, players update their continuation values according to the aggregation rule D^M and the new equilibrium is played.

In the notation I introduced before, the admissible set of the aggregation rule D^V is given by:

$$A^V(B,t) = A_m(B,t) \cap A_c(B,t)$$

whereas the admissible set of the aggregation rule D^M is:

$$A^M(B,t) = A_m(B,t) \cap (A_c(B,t) \cup A_g(B,t))$$

1.2.4 When does institutional change matter?

As can be calculated easily from the previous equations, the difference between the two rules is the interval $A_m(B,t) \cap (A_g(B,t) - A_c(B,t))$. This is the set of the projects that the mayor and governor would agree on, but the central government vetoes. It is important to analyze under what conditions we should expect this change in decision rule employed by the committee to impact the project outcomes. I will concentrate on the static change the Supreme Court decision brings. If the within period decision do not change with an institutional shift, we cannot expect to realize different projects over a period of time.

As mentioned before, the difference between the two decision rules is

$$A_m(B,t) \cap (A_g(B,t) - A_c(B,t))$$

. This set is empty if $A_g(B,t) - A_c(B,t)$ is empty. This happens if $A_g(B,t) \subset A_c(B,t)$. That is, if the admissible set of the governor is smaller than that of the central government, no changes will be observed. The other, less plausible, scenario is that $A_m(B,t) = \emptyset$.

In the case with preference shocks, recall that

$$Pr(D^V(\cdot) = 1) = Pr(d_m(\cdot) = 1)Pr(d_c(\cdot) = 1) \tag{14}$$

and

$$\begin{aligned} Pr(D^M(\cdot) = 1) =& Pr(d_m(\cdot) = 1)Pr(d_c(\cdot) = 1) + Pr(d_m(\cdot) = 1)Pr(d_g(\cdot) = 1) \\ & - Pr(d_m(\cdot) = 1)Pr(d_c(\cdot) = 1)Pr(d_g(\cdot) = 1) \end{aligned} \tag{15}$$

Thus, $\Delta D^{M-V} = Pr(d_m(\cdot) = 1)Pr(d_g(\cdot) = 1) - Pr(d_m(\cdot) = 1)Pr(d_c(\cdot) = 1)Pr(d_g(\cdot) = 1)$. Then, as $Pr(d_g(\cdot) = 1) \to 0$ or $Pr(d_c(\cdot) = 1) \to 1$, we will have $\Delta D^{M-V} \to 0$. This is quite intuitive. If center allows every project, taking away its veto right will not change anything simply because no project is ever vetoed. Similarly, if the governor rejects every project, she will never join a winning coalition, and therefore will never be pivotal.

Tying this back to the fundamentals of the model, we can note several things. First, $Pr(d_g(\cdot) = 1) \to 0$ is true if, say, w_g is very large compared to u_g. Or the particular type of the project is much more preferable to the other type. Similarly, $Pr(d_c(\cdot) = 1) \to 1$ is true when w_c is small and u_c is relatively large or the particular type of the project is much less preferable to the other type.

Another crucial finding of the model is that the Supreme Court decision affects each OCAD differently. To see this intuitively, I will retain the assumptions 1 and 2. Recall that, under these assumptions, $A_i(B, X, t)$ is an interval that includes 0. Since I will study the differential impact of the Supreme Court decision, now I will keep X as a determinant of the admissible projects. Under assumptions 1 and 2, as previously derived, we have

$$A^V(B, X, t) = [0, \min\{a_c(B, X, t), a_m(B, X, t)\}]$$

and

$$A^V(B, X, t) = [0, \min\{\max\{a_c(B, X, t), a_g(B, X, t)\}, a_m(B, X, t)\}]$$

.

Therefore, for OCADs with player preferences that imply

$$\min\{\max\{a_c(B, X, t), a_g(B, X, t)\}, a_m(B, X, t)\} = \min\{a_c(B, X, t), a_m(B, X, t)\}$$

institutional change will have no impact. This happens when $a_c(B, X, t) > a_g(B, X, t)$ or $a_m(B, X, t) < \max\{a_c(B, X, t), a_g(B, X, t)\}$. These are the municipalities where the set of admissible projects of the governor is a subset of that of the center. Alternatively, in the OCADs where the admissible set of the mayor is the smallest one, the institutional change won't have any impact as the governor won't be approving any projects that were previously vetoed. To recap, municipalities in the set

$$M_s = \{X | a_c(B, X, t) \geq a_g(B, X, t) \vee a_m(B, X, t) \leq \max\{a_c(B, X, t), a_g(B, X, t)\}$$

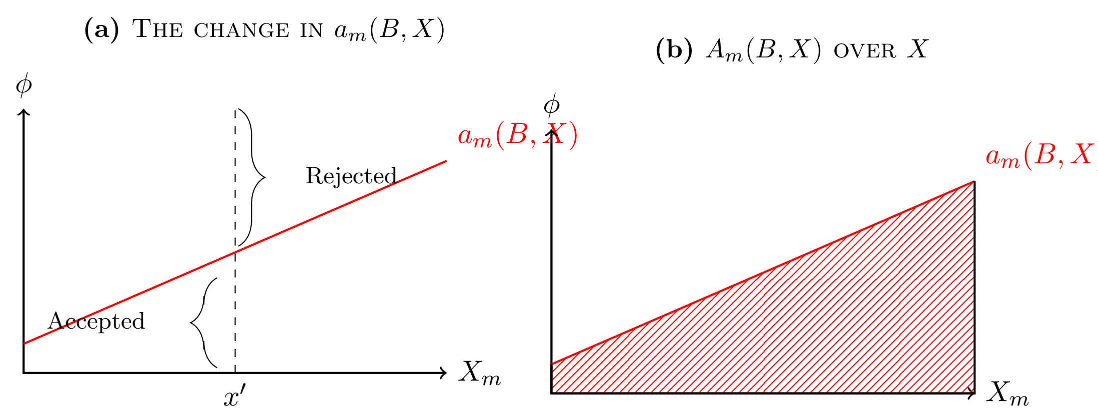
will not be affected by the institutional change. On the other hand, the municipalities in the set

$$M_c = \{X | a_c(B, X, t) < a_g(B, X, t) \wedge a_m(B, X, t) < \max\{a_c(B, X, t), a_g(B, X, t)\}$$

will be affected by the institutional change.

Next, notice that the model imposes a monotonic change in how effective the institutional change is. To see this, suppose that the project draws are the same for each municipality and the municipalities differ only in one dimension, X. Moreover, suppose that the utilities of the players are linear and monotonically increasing in X[9]. One can think about X as a variable that measures the *need* for projects in a specific type. In this case, the set of admissible projects will be increasing in X. This is simply because the continuation value will be increasing slower than the instantaneous utility function in X. This implies that the effect of the institutional change will be changing with X. To visualize a possible scenario, consider figure 1. On the left panel, we see how $a_m(B, X)$ changes with X. Note that the linearity of $a_m(B, X)$ is guaranteed by the fact that I assume u is linear in X. The projects with $\phi > a_m(B, X)$ are rejected and the projects with $\phi \leq a_m(B, X)$ are accepted. This leads to the $A_m(B, X)$ pictured in the left panel of figure 1.

Figure 1
ADMISSIBLE SET OF THE MAYOR

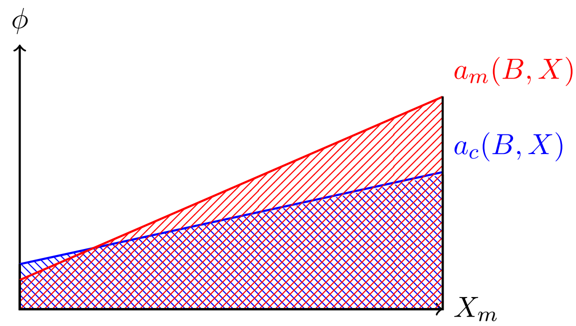


Next, consider the set of projects that will be admitted under D^V. For some arbitrary preferences for the center, figure 2 visualizes the set of projects that will be accepted under the veto power of the central government, conditional on being drawn.

The placement of the preferences of the governor will determine the impact of the institu-

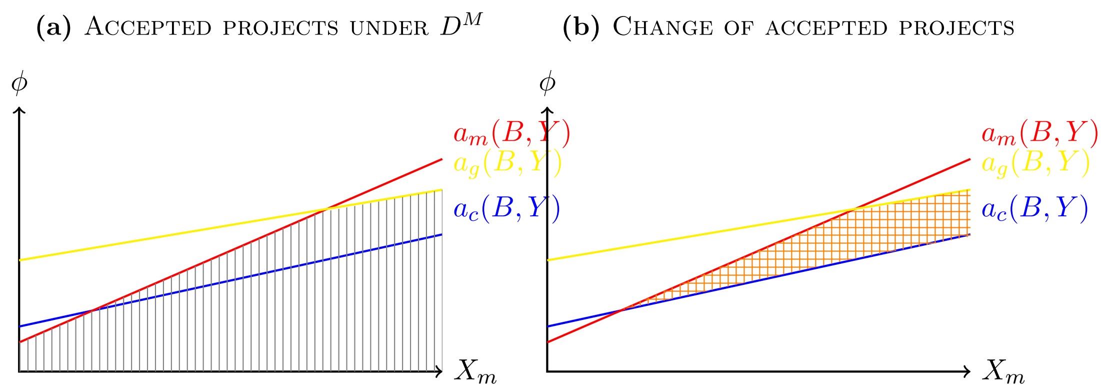
[9] A utility function of the form $u_i(\phi, X) = \alpha_\phi \phi + \alpha_X X$ will satisfy this where $alpha_X > 0$.

Figure 2
Set of accepted projects under D^V



tional shift. In figure 3, one possible such placement is shown. This placement results in an impact that changes with X. For very small X, the institutional change will have no impact. This is because the mayor of these municipalities is the most conservative among the committee members. On the other hand, in the municipalities with larger X, mayors accept larger projects whereas the center does not. Here, introduction of the governor allows the municipalities to pass projects with larger sizes.

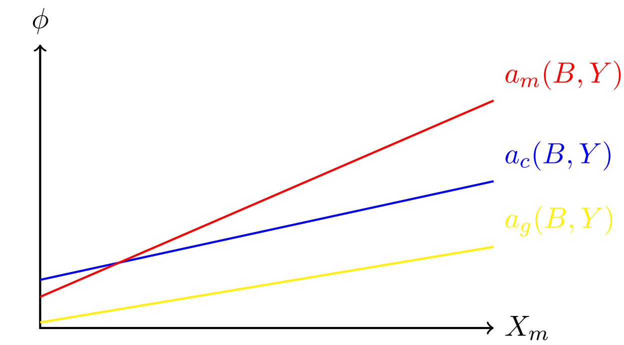
Figure 3
Change in the set of accepted projects after the institutional shift



Notice that other placements of the preferences are possible where the institutional shift has no impact on the accepted projects for any X. For example, in figure 4, the preferences of the governor is such that she is the most conservative player for all X. In this case, governor does not accept any project that center would not, and hence the institutional shift does not imply any structural change in the set of accepted projects.

Figure 4

An Alternative Placement of Preferences



1.2.5 Discussion of the model

The model allows me to study decentralization in a unique way where the committee voting aggregation rules can be thought as different levels of aggregation. A dictatorial aggregation rule represents the extremes in terms of decentralization. When this power is granted to the local government, the model represents a case of full decentralization. Whereas when it is allocated to the center, it represents full centralization. All the decision rules in between, where more than one player's votes are accounted for, we get various degrees of decentralization. This degree can change very narrowly. For example, the decision rule that requires a majority with or without the mayor as the proposer will lead to slightly different degrees, where the former is slightly more decentralized.

It is important to analyze the channels of the model that would imply different investment decisions in the beginning versus towards the end of the cycle, other than the decision rule. This is the advantage of using a model instead of a mean comparison approach. This helps me differentiate the effect of the Supreme Court decision, which is made within a single budget cycle, and other potentials reasons. As lemma 1.2.6 suggest, the set of admissible projects grow toward the end of the cycle. Hence, *any* project is more likely to be accepted toward the end of the cycle than it is in the beginning. Moreover, the size of the budgets, in the beginning and at the end of the cycle are significantly different. Hence, if the average size of the targeted transfer projects is lower (which *is* the case in the data), it could be more likely that the members approve targeted transfer toward the end of the cycle. Hence, the model offers various plausible channels that capture the changes in the investment decisions of an OCAD following the Supreme Court decision.

Another important aspect of the analysis shown above is the possibility of no impact of the institutional change. This result is obtained when the uncertainty introduces by the preference shocks were shut off. When these are introduced back into the model, unless $Pr(d_g(\cdot) = 1) \to 0$ or $Pr(d_c(\cdot) = 1) \to 1$, institutional change will impose changes, albeit possibly small. Notice that with reasonable assumptions on the distribution of the preference shocks, $Pr(d_g(\cdot) = 1) = 0$ or $Pr(d_c(\cdot) = 1) = 1$ will not be achieved in finite X. This implies that the introduction of uncertainty will imply changes at every OCAD. However, the degree with which the changes occur will depend on the characteristics of the municipality.

The model described in this section makes several restrictive assumptions to gain tractability. Most important of these is the finite horizon assumption. This assumption imposes limits to the set of equilibria where some plausible equilibria is eliminated. In a version of the game where horizon is infinite and players vote strategically, they can adopt complex strategies that allows different equilibrium paths through credible punishment. One such interesting case is logrolling. For example, players can first approve the projects that are in line with central government's preferences and then the ones in line with that of the mayor. In such a game, empirically studying the impact of an institutional change is very difficult as the set of equilibria is very large. Thus, I choose to limit the size of the set of equilibria to gain empirical tractability.

I discuss the model in more detail after I present the empirical specification.

1.3 Data

I have collected data from the Colombian Department of National Planning (DNP) for the OCAD projects between 2012-2014. The period from January 2013 to December 2014 constitutes a budget cycle I use for my empirical application. I merge this dataset with the Panel on Colombian Municipalities obtained from the University of Los Andes in Bogota, local and general election results in Colombia and DNP reports on municipal finances and administrations. These datasets contain the main characteristics of Colombian municipalities, including socio-economic variables and finances.

There is a total of 1123 municipalities in Colombia. 105 of these municipalities are dropped in my analysis. 81 of these municipalities have no OCAD budgets. These are small, rich and non-producer municipalities[10]. Moreover, 24 municipalities that have missing covariates are dropped. This leaves 1018 municipalities for my analysis.

1.3.1 Data from the OCADs

DNP keeps record of the sector, date of acceptance, cost and number of beneficiaries for each project. For the period that my data spans, DNP did not centrally collect data on votes at the project level which creates the biggest challenge in the identification of preferences. I neither observe the set of rejected projects. However, due to the structure of the OCADs, where projects are often communicated between the voting parties before the meetings, the actual number of rejected projects in an OCAD meeting is very low. A DNP bureaucrat who casts central votes in about 15 municipalities reported that she has actually seen only one project being rejected. Usually, the to-be-rejected projects are never submitted to a vote in the committee and are tabled before the meeting starts.

A total of 5419 projects were implemented in OCAD meetings in the 2013-2014 cycle I am using. Since each OCAD meets at most every two months, there are 12 possible meetings

[10] Each producer municipality is directly transferred a portion of the royalties it generated. Moreover, as mentioned above, non-producer municipalities are assigned budgets according to their population and need.

for each OCAD. OCADs do not meet in a two month time period either if there are no projects proposed by the municipalities or because there is no budget left for the OCAD to spend on new projects. DNP classifies projects within 20 different sectors[11] I give detailed summary statistics for each sector in the appendix. To ease the estimation procedure, I classify the sectors into two main categories. The first group is the "public goods and services" and the second one is the "targeted transfer goods". In table 1, I present how each sector is classified. The first type contains services like environment, transportation, risk management, education, health that are arguably public goods/services and are relatively harder to direct to a specific group. The other classification, on the other hand, includes sectors like dwellings, social inclusion, sports, business and agriculture. These are sectors that benefit a narrower portion of the population and are relatively easier to direct. For example, the sports projects can be purchasing jerseys for a local soccer team and agriculture could be purchasing machinery for local farmers.

Table 1
CLASSIFICATION

Classification	**Sectors Included**
Public Services & Goods	Environment and sustainability, Communication, Culture, Children, Defence, Education, Urban equipment, Energy, Health, Law, Transportation, Risk management
Targeted transfers	Business and Tourism, Agriculture, Dwellings, Sports, Social inclusion

This classification is in line with the previous literature that lays a dichotomy between the provision of public goods and transfers as a potential way to characterize government expenditure (see, for example, Persico and Lizzeri [2001]). Empirically studying this dichotomy, however, is difficult. The sectors included in the public goods type has potentially

[11] Online appendix provides basic descriptive statistics of the number and cost of projects by sector, respectively. The most common project type is transportation, which is mostly toward building new roads or upgrading and repairing the current ones. The most expensive project sector that we observe is clean water and basic sanitation services, which mostly consists of extending or building sewage or water systems.

broader beneficiaries. Providing electricity, building schools, sewage systems, justice halls and roads could potentially benefit everyone living in the municipality. On the other hand, building a sports complex, buying machinery for local farmers or providing housing for specific people is more directed toward a specific segment of the population. Besides this aspect, the sectors that are grouped under public goods are defined as the primary goals of the municipalities under Colombian law. Municipal government is responsible for water sanitation, elementary education, local electrification and communication, fire departments etc. There is no clear law that indicates that the municipality is responsible for providing housing or farm equipment. Overall, this categorization distinguish between efficient and inefficient redistribution through public expenditure.

Some municipalities implement multiple projects in a given OCAD meeting. In my empirical implementation, I allow an OCAD committee to spend only on one type of project each month. This is not exactly true in the data, since in about 3% of the period-municipality pairs, an OCAD invests in both type of projects. In these cases, I randomly assign projects to periods before and after so that I can aggregate the projects into a single category each period. Hence, the data I use to implement my model has either *public goods* or *transfer* spending for each OCAD meeting period. Moreover, I include the projects done before September 11, 2013 in period 4 and between September 11 and September 30th in period 5. This allows me to have a clear cut between periods of veto and no-veto where the Supreme Court decision is made right between the 4th and 5th OCAD meetings.

In the data, the average cost of a public goods project is \$ 670,000 whereas the average cost of a transfer project is \$ 390,000. Unsurprisingly, the public good projects cost less on average. The average project size does not change dramatically following the Supreme Court decisions. There is a slight decrease in the average sizes of both types of projects. Details on the descriptive statistics for average project costs can be found in table 3. An important insight this table gives us is that the maximum amount spent on a project increases after the revoking of the veto. This is indeed what the model described above would suggest as the set of admissible projects should increase with the addition of a new winning coalition.

Table 2
Descriptive Statistics for Project Sizes Accross Types

Sample	Public				Transfer			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
All	67	133	0.8	1526	40	85	0.56	1527
Before the SC decision	67.5	130	1.1	1288	43	76	0.8	705
After the SC decision	66	135	0.8	1526	39	89	0.56	1527

Note: In 10,000 2014 US Dollars.

Another important aspect of the data is the average number of projects per municipality. Municipalities usually do not adopt many projects: the median number of projects implemented in an OCAD in both types is 1. However, on average, OCADs are more likely to implement a public goods project. Mean number of public goods project is 1.7 versus 1 for the transfer projects. This information will be important when discussing potential reduced form approaches to the data. In essence, the fact that most of the municipalities have very few projects makes it very difficult to employ a differences in differences analysis to understand the impacts of the institutional shift.

Table 3
Descriptive Statistics for Number of Projects per municipality

	Public				Transfer			
	Mean	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Number of Porjects	1.7	1.3	0	8	.99	1	0	6

There are several other statistics I will use in the rest of the paper. I will use monthly total spending by type to assess the fit of my model to the data. Detailed descriptive statistics about these can be found in the online appendix which provides monthly average spending across municipalities and sectors. Figure 5 provides the graphic representation of total spending across periods. Other important statistics to look at are the initial allocated budgets and the remaining budgets after the cycle. As was discussed before, most municipalities do not spend the entire budget allocated. That is the reason I include a saving motivation in the model and it is helpful to discuss how significant this might be in the data. 924 out of 1018 municipalities have left less than $10000 in their budget. However, there is a lot of variation. The maximum amount of budget left in the bank by a municipality is $6.6 million. Table 4 provides details on the initial and remaining budget[12]. Moreover, it is important to note that, for the initial budget, I sum the unspent budget from the previous cycle in 2012[13] and the allocation in the 2013-2014 cycle. The variance in the unspent budget is interesting. The fundamental reason the municipal OCADs leave an unspent portion of their budget is not mainly due to saving concerns. On the contrary, many municipal governments do not have the administrative capacity to create, plan and design projects. Regardless of its size, the laws governing OCADs require sophisticated planning and feasibility reports for each project. Even though DNP provides assistance in these procedures, some municipalities lack the manpower to generate the required reports. Hence, in my empirical specification, I also capture this aspect of the OCAD spending.

Table 4

Descriptive statistics of initial and leftover budget (in 10000 2014 USD)

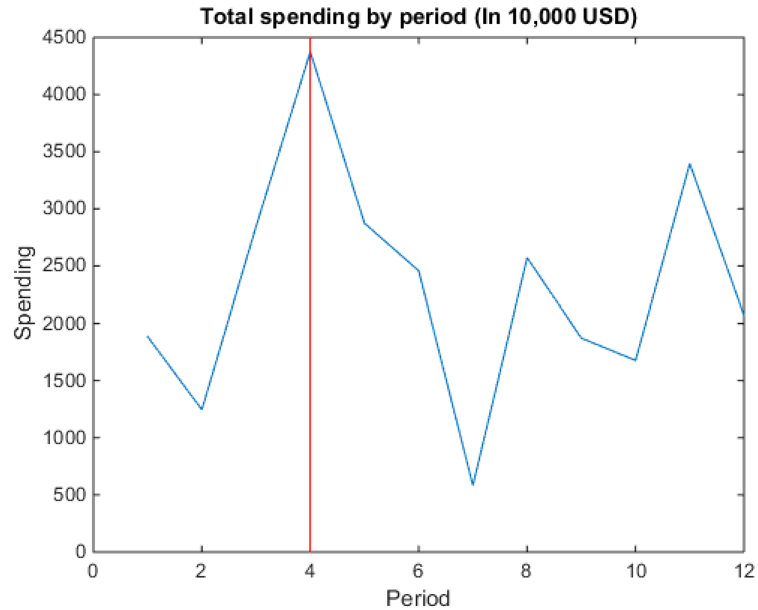
Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Budget	164.834	423.515	1.298	7884.818
Remainder	7.874	45.267	0	660.551
N		1018		

The last thing I want to point out about the data is its high frequency and seasonality. Figure 5 lays out the total spending by month. It can be seen that spending has high frequency and follows a seasonal behavior that is quite small in the beginning of each year. The DNP officers explained this trend by the seasonal responsibilities of the mayors and the governors. During January and February, they have a lot of bureaucratic responsibilities, which often delays project proposals. This implies that the probability a mayor proposes in a given month is not the same and follows a seasonal pattern. Hence, I will need to control for this in the empirical specification.

[12]Online appendix provides additional by-period statistics in the supplements

[13]The 2012 cycle included mostly producer municipalities and was structured differently.

Figure 5



1.3.2 Covariates

The covariates I use for my empirical application come from the larger Municipal Panel Data collected by the University of Los Andes and DNP reports. Hence the covariate data combines several resources, including the Colombian census, ministry reports and fiscal statements of municipalities. In my analysis, I include the most relevant variables I gathered from this data. In the empirical application, I will use each of these variables to explain the preferences toward either public good projects or targeted transfer projects. Table 5 provides the summary statistics and descriptions of the variables used in the empirical analysis. Column 1 gives the category the data will be used towards, 2 gives a short description and the rest are the usual statistics. Besides approximating preferences, I also use administrative efficiency to approximate the ability of a municipality to draw projects, which is also listed in the table.

Below is an itemized description of the variables.

- **Students per primary school:** Children in elementary school age per elementary

Table 5
SUMMARY STATISTICS FOR THE COVARIATES

Category	Description	Mean	Std. Dev.	Min.	Max.
Public Goods	Students per primary school	50	36	10	589
	Previous investment on public goods	0.91	0.04	0.58	0.99
	Municipal debt per capita	28	57	0	959
	Distance from the state capital (km)	123	89	0	926
Transfers	Self-sufficiency	47	22	0	98
	Mayor from president's party	0.24	0.43	0	1
	Mayor from governor's party	0.08	0.28	0	1
	Mayor from a local party	0.215	0.41	0	1
All	Administrative efficiency	0.52	0.5	0	1

school in the municipality. Reported by the Colombian Education Ministry at the end of 2012. Included because it is the most recent variable to capture the public good capital accumulated in a municipality.

- **Previous investment on public goods:** Percentage of total investment in public goods and services in 2012 by the municipal government as reported by municipal balance sheets.

- **Municipal debt per capita:** 2013 (accumulated) per capita debt of the municipal government in 2014 USD.

- **Distance from the state capital:** Distance of the municipality from where the governor resides[14].

- **Self-sufficiency:** Portion of the municipal budget that the municipality generates from its own resources. As this part of the budget is not earmarked, it is a good measure for how much *free* spending is available to the municipality.

- **Mayor from president's party:** A dummy variable indicating whether the mayor belongs to Presindent's party (Social Party of National Unity)

- **Mayor from governor's party:** A dummy variable indicating whether the mayor belongs to governor's party.

[14]In 12 cases, this data was approximated using GPS coordinates.

- **Mayor from a local party:** A dummy variable indicating whether the mayor belongs to a local party instead of a national one. It is equal to one if mayor's party does not run in more than 10 elections.

- **Administrative Efficiency:** DNP, based on a fixed formula, determines the administrative efficiency of each municipality. Used to approximate the probability of coming up with a project.

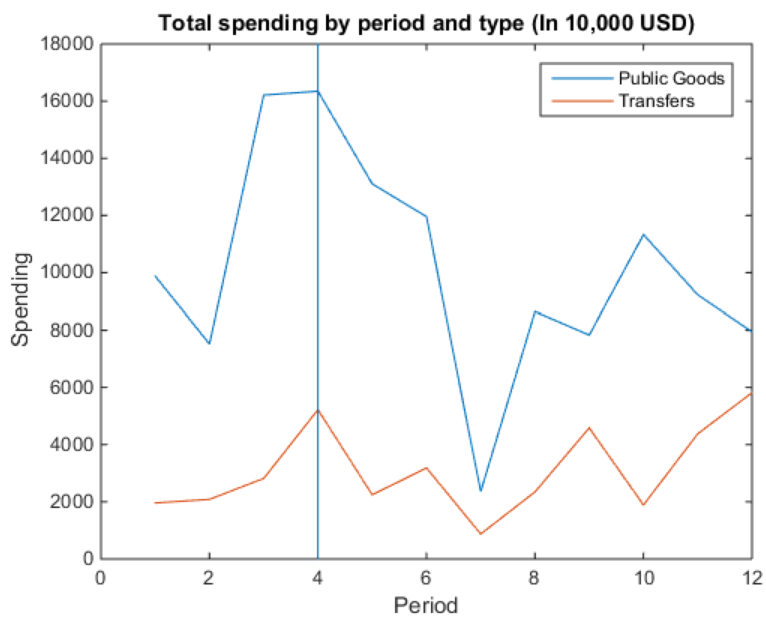
I will further discuss the use of each covariate in the empirical specification section. One thing I'd like to note here is the fact that I use *students per school* to approximate the need of each municipality for public goods. Unfortunately, the last census year in Colombia was 2005. Although this census contains many variables that could better measure the need for public goods, the Ministry of Education of Colombia each year publishes the number of schools and number of students for each Colombian municipality. Since this is the last available measure of public good capital in a municipality, I use it to encompass the need for public good capital at large. I report the correlations between this measure at 2005 and other possible measures of public good need measured in the 2005 census in the appendix.

1.3.3 The effect of the Supreme Court decision

In this section I analyze the changes in the composition of municipality spending after the central government's veto power was revoked by the Supreme Court. I first look at the general trends in the data around the institutional shift. Then, I provide a deeper analysis guided by the theoretical findings from the previous section and the fact that there are strong seasonal trends in the data.

The first thing I study is the evolution of spending around the institutional shift. Figure 5 provides the total amount spent in OCADs by month. There are two things to note about this graph. First, following the Supreme Court decision –marked by the red vertical line–, the spending seems to decrease significantly. Second, the general trend in spending follows a yearly cycle. As mentioned before, this is probably due to the yearly calenders of the mayors. Although the drop in the total spending is interesting, t is unclear whether it is spending toward public goods or transfers or both is the driving force of this drop. To investigate this further, I include figure 6, which shows the total spending patterns for different types of projects. This figure clearly shows that it was the spending in public goods that determined the fall in the overall spending. This could be due to several reasons. For example, most of the mayors may not prefer to spend their budget on public goods, but that was the only type of spending that the central government would not veto. However, it is impossible to distinguish this or other explanations from these graphs alone.

Figure 6



To understand the changes in the portion spent in different types, instead of looking at the overall trends in spending, I consider the share of the spending that goes to public goods across municipalities. Suppose $\phi_{m,t,s}$ is the spending of municipality m in period t to type s. Let $\rho_{m,t} = \frac{\phi_{m,t,pub}}{\sum_{s\in\{pub,tra\}} \phi_{m,t,s}}$, be the share of public good expenditure by municipality m in period t. Next, define the set $N_t = \{m \mid \phi_{m,t,s} > 0, \text{ for some } s\}$. The measure I consider is the average of $\rho_{m,t}$ over active municipalities, which is given by:

$$\mu_t = \frac{\sum_{s \in N_t} \rho_{m,t}}{|N_t|} \tag{16}$$

There are couple of things to note here. First, I take the averages across municipalities and not the entire spending portions. This is because, as suggested by the theory, municipalities are affected in different ways from the Supreme Court decision. This measure will allow me to differentiate between municipalities using critical covariates. Second, there are alternative measures that could be used instead of ρ. The desirable feature of ρ is that it epitomizes all the relevant information about a municipality-time observation into a single number. One aspect of this measure is that it does not take into account the OCADs that did not approve projects in a given time period. While taking the data at the face value, this is difficult to avoid since the data does not provide any information as to *why* the OCAD did not spend. Was it the mayor who did not propose a project or was the proposed project rejected by the committee? Thus, by excluding those municipalities, ρ provides a clean measure that epitomizes the spending decisions of active OCADs across sectors.

Figure 7 provides the graph of ρ_t over time and the linear fit before and after the Supreme Court decision. The graph suggest a change in trend after the Supreme Court decision, yet it is highly inconclusive. The jump observed at the institutional shift is statistically insignificant.

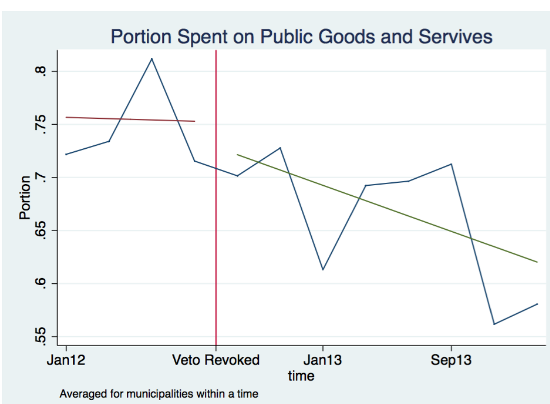


Figure 7

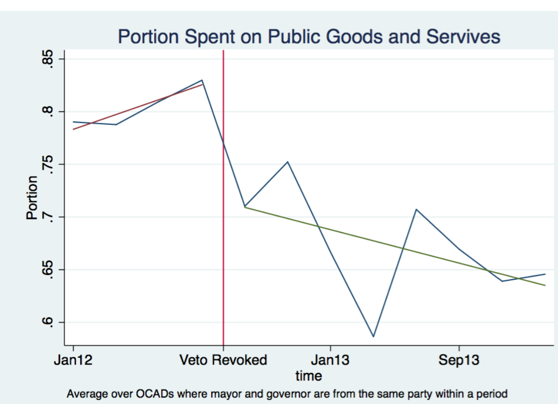


Figure 8

Recall that the theoretical section underlined that each OCAD will be distinctly effected by the change of institutions. This effect depends on the preferences of the players. We know that, in the veto period, governor does not contribute to the investment decisions. However, after the Supreme Court decision, mayor and governor can form a winning coalition and pass a previously unfeasible project. Thus, we should expect the shift in the institutions to effect mostly the municipalities that have favorable characteristics to the governor. The covariate I consider in that direction is whether the mayor and the governor are from the same party. We should expect that the municipalities with connections to the governor should face larger compared to the rest.

The average share of total spending on public goods and services by municipalities with mayors sharing the same party with the governor over time is given in figure 8. Comparing this to the figure 7, the drop is sharper, as the theory suggests. This drop is characterized by both a higher spending on public goods in the veto era and a lower spending on public goods after the revoking of the veto power. Thus, the municipalities with connections to the governor have a higher drop rate compared to the rest.

These graphical analyses, however, can be misleading for several reasons. First, as suggested by the graphs, there are strong seasonal changes in total spending and share of it that goes to a particular type. Both spending and public goods spending increase in the beginning of the year and gradually decrease thereafter. Second, the data has a high frequency, implying that a month to month comparison is not very informative. In fact, the drop around the Supreme Court decision is insignificant, probably due to the small sample size. I address the first issue by comparing the first nine months of 2013 (in which the central government held their veto power) to the first nine months of 2014 (in which the veto power was revoked). This allows me to isolate the same periods within a year and avoid comparing the first nine months of the year to 15 months which includes end-of-the-year period which evidently behaves different than the rest. Moreover, I address the frequency and sample size issues by aggregating municipal spending in the periods I consider. Hence, instead of looking at the monthly variation, I take the aggragate spending in nine-month periods.

There are certain things to note. First, not all municipalities are active in the periods I consider. Hence the following results are local to the active municipalities in the periods considered. Second, the remaining budget for the municipalities change between these periods. As such, they cannot be treated as a control versus treatment group. It could be that OCADs tend to spend their money on directed goods when not much is left in the bank. These could be dealt with a more structural approach, and I will address these when implementing my model.

To understand the precise statistics I report, let $\Phi_{m,V,s} = \sum_{t\in\{Jan13,\ldots,Sep13\}} \phi_{m,t,s}$. Similarly, let $\Phi_{m,V,s} = \sum_{t\in\{Jan14,\ldots,Sep14\}} \phi_{m,t,s}$. This implies that $\phi_{m,V,s}$ defines the expenditure of a municipality on type s during the first 9 months of 2013 (when the central government had the veto power) and $\phi_{m,NV,s}$ is the expenditure of a municipality on type s during the first 9 months of 2014 (when the central government did not have a veto power). Next, define $\rho^*_{m,k} = \frac{\Phi_{m,k,pub}}{\sum_{s\in\{pub,tra\}} \Phi_{m,k,s}}$ for $k \in \{V, NV\}$. Hence, ρ^* is the share of total spending that goes into public goods for a municipality m in era k. This implies that $1 - \rho^*$ is the portion spent on targeted transfers.

Table 6 provides the means of $rho*$ in the veto and no-veto eras, and the t-statistic of these means' difference. The fall in the share spent on public goods is 6 percentage points, and it is statistically significant. Note that the t-statistics is computing using unpaired samples. That is to say, I treat the mean of the share of expenditure on public goods in different eras as two independent observations. This is due to the fact that, as mentioned above, the median number of projects by a municipality is 1 in either type. This implies that there are not many municipalities that are both active in the January-September 2013 *and* the January-September 2014 periods. For robustness purposes, paired statistics are given in the appendix.

Even though the result is suggesting a significant effect of the Supreme Court decision, some caution is necessary. There is no reason to believe that the difference is the *causal* impact of the change in institutions. To the contrary, the OCAD decisions in the January-September

Table 6

CHANGE IN SHARE SPENT ON PUBLIC GOODS, ρ^*

Statistic	Veto Period	No Veto Period	Difference
Mean	.74	.68	.06***
St. Dev	(.38)	(.41)	(.02)
N	542	454	

Note: The t-statistic computed using unpaired samples, assuming the samples were independent and had equal variances.

2014 period are not the exact counter-factuals of the ones in the January-September 2014 period. Besides having different active municipalities in these periods that causes a selection issue, it is also important to note that the same OCAD is facing different problems in each of these periods. Remember that the Supreme Court decision happens within a specific budget cycle. Thus, the decision of a committee with a lot left in the budget cannot be the same with when it has little remaining. Due to the dynamics of the problem and the selection issue, we need to exploit the behavioral predictions of a model to uncover the true effects of the institutional change. In fact, after estimating the parameters of my model, I revisit this specific statistic I show and provide a deeper analysis of the effect of the institutional change.

Effect of the institutional change in different municipalities

As I mentioned before, theoretical analysis shows that the shift in institutions will affect the municipalities depending on the preferences of the players. Hence, the fall in the share spent in public goods shall have differential effects depending on the characteristics of the players. There are two particular cases that I consider here. First, I consider municipalities which are favorable to the central government. We should not expect large changes in their spending, because these municipalities, with the central government, always had their winning coalition in the OCAD. I consider two covariates to understand the proximity of a mayor to the central government. First is whether the mayor is from the party of the president. The other important factor the central government might consider is the debt of municipalities because of the long-lasting fiscal problems in Colombia. We would expect the central government to be more lenient toward municipalities with good fiscal records.

Next, consider the other extreme, where the governor and the mayor could have winning coalitions. These municipalities are expected to have large jumps in their spending portions, as these mayors can get into winning coalitions with the mayors to approve projects they previously could not. I use two covariates to measure the relationship between the mayor and the governor. First is whether they are from the same party. The second is the distance between the municipality and the state capital. For this last measure, a priori, it is hard to predict the effect of the institutional change. The jump for remote municipalities can be low if the relationship of the mayor and the governor gets better as the distance shrinks. This can happen, for example, because they see each other more frequently. Alternatively, the governor can be more lenient toward distant mayors as they pose less threat to his seat. However, if this measure truly approximates a relationship between the governor and the mayor, we expect it to imply different levels of changes at the extremes.

Table 7 provides results on these measures and is supportive of the theory. I use the entire sample, and the sample within specific groups to test the ideas developed above. The first column describes the sample created through restrictions. The first row is the entire sample. The second and third rows are party proximity of the mayor and the president. The forth and fifth rows are the same for the governor. Sixth and seventh rows are no debt and high debt municipalities respectively. Here, a high debt municipality is defined as having a debt per capita higher than the sample mean. Finally eight and ninth rows are the sample of municipalities that are close to or far from a major departmental city. A remote municipality is defined to have a higher distance than the departmental average. For each of the samples I provide $\bar{\rho}^*_{m,V}$ in the third column, $\bar{\rho}^*_{m,NV}$ in the forth column and $\bar{\rho}^*_{m,V} - \bar{\rho}^*_{m,NV}$ in the last one. The standard deviations are also given. Similar to my previous analysis, standard deviation of $\bar{\rho}^*_{m,V} - \bar{\rho}^*_{m,NV}$ is measured assuming the samples V and NV are independent and their variances are the same. Robustness checks with different assumptions are in the supplemental materials[15]. In general, the results are robust.

The effects are as the theory suggested. The shift in the share spent on public services in the overall sample is about 0.6 and is significant. As expected, the change in the spending of the municipalities with mayors sharing same party as the president is insignificant. It is, however, significant for the mayors that are not members of president's party. We also have clear evidence for the trend the theory expects for the governor's party. The municipalities where the mayor and the governor belong to the same party have a statistically significant 19 percentage points fall in the share they spend on public services. The fall diminishes to 5 percentage points when this is not the case. For municipalities with no debt, as the theory suggests, there is no significant shift in spending. However, for the municipalities with high debt, there is a significant shift above that of the entire sample. Changes in the remote municipalities, on the other hand, is significant and slightly above the 6 percentage point mark whereas the close-by municipalities have low changes that are statistically insignificant. This suggests that the governor gets more lenient toward the remote municipalities.

1.4 Estimation and Identification

1.4.1 Specification

First I start by specifying the instantaneous utility function u and the utility from the savings w. The preferences toward a project depends on the type of the projects and the characteristics of the municipality. I parametrize u as follows:

$$u_i(t, s, \phi, Y, \varepsilon) = (\theta_{i,s} Y_{i,m,s} + \varepsilon_{i,t,s})\phi \tag{17}$$

Hence, the data I use to estimate player i's utility from a project with type s is included in $Y_{i,m,s}$. This vector consists of a constant and two covariates that is type and player specific.

[15]See appendix for the same table without the equal variance assumption.

Table 7
DIFFERENCE STATISTIC WITHIN SOME SUB-SAMPLES

Sample	Statistic	Veto Period	No Veto Period	Difference
All	Mean	.74	.68	.06***
	St. Dev	(.38)	(.41)	(.02)
	N	542	454	
President and mayor same party	Mean	.77	.69	.06
	St. Dev	(.34)	(.44)	(.05)
	N	135	122	
President and mayor different parties	Mean	.75	.68	.07**
	St. Dev	(.39)	(.41)	(.02)
	N	416	330	
Governor and mayor same party	Mean	.80	.61	.19**
	St. Dev	(.34)	(.44)	(.08)
	N	52	44	
Governor and mayor different parties	Mean	.74	.69	.05**
	St. Dev	(.34)	(.44)	(.02)
	N	490	410	
No debt	Mean	.72	.67	.06
	St. Dev	(.40)	(.42)	(.03)
	N	253	223	
High debt	Mean	.77	.69	.08*
	St. Dev	(.37)	(.41)	(.04)
	N	157	123	
Far from a major city in the state	Mean	.75	.67	.08**
	St. Dev	(.38)	(.41)	(.03)
	N	265	201	
Close to a major city in the state	Mean	.75	.69	.05
	St. Dev	(.38)	(.42)	(.03)
	N	277	253	

Note: The difference statistic computed using unpaired samples, assuming the samples were independent and had equal variances.

Table 8 introduces the data I use for each player-type pair. The important thing to note about this specification is that it is very parsimonious. The linear structure of the model uses very few parameters to impose a structure on the utilities.

I measure *public good capital* using the elementary school age population per elementary school. This is the most recent measure available and is highly correlated with other measures of public good capital in 2005[16]. Public good capital captures the overall need for

[16]I discuss this variable more detail in the appendix.

Table 8
DATA USED

	Sector	
Player	**Public**	**Transfer**
Mayor	Public goods capital Recent public investment	Self-sufficiency Mayor from local party
Center	Public goods capital Debt per capita	Self-sufficiency Mayor from president's party
Governor	Public goods capital Distance from the state capital	Self-sufficiency Mayor from governor's party

public good projects in each municipality. Moreover, the mayor cares about the recent public investment made in her municipality, measured as a percentage of total municipal budget spent on public goods and services. This is important in terms of the preferences of the mayor, as a higher spending toward public goods may signal a preference toward public goods. On the other hand, the center cares about the debt per capita in a municipality. Decentralization led to a serious municipal debt crisis in the early 2000s and many municipalities were bailed-out by the central government. Hence, the central government's relationship with the mayors firmly depend on the debt of the municipality. Lastly, the governor values the municipalities' distance from the state capital when deciding on a public goods project. This can be interpreted either as a proxy for the relationship between the mayor and the governor; or, alternatively, it is possible that governor have differential preferences toward public goods depending on the distance.

As for the targeted transfer projects, I assume that each player values the *self-sufficiency* of the municipality. This is the share of the municipal budget that is generated from local resources like taxes. This portion of the municipal budget is not earmarked and hence can be spent in any way. Then, the mayors that generate most of their resources are able to transfer resources to targeted groups when they need to. In return, this implies that rejecting a transfer project from such a municipality can be useless, as the mayor can find other resources to implement the project. Moreover, the mayor's preferences depend on whether she belongs to a local party. Local parties are quite common in Colombian local politics. They usually operate within couple of municipalities or sometimes a single state. There is a tendency to think that these local parties belong to clientelistic networks and belong to the local elites, which are more prone for transfers. The center and the governor, on the other hand, value whether they share the same party with the mayor when evaluating a project of targeted transfers. These variables control for *vertical* clientelism, in that, governor or the center could be more lenient toward the mayor when they are politically connected.

Lastly, I assume that the preference shocks are distributed independently across players, periods and types. That is, $\varepsilon_{i,t,s} \overset{iid}{\sim} N(0,1)$. Independence is admittedly a very restrictive assumption. If one thinks of the preference shocks as an unobserved quality of the project, it will likely be correlated across players. Moreover, distributions can potentially depend on types (for example, transfer projects may have systematically higher variance than the public goods ones). However, this assumption improves tractability and provides computational speed. Moreover, the identification of preference shock parameters is difficult given that they interact with the level of the utility imposed by the preference parameters.

Next, I discuss the specification of the utility from saving the function, which I take as homogeneous and simply an identity function. Thus,

$$w_i(x) = x \tag{18}$$

Note that this assumption is just a *normalization.* Since any project is ultimately compared to the saving option, it would be difficult to identify any extra parameters in the savings utility in addition to the instantaneous flows. Homogeneity assumption is not strong and it keeps the parsimonious structure. This is simply another normalization as the heterogeneity in saving tendency will, in fact, be captured by the *level* of utilities.

For the project arrivals, I assume a three-fold process. First, with probability $\Pi^{draw}(D_{admeff})$, a municipality is *able* to draw a project. Here D_{admeff} is a dummy variable indicating whether the municipality received high rates from the DNP in their bureaucratic quality. I choose this specification because some of the municipalities have very few projects, and this, in general, is correlated with their administrative efficiency, a measure DNP assigns to each municipality yearly. $\Pi^{draw}(D_{admeff})$ is specified by:

$$\Pi^{draw}(D_{admeff}) = \frac{exp(\theta_{d1,t} + \theta_{d2} D_{admeff,m})}{1 + exp(\theta_{d1} + \theta_{d2} D_{admeff,m})} \quad (19)$$

$\Pi^{draw}(D_{admeff})$ depends on time because of the seasonal patterns shown in the data section. I assume $\theta_{d1,t'} = \theta_{d1,t'+1}$ for $t' = \{1, 3, 5, 7, 9, 11\}$ to keep the parameter space parsimonious.

If a project was drawn, next, it's type is determined randomly. The type is public goods with probability Π^{pub}. I assume this probability is homogeneous across municipalities and constant within periods. It is given by:

$$\Pi^{pub} = \frac{exp(\theta_{type})}{1 + exp(\theta_{type})} \quad (20)$$

Finally, the size of the project is given by $\phi_{t,s} = \tilde{\phi}_{t,s} B_{m,t}$ where $\tilde{\phi}_{t,s} \sim Beta(1, \tilde{\beta})$ and $B_{m,t}$ is the available budget in period t. Here, I parametrize the distribution using the equation

$$\tilde{\beta}(B_{m,t}) = \frac{2}{\pi} b_{1,s} \arctan\left(\frac{B_{m,t}}{b_{2,s}}\right) \quad (21)$$

Here $b_{1,s}$ controls the maximum $\tilde{\beta}$ can achieve whereas $b_{2,s}$ controls the speed with which it achieves that value. The larger $b_{2,s}$, the slower $\tilde{\beta}$ approaches $b_{1,s}$ as $B_{m,t}$ grows. Note that the mean of the random variables drawn from $Beta(1, \tilde{\beta})$ *increases* as $B_{m,t}$ decreases. Hence, this specification allows the possibility that OCADs with small budgets will have to spend a large portion of their budget into projects and vice versa. This is important, because when municipalities have small budgets left, they tend to spend a larger portion of it to the projects. Therefore, I capture the importance budget plays in determining the size of the project.

Since it is not identified, I set $\beta = 0.98$. This gives me a total of 30 parameters to estimate. 18 of these are preference parameters and the remaining 12 are distribution parameters.

1.4.2 Discussion of the empirical model

My modeling approach impose several restrictions on players' behaviors. Moreover, there are certain aspects of the setting that I choose not to model, which also grants discussion. In this section, I go through these and discuss their potential impact on the parameters estimated.

Finite horizon

As discussed while presenting the model, the finite horizon assumption has important implications with respect to the possible equilibria of the game. For example, an equilibrium with logrolling where one type of projects are implemented first and the other second is ruled out. In finite horizon, players cannot credibly commit to stick with such a strategy.

The OCADs continued through to 2015 with the same players. However, six months after the span of my data ends, ta local election took place. This is particularly important because the term-limits for mayors and governors in Colombia is a single cycle. Therefore mayors and governors in the OCADs necessarily changed soon after my data ends. However, it is not uncommon that a strongly connected candidate runs in place of the incumbent and represent its legacy.

It is also useful to think about the implications of the alternative model with infinite horizon. In this class of models, instead of a utility from savings, players would value the funds in the bank because of the future projects. At the end of each cycle, OCADs would draw another budget and continue with the same stage game. Now, in this setting, consider a municipality with low administrative efficiency. For this municipality, the expected utility from the next budget cycle would be lower than that of the saving motive finite horizons model. This is because it is very hard to draw projects for this municipality in the next cycle as well and hence a dollar has a different value to it than it does for an efficient municipality. This would imply that the inefficient municipalities would be *more* willing to take on a project at any time as their opportunity cost of forgoing a project is lower. The finite horizon assumption does not completely undermine this effect, however. In any period other than the last one, inefficient municipalities have lower opportunity costs simply because it is less likely for them to draw projects in the future. However, in the last period, finite horizon model assumes an equal opportunity cost for all types of mayors.

Moreover, the infinite horizon model would be able to generate logrolling. However, notice that in this class of models, the set of equilibria is very large and it is empirically impossible to differentiate the effect of Supreme Court decision from that of a potential logrolling.

Presidential Election of 2014

The first round of Colombian Presidential Election of 2014 was held on May 25, 2014. This date coincides with the span of my sample. In particular, it is after the central government lost its veto power. Both this election and the second round on June 15th were close elections which resulted in the reelection of President Santos. It is not clear whether Santos used the OCADs as a method of clientelism. It is possible that the central government acted more lenient in accepting transfer projects around the election to collect the support of local mayors. If this channel truly exists in my sample, this would upward bias my estimate of the impact of the veto.

No anticipation of the Supreme Court Decision

An important assumption I make is that players are not able to anticipate the change in the institution. There are several reasons to believe that this was actually true. A lawyer, who was not affiliated with any local or state government, filed a suit to the Supreme Court suggesting that the central government's veto power was unconstitutional. This case did not make the headlines in the media until the actual Supreme Court decision. In fact, most of the DNP bureaucrats were not even aware of the case until after the decision was made. It is also very unlikely any mayors anticipated the decision, as the lawyer was based in Bogota and was not representing any entity.

It is still useful to discuss the implications of this assumption. Suppose that a mayor anticipates the decision. If her preferences are aligned with that of the center, anticipation will not have an impact on the approved projects. On the other hand, if the preferences of the mayor and the center is different, she will be more likely to *not* to propose projects while the central veto lasts. Proposing these projects in the non-veti period will have a higher probability of success with the vote of the governor. As will be discussed in the next section, most mayors prefer transfer (public goods) projects more (less) compared to the central government. Hence, if all the mayors had anticipated the decision, we'd expect mostly the municipalities whose mayors prefer public goods to be active in the veto period. Hence, anticipation by the mayors will upward bias the estimate of veto's effect.

Alternatively, if the central government anticipates the Supreme Court decision, its continuation payoff will be lower. This is straightforward to see, as there will be a positive probability of future projects that the center would otherwise veto. Under some assumptions, this will imply that the representative will be more lenient toward projects –that is, it will be more likely to accept projects that it did not before. Hence, the bias on the estimate of veto's effect would be downward.

1.4.3 Identification

Several characteristics of the data make identification very difficult. The first of these is that I only observe the accepted projects. Second, I do not observe individual votes of the players. On the other hand, the institutional shift *and* several exclusion restriction assumptions I make helps identification. To get a clear understanding of the use of the institutional shift, I present a formal identification method in the appendix. However, this has to be done by making strong assumptions. I assume out the dynamic aspect of the game, and suppose that I have *cross-sectional* observations from two sets of infinitely many OCADs, one group with the veto aggregation mechanism and other with the majority. Moreover, I assume that the municipalities differ only on one dimension, X, with infinite support. Finally, to focus on the preference parameters, I assume that the distribution of the projects are known. Under these assumptions, I can show that the observations from the data with veto power provides the identification of the preference parameters of the center and the mayor. However, labeling these parameters to a specific player is impossible using only this information. To solve this, I exploit the dataset with the majority aggregation rule. The intuition behind the identification argument relies on taking X to infinity. I show that each possible labeling generates a different rate of convergence of the probability of observing specific projects as X approaches to infinity in the two datasets. This differentiates each possible set of labeling and provides the identification of the exact preferences of the center and the mayor. Using the labels and the information from the accepted projects with the majority aggregation rule, I finally show that I can recover the preference parameters of the governor. Section 1.8 provides the formal details on this argument.

In the actual estimation, I use exclusion-restrictions to identify the parameters of my model. As mentioned in the previous section, in the public sector utility, public good capital affects every player. On top of this, the mayor (and only the mayor) cares about the previous expenditure on public goods. Center values the per capita debt of the municipality relative to the rest of the country and the governor is concerned about the distance of the municipality from where he resides. Hence, a change in any of these variables will effect the probability of voting only through the relevant player. Similarly, for the utility driven from private transfers, ability of a municipality to generate its own resources effects every player. Whether the ruling party is local effects only the mayor and sharing the same party effect the center and the governor. Please see table 9 for the exclusion restriction summary. Note that similar ideas for the distribution of the projects hold.

Before going into the formal representation of identification, to see the use of these variables in a more intuitive level, note that not observing the rejected projects may imply that we can simultaneously increase the probability of a project being drawn and decrease the probability of acceptance of a project. However, changing the administrative efficiency effects the probability of drawing a project but not the probability of acceptance. Similarly, not observing the votes may imply that two different sets of preference parameters can generate the same probability of acceptance for each project. Yet, keeping all the variables same and only changing, say, the party affiliation affects the probability of observing a project only through the preferences of the relevant player.

To make this argument more formal, consider the probability of observing a specific project ω. There are two components to this: the probability of drawing ω and the probability that it is accepted by the committee given the aggregation rule. The first component is proportional to the multiplication of the probability of drawing a project, probability that the project is of a particular type s and the probability of drawing a project with size ϕ

$$\Pi^{draw}(D_{admeff}, t)\Pi^{s} f_{Beta}(\phi|B_t, s)$$

The second component, on the other hand, depends on the aggregation rule. It is equivalent to both the center and the mayor accepting the project in the first 4 periods and to the mayor and either the center or the governor accepting thereafter. It is given by:

Table 9
List of Exclusion Restrictions

Variable	Player
Public Sector:	
Students per school	All
Portion of public investment	Mayor
Debt per capita	Center
Distance from the governor	Governor
Private Sector:	
Self-sufficiency	All
Loacl party	Mayor
Share same party	Center
Share same party	Governor
Probability of Project Proposal:	
Administrative Efficiency	All

$$Pr(D^*(t,B,\omega)=1)=\begin{cases} Pd_m Pd_c + Pd_m Pd_g - Pd_m Pd_c Pd_g & \text{if } t>4 \\ Pd_m Pd_c & \text{otherwise} \end{cases} \tag{22}$$

where $Pd_i = Prob(u(\omega) > \beta O(B,t,\phi,Y,X))$.

Hence, the probability that I observe a specific project is:

$$Pr(\phi_{t,s}) \propto \Pi^{draw}(D_{admeff},t)\ \Pi^s\ f_{Beta}(\phi_t|B_t)\ Pr(D^(t,B,\omega)=1) \tag{23}$$

This equation helps the understanding of identification through exclusion restrictions. Similar equations are easy to drive for the non-veto period and alternative type of projects. The first term,

$$\frac{exp(\theta_{d1,t}+\theta_{d2}D_{admeff,m})}{1+exp(\theta_{d1}+\theta_{d2}D_{admeff,m})}$$

depends only on period and administrative efficiency and does not depend on type or the size of the project. The second term, $\dfrac{exp(\theta_{type})}{1+exp(\theta_{type})}$, on the other hand, is constant across periods and depends on the type. $f_{Beta}(\phi_t|B_t, b_{pub})$ depends only on the budget and the type. Finally, preference parameters depend on $Y_{i,m,pub}$, budget and the period.

To have a clear idea as to the use of exclusion restrictions, let $Z_t(X)$ be the set of municipalities that share characteristic X at period t. Next, consider the statistic

$$\frac{Pr(\phi_{pub} > 0|Z_t(Y_{i,m,s}, D_{admeff}, B_t))}{Pr(\phi_{tra} > 0|Z_t(Y_{i,m,s}, D_{admeff}, B_t))}$$

for all t and $\phi > 0$. This is the odds of observing public goods versus a transfer project, given a period t and same characteristics. This statistic contains information only on θ_{type} because these municipalities share the ability to draw a project as well as the probability to pass a project within a given time period. Next, consider the statistics:

$$\frac{Pr(\phi_t > 0|Z_t(Y_{i,m,s}, B_t))}{Pr(\phi_t > 0|Z_t(Y_{i,m,s}, B_t)))}$$

for all t and $\phi > 0$. This statistic is the odds that a project will be observed in an efficient versus inefficient municipality, given the same characteristics. Since now we know θ_{type}, this statistic contains information only on θ_{d2} since the municipalities in $Z_t(Y_{i,m,s}, B_t)$ differ only in the dimension of administrative efficiency. This implies that θ_{d2} can be identified through this statistic. Next, the statistic:

$$\frac{Pr(\phi_{t,s} > 0|Z_t(Y_{i,m,s}, B_t))}{Pr(\phi_{t+2,s} > 0|Z_t(Y_{i,m,s}, B_t))}$$

for all odd t, all s and $\phi > 0$ will provide information only on $\theta_{d,t}$. Hence, at this point, I showed that we can identify $\Pi_{draw,t}$, Π_{type}. Next, notice that the budget a municipality has left at period t, B_t, affects both the distribution of project sizes and the opportunity cost of a project. However, keeping the budget constant and changing $Y_{i,m,s}$, one by one at each t will provide information on the preference parameters. For example, to identify the preference parameters of the central government, the statistic

$$\frac{Pr(\phi_{sec}|Z(Y_{mayor,m,s}, B))}{Pr(\phi_{sec}|Z(Y_{mayor,m,s}, B))}$$

for $t \leq 4$ and all ϕ is used. Similar computations are possible for other players. Finally, knowing the preference parameters, identification of $b_{1,s}$ and $b_{2,s}$ is possible using the observed distribution of project sizes.

Thus every single parameter effects the probability of observing a specific statistic in a distinct way, which provides identification. Notice that the preferences of the governor can only be recovered through the observations in the no-veto period ($t > 4$).

Moreover, as mentioned earlier, the preference parameters can be recovered using the differential changes in the accepted projects before and after the Supreme Court decision. To see this, for example, consider the statistic:

$$\frac{Pr(\phi_{sec,t}|Z(Y_{mayor,m,s}, Y_{center,m,s}, B))}{Pr(\phi_{sec,t'}|Z(Y_{mayor,m,s}, Y_{center,m,s}, B))}$$

where $t \leq 4$ and $t' > 4$.

In this statistic, $Pr(\phi_{sec,t}|Z(Y_{mayor,m,s}, Y_{center,m,s}, B))$ will not change with the preferences of the governor and hence is constant up to D_{admeff}. However, the denominator,

$$Pr(\phi_{sec,t'}|Z(Y_{mayor,m,s}, Y_{center,m,s}, B))$$

changes with the preferences of the governor and hence the only variables that effect this statistic is the preference parameters of the governor. Similar statistics can readily be found for other players.

Thus, the preference parameters are *over-identified*. They can be identified due to the exclusion restrictions and the fact that the effect of the Supreme Court decision is differential on each municipality.

1.4.4 Likelihood and the Estimation Method

I use maximum likelihood estimation techniques to estimate the parameters of my model. I discuss the intuition behind the likelihood construction here and leave the exact derivation to the appendix. For completeness, although it has been derived above, consider the probability of observing a specific project. With the current specification, the probability a specific project $\omega = \left(\phi, s, \{\varepsilon_{i,t,s}\}_i\right)$ will be accepted in the committee is:

$$Pr(D^*(t,B,\omega)=1) = \begin{cases} Pd_m Pd_c + Pd_m Pd_g - Pd_m Pd_c Pd_g & \text{if } t > 4 \\ Pd_m Pd_c & \text{otherwise} \end{cases} \tag{24}$$

where $Pd_i = Prob(u(\omega) > \beta O(B,t,\phi,Y,X))$. The precise probability a project ω will be accepted by player i is

$$Pr(\varepsilon_{i,t} \geq \beta O(B,t,\phi,Y,X)\phi^{-1} - \theta_{i,s}Y_{i,m,s}) = 1 - \Phi(\beta O(B,t,\phi,Y,X)\phi^{-1} - \theta_{i,s}Y_{i,m,s})$$

. Then, the probability of observing a project of size ϕ in sector s at time t becomes:

$$Pr(\phi_{t,s}) = \Pi^{draw}(D_{admeff}, t)\Pi^{Pub} f_{Beta}(\phi_t | B_t) Pr(D^*(t,B,\omega)=1) \tag{25}$$

Next, I look at the likelihood of observing *no* project at a given t. This is possible due to one of the two reasons. Either the mayor failed to come up with a project, which happens with probability $(1 - \Pi^{draw}(D_{admeff}, t))$ or the drawn project was rejected by the committee. The latter happens with probability $(1 - Pr(D^*(t,B,\omega) = 1)$ for a specific project. The probability that *any* project was rejected is $\int_{\tilde{\omega}} g_\omega(\tilde{\omega}|B)(1 - Pr(D^*(t,B,\omega) = 1)d\tilde{\omega}$ where g_ω is the implied distribution of ω[17] Hence the contribution of observing no projects implemented

[17] Recall that the random variables in ω are ϕ and ε.

in a given period to the likelihood is:

$$
\begin{aligned}
Pr(\omega_t = \emptyset) = \quad & (1 - \Pi^{draw}(D_{admeff}, t)) \\
& + \sum_{s=pub,tra} \Pi^{draw}(D, t)\ \Pi^{s} \left(\int_{\tilde{\omega}} g_{\omega}(\tilde{\omega}|B)(1 - Pr(D^{*}(t, B, \omega) = 1) d\tilde{\omega} \right)
\end{aligned}
\tag{26}
$$

The first part of this equation is the probability that the mayor failed to propose a project whereas the second part is the probability that *a* drawn project was rejected.

Equations 25 and 35 provide the basis of the likelihood. For the complete derivation, see the appendix. Estimation of the game relies on the methods developed by the dynamic discrete choice literature. I can solve the game backwards for every set of parameters and compute the equilibrium of the game. Then, according to the functions defined in this section, I can compute the likelihood.

1.5 Results

Tables 10 presents the estimates on preference parameters and table 11 presents results on the distribution parameters. These parameters are difficult to interpret on their own. Even though the exact amount has to be determined by the specific characteristics of a municipality, in table 10, one can see that the public goods projects are, in general, more preferred by the central government compared to the other players. One easy way to see this is to look at the constant term in the utilities. This parameter is reported in the first row of each player panel and is 1.88 for the center, 1.61 for the mayor and 0.78 for the governor. On the other hand, for the transfer projects, mayor has the largest constant utility parameter at 1.5, followed by center's 1.11 and governor's 0.8.

It is also possible to make two important observations from table 11. First, we see the seasonal trends to be very important in the estimates for the ability to draw. Note that as the estimate for a particular period decreases, the ability of a mayor to draw a project decreases as well. Hence, the observation of low number and size of projects in the beginning of the fiscal year is reflected by the smaller estimates for $\theta_{d,Jan-Feb}$. Moreover, estimates suggest that administrative efficiency is an important determinant of the ability of drawing a project. Estimates also show that drawing a public goods project is more likely than drawing a transfer project. Finally, estimates on the size of the cost shows that, given the same budget, the cost of drawing a public good project will be higher on average.

As mentioned, these tables are difficult to interpret without thinking about them though a particular municipality. This is what I address in the next section. Specifically, I consider a hypothetical median municipality to understand the implications of the estimated parameters.

Table 10

PREFERENCE PARAMETER ESTIMATES

Type	Player	Variable	Estimate	95% Confidence Interval
Public Good Projects	**Mayor**	Constant	1.61	(1.23, 1.99)
		Public Good Capital	0.07	(0.03, 0.11)
		Recent Investment	0.01	(0.004, 0.016)
	Center	Constant	1.88	(1.45,2.31)
		Public Good Capital	0.29	(0.13, 0.45)
		Municipal Debt	0.003	(0.002, 0.004)
	Governor	Constant	0.78	(0.52 , 1.04)
		Public Good Capital	0.28	(0.23, 0.32)
		Distance from Capital	-0.002	(-0.005, 0.001)
Targeted Transfer Projects	**Mayor**	Constant	1.5	(1.3, 1.7)
		Self Sufficiency	0.005	(0.001, 0.009)
		Local Party	0.7	(0.6,0.8)
	Center	Constant	1.11	(0.97, 1.25)
		Self Sufficiency	0.006	(0.001, 0.011)
		Party Affiliation	0.82	(0.71, 0.93)
	Governor	Constant	0.8	(0.4, 1.2)
		Self Sufficiency	0.0006	(-0.003, 0.0036)
		Party Affiliation	0.91	(0.74,1.08)

Table 11
Distribution Parameter Estimates

Distribution	Variable	Estimate	95% Confidence Interval
Ability to Draw	Jan-Apr 13	-1.3	(-2.13, -0.47)
	May-Aug 13	-0.5	(-0.1, 1.1)
	Sep-Dec 13	-0.8	(-0.1, -1.5)
	Jan-Apr 14	-2	(-3.2, -0.8)
	May-Aug 14	-1.4	(-2.63, -0.17)
	Sep-Dec 14	-0.6	(-1.41, 0.21)
	Administrative Efficiency	1.4	(0.3, 2.5)
Type of the Projects	Public Goods	0.19	(0.15, 0.23)
Size of the Cost	$b_{1,pub}$	3.6	(1.1, 4.9)
	$b_{2,pub}$	204	(42, 366)
	$b_{1,tra}$	15.7	(3.1, 28.3)
	$b_{2,tra}$	712	(341, 1083)

1.5.1 Interpretation of the estimates

I begin by presenting the results on the distribution of project draws. Tables 12 and 13 describe the results on the estimated distributions of projects draws. The first set of results, shared in table 12, show that the ability of the mayors to propose projects differ significantly across periods. Note that these are constant across all municipalities up to the administrative efficiency of the municipality..In fact, estimates show that the efficiency of the municipal bureaucracy is an important determinant of the ability to draw projects. An efficient municipality is almost twice more likely to draw a project. In the periods within January-April 2013, an efficient municipality has a 52% chance of drawing a project whereas an inefficient one draws projects only 21% of the time. These numbers are 71% and 38%, respectively, in the May-August 2014 period. The period where the municipalities are least likely to draw a project is the January-April 2014 period and they are most likely to draw a project in Sept-Dec 2014. Next, table 13 describes the probability of drawing different types of projects. Structural estimates suggest that drawing a public good project is more likely, however the difference is not stark. A drawn project is a public goods one with probability 0.55 and a transfer project with probability 0.45. Finally, 14 describes the beta distributions for the draws of the project sizes. I find that the mean size of a public goods project is larger than the mean size of a targeted transfer project. This difference increases as the budget increases. The mean public project size is $350,000 when drawn from a median sized budget ($560,000). The same statistic for a transfer project is $300,000. The difference between these expectations increase as the budget increases. When an OCAD with a budget at the 75th percentile draws projects (slightly above $1million dollars), the numbers are $590,000 and $440,000 respectively.

Table 12
Distribution of projects

	Type of Municipality	**Jan-Apr**	**May-Aug**	**Sept-Dec**
Probability of drawing a project in 2013	Efficient Administration	52%	71%	65%
	Inefficient Administration	21%	38%	31%
Probability of drawing a project in 2014	Efficient Administration	35%	49%	69%
	Inefficient Administration	11%	19%	36%

Table 13
Probability of types, conditional on drawing one

Public good project	**Transfer project**
55%	45%

Table 14
Draw of project sizes

	Budget	**Public**	**Transfer**
Mean draw	25 pctile budget (30)	23	21
	Median budget (56)	35	30
	75 pctile budget (114)	59	44

Next, I present the interpretation of the estimation results of the preference parameters. The first thing I study is the tendencies of different players to accept specific types of projects. To do this in an intuitive way, I present the results concentrating on a hypothetical *median* municipality. This municipality has the median characteristic in every characteristic dimension. That is, all the characteristics used to approximate the preferences of players over the types of public good or transfer projects are taken to be at the median level. Moreover, administrative efficiency –which is a dummy variable– is taken to be $\frac{1}{2}$. I consider an OCAD located at this municipality with a median budget ($ 560,000) deciding on whether to approve a mean draw of each type of project presented in table 14. As it is important to have an understanding of the effect of each variable, while keeping everything else at the median level, I report the acceptance probability for each player when a single characteristic is moved to a high or a low level. That is to say, suppose we are interested in examining the impact of a change in public goods capital on the probability of accepting a public good project. To do this, I keep every other characteristic at the median level and examine the probabilities of acceptance at hypothetical municipalities where the public goods capital is at the 90th or the 10th percentile of public good capiral observed in the data. Table 15 presents the probabilities of acceptance for public goods projects for each player whereas table 16 represents the table for the transfer goods projects. Each table reports the probability that a player accepts the corresponding project type at its mean cost, taken as the mean probability a project is accepted during the veto and non-veto periods. Note that changing a variable has two distinct effects. First, it effects the utility obtained from a specific type of project, which has a direct impact on the probability of accepting that type of projects. Second, by influencing the ratio of the utilities a player can obtain from the different types, it effects the probability of acceptance of the other type of projects. To capture all of these effects, the tables have four panels. In the upper panel, I consider the effect of changing the variables that are included in the preferences over public good projects in the veto period (the probability of the governor is not reported in this panel). In the second upper panel, same is done for the variables concerning the preferences toward transfer projects. The other two panels provide the same for the non-veto periods, respectively. In both tables, the first column reports the player. The second column provides the probability that each player accepts the certain type of project under consideration when all the variables are at their median. In the following columns, I replace variables one by one to their *high* (90th percentile) or *low* (10th percentile) levels.

There are several things to note about this exercise. The probability a player accepts a project in the veto and non-veto periods changes due to two distinct effects. First, altering the aggregation rule affects the continuation value of the players as it changes the set of admissible projects. Second, as players get to the end of the game, the probability they accept the projects increase as was proved in the theory section. I will look at the latter effect in more detail in the following section.

Analyzing table 15 provides important details about the preferences of the players. The straight-forward observation is that the central government is more likely to accept a public good project compared to the mayor under both regimes. For example, when all the variables are at their median levels, mayor accepts a mean sized public goods project with 0.55 probability whereas the central government accepts them with probability 0.82. The difference between these probabilities decrease when the municipality has low public capital[18], high recent investment in public goods and low debt. Moreover, we see that the preference parameters used to predict the utilities from transfer projects also impact the probability a player accepts public good projects. For example, if the mayor and the president share the same party, center is less likely to accept a public good project of a mean size (as his utility from transfer good projects increase) and the mayor is slightly less likely to do so as well (as his continuation payoff changes). Moreover, a mayor with low fiscal independence is also more likely to invest on public good projects. Next, considering the preference of the governor toward public good projects, we see that they lay somewhere between that of the center and the mayor, albeit are closer to that of the center. This is true more so when the the municipality is close to the state capital. Moreover, as expected, preferences of the mayor and the governor are more aligned when they are affiliated with the same political party.

Analysis of table 16 reveals the preferences of different layers of government toward transfer projects. On average, the mayor is more likely to accept a transfer project compared to the center. In the veto period, the mayor of the *median* municipality is 26 percentage points more likely to accept a drawn transfer project compared to the center. Moreover, center is very lenient toward transfer expenditure when the mayor is from the same party with

[18]Note that public capital is calculated using number of schools age children per elementary school. Hence, a larger estimate on this parameter implies that low public good capital is related with higher utility.

Table 15
PROBABILITY OF ACCEPTING A MEAN SIZE PUBLIC GOOD PROJECT

VETO

	All Median	Pub Cap H	Pub Cap L	Prev Inv H	Prev Inv L	Debt H	Debt L	Distance H	Distance L
Mayor	.55	.51	.58	.57	.54	.55	.56	.56	.55
Center	.82	.79	.86	.82	.83	.87	.80	.83	.83

	Fis Ind		Local	Pres	Gov
Mayor	.54	.58	.51	.54	.56
Center	.82	.84	.83	.80	.83

NO VETO

	All Median	Pub Cap H	Pub Cap L	Prev Inv H	Prev Inv L	Debt H	Debt L	Distance H	Distance L
Mayor	.67	.67	.70	.69	.66	.68	.67	.69	.67
Center	.92	.90	.95	.92	.93	.94	.90	.93	.92
Governor	.80	.76	.85	.80	.80	.81	.80	.77	.86

	Fis Ind		Local	Pres	Gov
Mayor	.67	.69	.65	.68	.68
Center	.92	.93	.93	.91	.93
Governor	.80	.80	.81	.81	.77

the president or when the municipality is fiscally independent. In fact, the probability of accepting a mean sized transfer project is almost the same when the mayor and the president share the same party. In general, the disagreement between the mayor and the center toward transfer projects decreases as the public capital in the municipality increases, municipal debt decreases, and, as mentioned, they share the same party. Governor, on the other hand, is the least likely player to accept a transfer project. This, however, completely changes when the governor and mayor share the same party: in this case they have around 64% chance of accepting a transfer good project at mean size. However, in general, the preferences of the governor is very much aligned with the center. This is true especially when there is no alignment of political parties, the municipality is far away from the state capital, the municipal debt is low and the municipality cannot generate its own resources.

Table 16
PROBABILITY OF ACCEPTING A MEAN SIZE TARGETED TRANSFER PROJECT

VETO

	All Median	Pub Cap H	Pub Cap L	Prev Inv H	Prev Inv L	Debt H	Debt L	Distance H	Distance L
Mayor	.51	.52	.50	.51	.51	.51	.51	.51	.51
Center	.25	.29	.21	.25	.25	.19	.28	.25	.25

	Fis Ind H	Fis Ind L	Local	Pres	Gov
Mayor	.56	.47	.73	.49	.51
Center	.30	.21	.25	.51	.25

NO VETO

	All Median	Pub Cap H	Pub Cap L	Prev Inv H	Prev Inv L	Debt H	Debt L	Distance H	Distance L
Mayor	.63	.64	.63	.63	.65	.64	.63	.63	.64
Center	.44	.47	.42	.45	.44	.39	.47	.44	.45
Governor	.35	.38	.33	.36	.35	.36	.36	.37	.33

	Fis Ind H	Fis Ind L	Local	Pres	Gov
Mayor	.69	.60	.83	.63	.63
Center	.50	.40	.44	.71	.45
Governor	.37	.35	.36	.37	.65

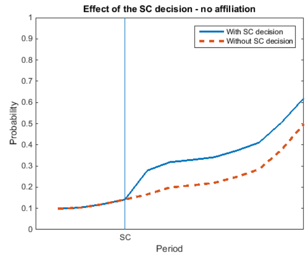
It is interesting that the center and the governor are very close in terms of their preferences. The first order implication of this finding is that the change in the admissible set of projects after the Supreme Court change should not be very large. Or in other words, had the governors' preferences were closer to that of the center's, we would have seen a much larger shift in the types of investments following the Supreme Court decision in September 2013.

Even though these statistics are very useful in analyzing the preferences of players toward different types of OCAD projects, the Supreme Court decision would have differential impacts on municipalities. To investigate this, we should differentiate the dynamic change in the probabilities of accepting a project from the effects of the change in the aggregation rule. This is exactly what next section does.

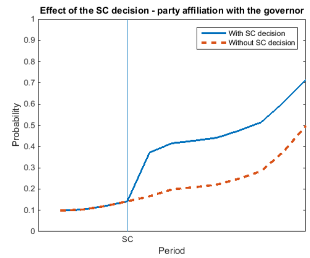
What drives the voting behavior?

The estimates presented in the previous section shows several things. Among them is the finding that the mayor is more likely to accept transfer projects compared to the center. Hence, the fundamental effect of the Supreme Court decision will be through the probability of accepting transfer projects. To see the change in the acceptance rates of transfer projects, I consider three municipalities with the same (median) characteristics, except their party affiliation. In figure 9, I plot the probability that a mean sized transfer project passes from the OCAD committee, conditional on political affiliations. Specifically, the first figure 9a in the panel presents the statistic for a municipality that is affiliated neither with the president nor with the governor. The same is done for an OCAD where the mayor and the governor share the same party in figure 9b and where the mayor and the president share the same party in figure 9c. In each figure, the x-axis is the periods and y-axis is the probability a transfer is accepted. The dotted red lines represent the probability that OCAD accept a mean sized transfer project had the central government kept its veto power whereas the blue solid lines are the actual probabilities estimated. The difference between these lines, hence, is the *dynamic effect of the institutional change.*

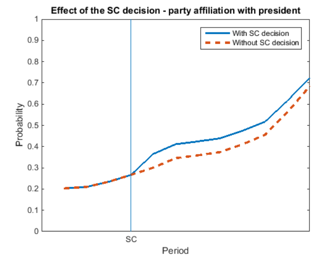
There are several things to note. First, there is a clear difference between the impact of the institutional change on *all* of the municipalities. This is expected. Remember that the probability of accepting a project in the veto area is equal to the probability that the center *and* the mayor accept the project at the same time. In the non-veto period, on top of this, there is the possibility that the mayor *and* the governor form a winning coalition. Hence, the probability of accepting any project in the non-veto period is higher. However, the



(a) No affiliation



(b) Governor mayor affiliated



(c) President Mayor affiliated

Figure 9

Changes in the probability of accepting a transfer project with different affiliations

OCADs are affected in different magnituted. The OCADs where the mayor and governor share the same party is effected the most whereas the OCAD where the mayor and the president are from the same party is effected least. This is very intuitive given the results presented in tables 15 and 16. I have mentioned in the analysis of these tables that these affiliations were critical in understanding the preference alignments of the players. Indeed, the figures clearly lay out the ideas put forward there. The impact of the Supreme Court decision is the lowest in the OCADs with a mayor that shares the same party with the president because the projects the mayor do not need the vote of the governor to pass projects. On the other hand, the OCADs where the mayor share the same party with the governor are the ones that have the largest change in the probability of accepting a project because the governor and the mayor can pass projects that were previously vetoed by the central government.

Lastly, there is a dynamic dimension of the effect of the institutional shift. As the players are more likely to accept projects as the game nears an end, the impact of the Supreme Court decision shrink as the periods grow. However, this result might have been different in an infinite horizon game.

1.5.2 Fit of the model

I consider four statistics to assess the fit of my model. The first two are the average number of projects by the OCADs and the mean project size. These are separately analyzed by the aggregation rule and types of the projects. I also consider the share of spending on public goods in the veto and non-veto eras, and the entire data. Finally, to assess the dynamic fit of the model, I consider the total expenditure by the type of projects over periods.

Table 17 provides the fit of the model for the first three statistics. In general, model fits the data very well. The mean number of projects for both types of projects and aggregation rules is predicted well by the model. The predicted numbers are always within 5 percentage deviation from the data. For the average project size, the model is able to fit the data well in the veto period, however, it significantly under predicts the average size of projects in the non-veto period. This is because the model predicts a higher total spending in the veto period. Hence, in the non-veto area, model leaves less budget to the municipalities than the amount they have data.

Table 17
FIT OF THE MODEL 1

	Period	**Public**		**Transfer**	
		Data	Model	Data	Model
Mean number of projects	Veto	1	.96	.71	.67
	No Veto	.60	.64	.28	.24
Mean project size	Veto	67	69	43	44
	No Veto	66	59	40	36

Moreover, model's fit to the share of spending on public goods is very well. The model predicts the veto era statistic to be 76%, whereas the real value is 74%. Similarly, the share of public good spending for the first nine months of the non-veto period is predicted to be 66% compared to the data moment of 69%. The same statistic using the entire months in the non-veto era is 65% whereas the data value is 64%. Finally, the predicted share spent for the entire sample is 70% compared to the data moment of 68%. Hence the model tends to very slightly over-predict the overall share of spending on public goods. However, the fit of shares of spending to the data is exceptionally well in general.

Table 18
FIT OF THE MODEL 2

		Data	Model
Share of spending on public goods	Veto	.74	.76
	No Veto (First 9 moths)	.69	.66
	No Veto (All months)	.64	.65
	Entire Sample	.68	.70

Figure 10 and 11 plots the dynamic fit of the total spending on public good and transfer projects, respectively. The model predicts the dynamic trends relatively well. However, the expenditure on public goods is under predicted. The expenditure on the transfers are, on the other hand, are predicted relatively good for the veto period although the dynamic fit of the post-veto period is inconsistent. The average spending is however is predicted very well.

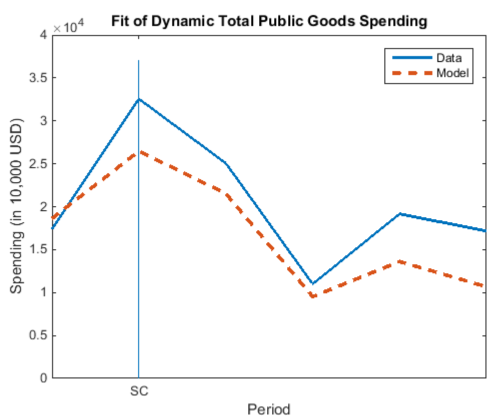


Figure 10

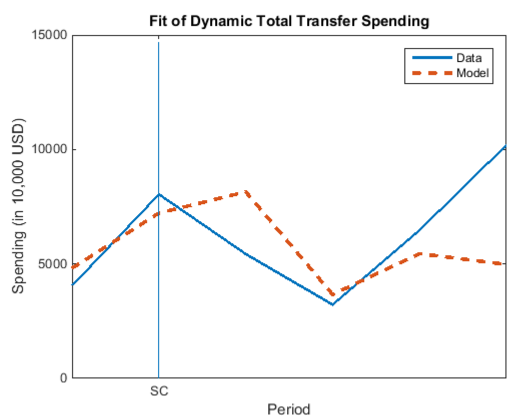


Figure 11

1.6 Public investment under alternative decentralization institutions

In this section, I present the results of the counterfactual analysis I perform. First, I examine the the true effect of the institutional shift and compare it to the mean comparison exercise performed in the data section. Next, I consider the variation in public investment choices under alternative decentralization institutions.

There is an important limitation to the counterfactual exercises I can run. In the data, all the projects are proposed by the mayor. This implies that the arrival distributions of projects are mayor specific. That is to say, had the central government designed the projects to implement from the allocated budget, they could prefer, say, much larger projects. There is no plausible way to deduce this distribution from the available data. However, I still do provide centralization results keeping in mind the very strong restriction of constant project draws imposes on my analysis.

The rest of the analysis relies on the data reported in table 19. This table provides the mean project size and mean project number for each types of investments under alternative voting aggregations governing the OCAD. Moreover, the share of expenditure on public goods are reported under each decentralization institution. The first column reports the statistics for the data. The second column is for the case where a complete decentralization mechanism is used and the third column is for the complete centralization. The last two columns are for he cases with no changes in the institutions and with no central veto for the entire budget cycle, respectively.

1.6.1 The true effect of the institutional shift

I compute the true effect of the institutional shift by comparing the share of expenditure on public goods in the counterfactual where veto power was never revoked (4th column of table

Table 19
COUNTERFACTUALS

	Sector	Data	Full Dec.	Full Cent.	No change	Majority
Mean project number	Public	1.71	1.25	1.81	1.76	1.67
	Transfer	0.94	1.15	0.81	0.90	0.99
Mean project size	Public	67	67	72	69	68
	Transfer	40	49	39	39	42
Share of expenditure on public goods		.68	.57	.78	.73	.68

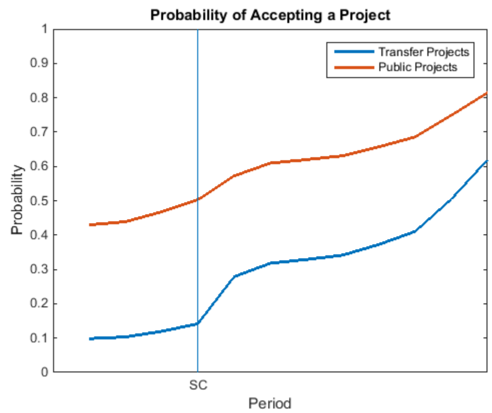
19) and the counterfactual where the majority rule was used in all the periods (5th column of table 19). The voting aggregation that assigns a veto power to the central government has 73% share of public goods expenditure whereas the system with majority voting has 68%. Thus, the difference of the share of expenditure on public goods is 5 percentage points between the two systems. This is less than the mean comparison performed in the data section, which was 6 percentage points.

This implies that the structural estimate is about 17% smaller than what the mean comparison suggests. As mentioned, some of the impact deduced from the mean comparison exercise is due to the dynamic effects that is not possible to distinguish with these methods. It becomes more likely to invest on transfer projects as the budget cycle comes to an end. Hence, my model shows that only 83% of the impact found through mean comparison is actually attributable to the shift in the institutions. The rest is due to the dynamic incentives of the players.

In fact, a quick analysis of figure 9a also outlines this dynamic effect. The probability of accepting a project increases as it gets closer to the end of the cycle. This effect will be more punctuated when the baseline probability of acceptance is lower, which is the case for the transfer goods and projects. This is clearly seen in the following figure 12. Here, the

Figure 12

Dynamic probability of accepting different types of projects



line in red is the probability a public good project is accepted and the line in blue is the probability a transfer good project will be accepted in a median municipality. It is clearly seen that the probability that a project is accepted increases as the end of the budget cycle approaches. This effect is larger for transfer good projects. Thus, even in the absence of a Supreme Court decision, the share of budge spent on public goods would decrease in the last part of the budget cycle.

1.6.2 Different decentralization mechanisms

Complete Decentralization

Complete decentralization coincides with the voting mechanism that takes only the vote of the mayor into account. Hence, this is a mechanism that assigns the decision of the mayor as the decision of the committee. Since mayor relatively prefers the transfer projects to the public goods, the expectation is that in completely decentralized system more of the budge would be allocated on transfer goods. However, there is also another subtle effect. Since the preference shocks are independent across the players, it is always more likely that the mayor will accept a project than will the mayor *and* the center. This is simply because $A \times B < A$ when A and B are less than 1. A priori, it is unclear which effect will be the dominant one.

Results are given in the second column of table 19. We see that the mean size of the public goods projects remain same at \$670,000 whereas the mean transfer project size increase from \$400,000 to \$500,000. Hence the increase is about 20% for the transfer project size. Moreover, the average number of public good projects enacted by the OCADs decrease from 1.71 to 1.25. In contrast, the number of projects enacted in targeted transfers increased from 0.94 to 1.15. These together imply that the share of expenditure on public goods must decrease. In fact, only 57% of the expenditure is on public goods, compared to the 68% in the data.

Thus, decentralization increases the amount of expenditure on transfer goods substantially. This effect happens both in terms of the sheer number of projects and the size of the projects. Thus a complete decentralization of the OCAD system would imply a larger chunk of the royalties to be spent on targeted transfers instead of public goods.

Complete Centralization

As mentioned above, the counterfactual where the center obtains the entire power cannot be approximated with the data at hand precisely. However, for completeness, I present the results regarding centralization in the third column of table 19. I find that, under centralization, the mean size and number of public goods projects would increase whereas the transfer projects would decrease relative to what we observe in the data. Moreover, the share of expenditure that goes to public goods would increase to 78%.

The mean number of projects increase from 1.71 to 1.81 for public goods projects, whereas the mean number of transfer projects decrease from 0.94 to 0.81. On the other hand the size of an average public prohect increased from \$670,000 to \$720,000. The decrease in the size of the transfer projects, on the other hand, was more subtle. This decreased from \$400,000 to \$390,000. The impact of this centralized version is due to the fact that the center is much more likely to accept a project draw that belongs to public goods.

Intermediate degrees of decentralization

Majority rule with mayor as the proposer. One of the voting aggregation mechanisms we observe in the data is the majority rule with the mayor proposing. In fact, this is the rule that was in use after the Supreme Court decision. Under this rule, the mayor is always in the coalition that passes a project because she is the proposer. Moreover, either the governor or the mayor has to vote in favor of the project. Thus, this voting mechanism is somewhere between complete decentralization and complete centralization. However, since the vote of the center is not necessary to pass a project, this system can be argued to be closer to the full decentralization case than it is to the centralized case.

The counterfactual statistics for this case is given in the last column of table 19. The changes between this system and what we observe in the data is not stark. This is reasonable, because two thirds of the sample was actually under this regime. The number of public good projects slightly decrease whereas the number of transfer projects slightly increase. On the other hand, the size of the projects remain roughly the same. Of course, since not much has moved in these variables, the share of expenditure on public goods also remained in the same level.

Majority rule with mayor as the proposer and central veto right. Under the voting rule where the central government has the veto power and the mayor has the proposal power, every successful project has to be approved both by the center and the mayor. Compared to the previous voting rule I analyzed, this rule is a closer to full *centralization* as it requires the approval of the central government for any project to be successful.

In fact, even though we observe this voting rule in the first third of the sample, had the decision by the Supreme Court was never made, the outcomes in the OCADs would change considerably. Both the size and the number of the public goods projects increase relative to the data, whereas the opposite is true for the targeted transfer projects. Hence, the share of expenditure on public goods also increases relative to the data, from 73% to 68%.

1.7 Concluding Remarks

Political decentralization can be carried out at many different degrees and layers. This depends on the specific diffusion of power to different levels of the government specified by the decentralization institutions. I take this heterogeneity in decentralization institutions seriously and develop a theoretical and empirical methodology to assess its impacts on public investment choices. I implement my model using a unique institution employed by the Colombian government to distribute the royalties it gathers from natural resources. This is the first paper that carries out such an analysis. It is argued that it is crucial to understand the preferential differences between layers of government over the types of public investment to successfully carry out such analysis. This is a first step toward the analysis of the heterogeneity of decentralization institutions. Hence, the framework I offer extends the existing empirical literature on political decentralization. In fact, the paper implies that designing successful decentralization institutions depends on a clear understanding of the preferences of stakeholders and how each possible decentralization institution aggregates them.

The main result of the paper is that complete decentralization would result in a decrease if the share of expenditures on public goods and services in Colombia. This drop is more significant for the municipalities that have mayors in local parties or those that share the same party with the state governor. Moreover, this difference would increase with the stock of the public good. This implies that the municipalities with lower stocks of public goods will continue investing on public goods despite gaining the entire control over their budget. One other result of the paper is that the governors have similar preferences to the central government, which implies that the Supreme Court decision regarding the voting rule employed by the OCAD committees had less effect than would have complete decentralization.

Another interesting result is that institutions do not *always* impact economic outcomes. In the theoretical section, I laid out conditions that implied no change in the composition of public investments following the Supreme Court decision that changed the voting mechanism used in the OCADs. Hence, the debates on institutions and their economic implications have to consider the various stakeholders and their preferences. Even though one intuitively expects the institutions to be a main determinant of the outcome, my theoretical analysis shows that, depending on the preferences of the decision makers, institutions can lead to same or very similar economic outcomes. This has also been verified in the empirical section, where I showed that the OCADs where the mayoral preferences aligned with that of the center did not face substantial changes in the composition of their investments.

The method I outline in this paper is easily applicable to other countries as long as a clear mapping from the institutions to an aggregation mechanism can be found. As mentioned earlier, this paper is a first step toward understanding the heterogeneity of decentralization institutions and hence opens up new and fertile questions. My approach has taken the preferences of different layers of government as granted. Further research on the sources of this heterogeneity of preferences would be an immediate avenue to broaden this research agenda. Moreover, it is critical to build a theoretical and empirical method to study the process of design of decentralization institutions. This would require, of course, the analysis of public welfare due to the public investments. When the next Colombian census is available, welfare implications of different degrees of decentralization institutions studied in this paper would be an important avenue of investigation.

1.8 Appendix to First Chapter: Formal Identification of a Simpler Model:

Several characteristics of the data make identification very difficult. First is that we observe only the accepted projects. Second, we do not observe individual votes of the players. And lastly, we have a dynamic setting which ties each decision to the expectation of future projects etc.

To simplify the identification argument, I will assume out the dynamic aspect of the game now and simplify the model to make the identification argument more formal. Consider a setting where we have data of the decision of a committee available from two periods, with no dynamic behavioral relation among them but with different aggregation rules. In both periods, committee members vote on projects that they draw from a distribution[19]. To further simplify the setting, suppose that there is only one kind of project and the preferences of player i toward a project of size ϕ is given by

$$\alpha_{0i} + \alpha_{1i}X + \tilde{\alpha}_{2i}\phi + \varepsilon_i \tag{27}$$

Players also value the savings in a linear way. Hence, player i accepts a draw (ϕ, ε) if and only if

$$\alpha_{0i} + \alpha_{1i}X + \tilde{\alpha}_{2i}\phi + \varepsilon_i \geq B - B + \phi = \phi \tag{28}$$

Letting $\alpha_{2i} = \tilde{\alpha}_{2i} - 1$, we have that a project is accepted by player i if and only if

$$f(\bar{\alpha}) = \alpha_{0i} + \alpha_{1i}X + \alpha_{2i}\phi \geq -\varepsilon_i \tag{29}$$

Hence, we have $d_i(\phi, X, \varepsilon) = \mathbb{1}\{\alpha_{0i} + \alpha_{1i}X + \alpha_{2i}\phi \geq -\varepsilon_i\}$

Moreover, I assume that $\phi \sim F$ and $\varepsilon_i \sim G$. I'll specify these distributions as need be. In the two periods of observations we have, the decision rule is different. In lieu with my setting, I assume:

[19] One can justify this setting in many ways. For example, think about having a large number of municipalities at each X, and they only get to choose once.

$$D^V(\{d_i\}_{i\in N}) = \begin{cases} 1 & \text{if } d_p = 1 \text{ and } d_c = 1 \\ 0 & \text{otherwise} \end{cases}$$

The second one is the majority rule with a proposer p:

$$D^M(\{d_i\}_{i\in N}) = \begin{cases} 1 & \text{if } d_p = 1 \text{ and } \sum_i d_i \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

Given the distribution of X, define the object

$$(\phi|X,\theta,D) \to \rho(\phi|X,\theta,D)$$

where θ is the set of all parameters.

Geometrically, one can think about this system quite simply as linear lines (except the parts where ϕ is equal to one or zero). When the drawn $\phi + \varepsilon_i$ lies below this line, i accepts the offer and when it is above i rejects. Hence, conditional on being drawn, a *phi* that lies below the line has a higher probability of being accepted than a ϕ above the line.

Definition 1. *Let θ^* and $\theta^{'}$ be two sets of parameters. θ^* is identified relative to $\theta^{'}$ if $\theta^* = \theta^{'}$ implies $\rho(\phi|X,\theta^*) = \rho(\phi|X,\theta^{'})$ for all X. θ^* is globally identified if it is relatively identified to all $\theta \in \Theta$.*

Lemma 1.8.1. *Preferences of the governor cannot be identified with the data from the first period.*

Proof. This is obvious since the preferences of the governor does not matter for D_V. Hence, conditional on other parameters, any values for the governor's preference parameters will create the same probability distribution ■

Lemma 1.8.2. *Given data from a single period, identification of the preferences of individual members is not possible. That is to say, even if all the parameters are identified, it is not possible to link them to specific players.*

Proof. Suppose $\theta^* \in \Theta$ is globally identified. Recall that $\theta^* = \{\{\alpha_{0i}, \alpha_{1i}, \alpha_{2i}, i = \{m,c,g\}\}, \gamma\}$ where γ is the distribution parameters. Let θ' be the set of parameters where we exchange α^*_{km} with α^*_{kg}. Obviously, $\rho(\phi|X,\theta^*) = \rho(\phi|X,\theta')$ but we have exchanged the preferences of two players ■

Define the set

$$\Theta^{NS} = \{\theta \in \Theta \mid \{\alpha_{ki}, \alpha_{kj}\}_k \in \theta \implies \{\{\alpha'_{ki}, \alpha'_{kj}\}_k \notin \theta\}, \alpha_{ki} = \alpha'_{kj} i, j \in \{m, g, c\}\}$$

. That is, Θ^{NS} is the set of parameters where the preference parameters have no symmetries. If a specific set of parameters $\{\alpha_{ki}, \alpha_{kj}\}_k$ is in the set Θ^{NS}, $\{\alpha'_{ki}, \alpha'_{kj}\}_k$,, where $\alpha'_{ki} = \alpha_{kj}$ is not. Lemma 1.8.2 tells us that, using data only from one period, we can only hope to identify θ relative to the set Θ^{NS}.

For the remainder, I show that from the veto period, we can identify the set of preferences for the municipality and the governor. Next, I show that the majority period will help us identify the governor's preferences. Using these two information then, we can link the parameters estimated to specific players.

The following assumption helps concentrate on only the preference parameters.

Assumption 3. *$\varepsilon_i \sim N(0,1)$, i.i.d. and $F \sim Beta(a,b)$ - a,b known.*

Assumption 3 makes parametric assumption regarding the draws of ε and ϕ. Note that, given that these parametric distributions are known, identification reduces only to the preference parameters, which is at the hearth of this argument. Note that for what follows, these assumptions are much more restrictive than what is needed. In essence, continuous strictly increasing CDFs are sufficient to show the propositions in this section.

Next, assumption 4 assumes that X has full support on the positive real numbers.

Assumption 4. *$Pr(X \in (l,m) \subset \mathbb{R}_+) > 0$.*

Assumption 5. *At the true set of parameters, θ^*, $Pr(X|\frac{\alpha_{0i}}{\alpha_{2i}} + \frac{\alpha_{1i}}{\alpha_{2i}} X \in (0,1)) > 0$.*

Note that, with these assumption, the probability of observing a project with size ϕ is equal to the probability of drawing it and that $D = 1$. That is, for the veto period

$$\begin{aligned} Pr(\phi|X,\theta) &= Beta(\phi|a,b) \times Pr(D(\phi,X,\varepsilon) = 1|X) \\ &= Beta(\phi|a,b) \times Pr(d_m(\phi,X,\varepsilon) = 1|X) \times Pr(d_c(\phi,X,\varepsilon) = 1|X) \\ &= Beta(\phi|a,b) \times \prod_{i=m,c} [1 - \Phi(-\alpha_{0i} - \alpha_{1i}X - \alpha_{2i}\phi)] \end{aligned}$$

Proposition 1. *Given Assumptions 3, 4 and 5; set of parameters for the municipality and the center, θ^* is identified relative to the set Θ^{NS} from observations with the decision rule D_V.*

Conceptually, we are trying to place two hyperplanes in three dimensions. The proof shows that given the observed outcomes of the accepted projects, there is only one way we can do this. The method relies on functional forms and utilizes the linearity of the utility functions and the non-linearity of the probability distribution of ε.

Proof. Suppose not. That is suppose there is $\theta \in \Theta^{NS}$ such that $Pr(\phi|X,\theta) = Pr(\phi|X,\theta^*)$ a.e. I will show that the set $\{\theta|Pr(\phi|X,\theta) = \mu\}$ is singleton. Let

$$\theta_i^* = \begin{bmatrix} \alpha_{i0} \\ \alpha_{i1} \\ \alpha_{i2} \end{bmatrix}$$

and

$$Z = \begin{bmatrix} 1 & X & \phi \end{bmatrix}$$

Within the set $E \subset \mathbb{R}_+$, the probability mass of $\phi \in P \subset [0,1]$ that will be observed is equal to:

$$\begin{aligned} Pr(\phi|X \in E, \phi \in P) &= \int_{\phi \in P} \int_{X \in E} Pr(\theta_1^* Z > -\varepsilon_1) Pr(\theta_2^* Z > -\varepsilon_2)\ dG(X)\ dBeta(\phi|a,b) \\ & \int_{\phi \in P} \int_{X \in E} (1 - \Phi(\theta_1^* Z)) \times (1 - \Phi(\theta_2^* Z)) dG(X)\ dBeta(\phi|a,b) \end{aligned}$$

Since all distributions and arguments are continuous, it must be that at each (ϕ, X) pair, for any observationally equivalent θ^* and θ pair, we must have

$$Pr(\theta_1^* Z > -\varepsilon_1) Pr(\theta_2^* Z > -\varepsilon_2) = Pr(\theta_1 Z > -\varepsilon_1) Pr(\theta_2 Z > -\varepsilon_2) \quad (30)$$

$$(1 - \Phi(\theta_1^* Z)) \times (1 - \Phi(\theta_2^* Z)) = (1 - \Phi(\theta_1 Z)) \times (1 - \Phi(\theta_2 Z)) \quad (31)$$

for all $X \in \mathbb{R}_+$ and $\phi \in [0,1]$. Let $\eta \in \mathbb{R}_+$. Let $Z_\eta = \begin{bmatrix} 1 & X+\eta & \phi \end{bmatrix}$. Note that, $\theta_i Z_\eta = \theta Z + \alpha_{i1}\eta$.

Now, notice that as we change η, $\theta_i Z_\eta$ changes linearly for each player. However, $(1-\Phi(\theta_i^* Z)$ moves non-linearly. Also this movement depends on whether $\theta_i^* Z$ is greater than 0 or not. For each θ, Z, η, define, the function $f : \mathbb{R}^3 \times \mathbb{R}^2_+ \times [0,1] \Rightarrow [0,1]$, such that $(1-\Phi(\theta_i^* Z_\eta) = (1-\Phi(\theta_i Z) + f(\theta, Z, \eta)$. The important thing to note is that $f(\theta, Z, \eta) \neq f(\theta', Z', \eta)$ for $\theta Z \neq \theta' Z'$.

Since θ and θ^* are observationally equivalent, 31 implies:

$$(1-\Phi(\theta_1^* Z_\eta)) \times (1-\Phi(\theta_2^* Z_\eta)) = (1-\Phi(\theta_1 Z_\eta)) \times (1-\Phi(\theta_2 Z_\eta)) \tag{32}$$

$$(1-\Phi(\theta_1^* Z))f(\theta_2^*, Z', \eta) - (1-\Phi(\theta_2^* Z))f(\theta_1^*, Z', \eta) + f(\theta_1^*, Z', \eta)f(\theta_2^*, Z', \eta) \tag{33}$$

$$= (1-\Phi(\theta_1 Z))f(\theta_2, Z', \eta) - (1-\Phi(\theta_2 Z))f(\theta_1, Z', \eta) + f(\theta_1, Z', \eta)f(\theta_2, Z', \eta) \tag{34}$$

But some algebra easily shows that this is only possible if $\theta_i^* = \theta_j$ where $i \neq j$. the intuition behind is very simple. If $\theta_i^* \neq \theta_j$, then we can find some X such that the *directional* derivatives of these hyperplanes are different. Taing this idea a step further, for each two hyperplanes, we can find some X where the directional derivative of the original hyperplanes cannot be the same. This in return implies that $\theta \notin \Theta^{NS}$ Q.E.D.

Next lemma shows that one can get the identities of the estimated parameters using the data from the second period where the aggregation rule D_{NV} is used. However, we have to make the following, somehow restrictive, assumption.

Assumption 6. $\alpha_{2g} \in \left[min\left\{ \frac{\alpha_{1m}}{\alpha_{2m}}, \frac{\alpha_{1c}}{\alpha_{2c}} \right\}, max\left\{ \frac{\alpha_{1m}}{\alpha_{2m}}, \frac{\alpha_{1c}}{\alpha_{2c}} \right\} \right]$.

This assumption guarantees that the governor's preferences lies somewhere between that of the mayor's and center's. I will not impose this restriction in the estimation and hence will be able to check it afterward.

The difficulty of figuring out which player each set of preference parameters identified by the previous Lemma stems from the fact that we cannot observe $Pr(d_g(\phi, X, \varepsilon) = 1)$ directly. To understand this clearly, consider a set $E \subset [0,1] \times \mathbb{R}_+$. Abusing some notation, let $p_i = Pr(d_i(\phi, X, \varepsilon) = 1|E)$. Then, what we observe under the rule D_V is

$$Pr(D(\phi, X, \varepsilon) = 1|E, D_V) = p_m \times p_c$$

and

$$Pr(D(\phi, X, \varepsilon) = 1|E, D_{NV}) = p_m \times p_c + p_m \times p_g - p_m \times p_c \times p_g$$

Note that we observe $Pr(D(\phi, X, \varepsilon) = 1|E, D_V) = \mathbb{P}_{E,V}$ and $Pr(D(\phi, X, \varepsilon) = 1|E, D_{NV}) = \mathbb{P}_{E,NV}$ in the data. Also, we observe, p_1 and p_2, without the ability (for now) to specifically match them to municipality versus the center.

Now, consider $\frac{\mathbb{P}_{E,NV}}{\mathbb{P}_{E,V}} = 1 + p_g(1/p_c - 1)$. This is the change of projects in the set E. Hence, letting $p_c = p_1$ or $p_c = p_2$ will change the *imputed* p_g, but not the actual observed outcome. Therefore, I propose another method of identification, which relies on 4 and 6. This will enable me to construct an identification at infinity argument.

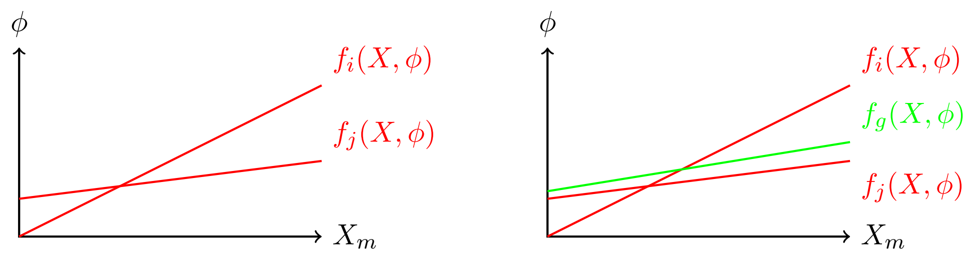
The idea of identification is the following. As we move along X dimension, α_{2i} –which is the coefficient in front of X– determines a certain rate of convergence of the observed distribution of $P(\phi|X)$ in either dataset. Hence, the identity matters with respect to the rate of convergence, which is determined by α_{2i}.

Recall that $P(\phi|X, D_{NV}) = p_m p_c + p_m p_g - p_m p_g p_c = p_m(p_c + p_g - p_c p_g)$. Take an example, suppose that both α_{2i}, $i = 1, 2$, are positive. Then, as we move along X, projects will become more and more admissible by both the center and the mayor. Hence, at the limit, we will have *every* project will be accepted. Hence, loosely speaking, $\lim_{X \to \infty} Pr(D_V(\phi, X, \varepsilon) = 1) = 1$. However, now suppose $\alpha_{21} > \alpha_{22} > 0$. If we can somehow *control* for p_g, we can say something about the rate of convergence of $D_V(\phi, X, \varepsilon) = 1$. In fact, suppose $\alpha_{21}, \alpha_{22} > 0$ and assumption 6 holds. Since $D_V(\phi, X, \varepsilon) = p_m p_c + p_m p_g - p_m p_g p_c$, the speed of convergence of *plim* $Pr(D(X, D_V) = 1)$ will give us an idea about the identities. To see this, suppose $2 = m$. Then the convergence speed will be equal to that implied by $Pr(d_1(X, D_V) = 1)$. Otherwise, it will be equal to some rate $O(x_n) < Pr(d_1(X, D_V) = 1)$. That is to say, if the speed of convergence of *plim* $Pr(D(X, D_V) = 1)$ is the same as $Pr(d_2(X, D_V) = 1)$, $2 = m$. Otherwise, it will be equal to $1 = m$. This is the basic idea behind identification.

Perhaps an easier way to understand this is graphical. First, consider the loci of the preferences of the players over the (X, ϕ) plane, denoted by $f_i(X, \phi)$ for player i. These are given in 13. Notice that a project below the locus of player i has a higher probability of being accepted than one above, given X. On the left panel, the loci of the preferences of the center and the mayor. Notice that, since we do not know, at this point, which locus belongs to which player, I denote the loci by i and j. The assumption 6 implies that the preference of the governor will lay somewhere between these two loci, as shown on the right panel.

Figure 13
PREFERENCES OVER X OF PLAYERS

(a) D^V **(b)** D^M



The idea of identification relies on the behavior of the projects around this loci when X is taken to infinity. Consider a sequence of $\eta_n = (\phi_n, X_n)$ where $min\{f_m(\bar{\alpha}; X_n, \phi_n), f_c(\bar{\alpha}; X_n, \phi_n)\} = 0$. This sequence is shown in figure 14 and is on the lowest locus identified from the observations with the veto. Now, consider the probability each player accepts these projects, η_n. Without loss of generality, assume that $f_j(X, \phi) < f_i(X, \phi)$ for large X, like shown in the figure. Given that each player accepts a project on its locus with probability $\frac{1}{2}$, we have $Pr(d_j(\eta_n)) = 0.5$ for large enough X. This also implies that $Pr(d_i(\eta_n)) = 1$ as X approaches infinity, as the projects are gradually getting further away from f_i, as shown in the right panel of figure 15.Due to assumption 6, this is also true for *any* preferences of governor, as shown in the right panel of figure 15.

Figure 14
THE SEQUENCE η_n

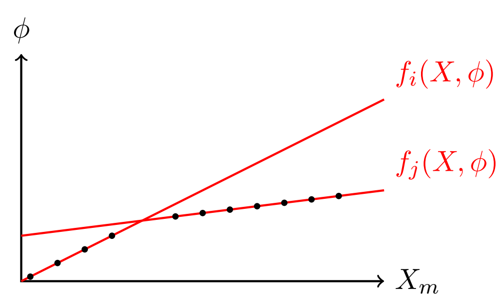
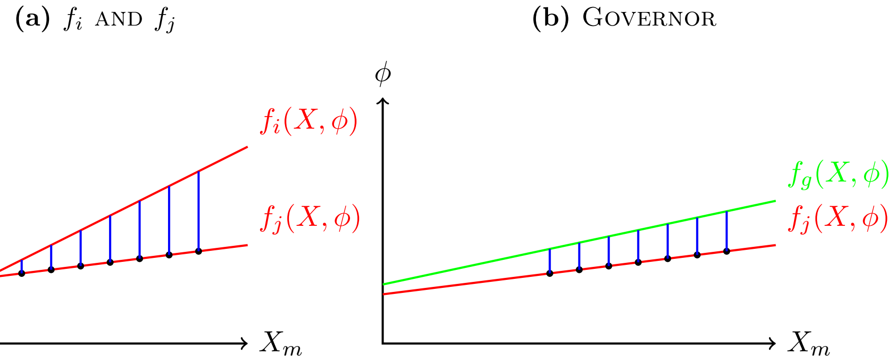


Figure 15

PROBABILITY OF ACCEPTING η_n



The probability of observing projects in the η_n, as X approaches infinity has different implications for D_V and D_M. As both the center and the mayor has to approve the projects in D^V, and one of these players accept these projects with probability 0.5 and the other with 1 when X is taken to infinity, we will observe only half the drawn projects (remember that projects distributions are known, which is guaranteed by exclusion restrictions in the actual model). Now, in D^M, what portion of these projects will be observed will depend on whether $m = i$ or $m = j$. If the former is correct, then the mayor accepts all the projects around η_n whereas if the latter is correct, than she accepts only half of the arrived projects around η_n. Since the governor accepts all the projects in η_n when X is large, this means that the probability of observing η_n in D_M will be 1 if $m = i$ and 0.5 if $m = j$. Hence, we can label the preferences with identities using this method. The following lemma makes this argument formal.

Lemma 1.8.3. *Given assumptions 3, 4, 5 , 6, one can match the identities of the municipality and center to their preferences using data from both D_V and D_{NV} periods. That is, one can identify the preference parameters of municipality and center.*

Proof. Let the loci of the preferences and the η_n defined as above. For some $\varepsilon > 0$, consider the set

$$\Omega(X) = [(\phi - \varepsilon, \phi + \varepsilon)|(\phi, X) \in \eta_n]$$

Absent the committee, we know the arrival rate of the projects in the set $\Omega(X)$ for each X. For one of the identified set of parameters from the previous lemma, we must have that the probability of accepting a project in this set equal to $\frac{1}{2}$. Denote this set of parameter as j. If in the set of observations fom D^M, we have $\mu(\Omega(X)|D^V) = \mu(\Omega(X)|D^M)$ as $X \to \infty$, we must have $j = m$.

Otherwise, if $\mu(\Omega(X)|D^V) < \mu(\Omega(X)|D^M)$ as $X \to \infty$, it must be that $i = m$ ■

The identification of the preferences of the governor, knowing the preferences of the mayor and the center is a corollary to proposition 1.

Derivation of the Full Likelihood

Denote the observed project in municipality m, in period t as $\omega^O_{m,t}$. When municipality m in period t implements a project in period t, $\omega^O_{m,t}$ records the type s and size ϕ of the project. Otherwise, $\omega^O_{m,t} = \emptyset$.

Denote the parameters we need to estimate by θ and characteristics of a municipality m, by Z_m. Recall that the probability of observing a project of size ϕ and type s is

$$\begin{aligned} Pr(\omega^O_{m,t} = \{\phi, s\}) &= \Pi^{draw}(D_{admeff}, t; \theta_{draw})\Pi^{Type}(\theta_{type}) \\ &\quad \times f_{Beta}(\phi_t|B_t; \theta_{beta})Pr(D^*(t, B, \phi, s = 1; \theta_{pref})) \end{aligned}$$

similarly, the probability of being rejected is:

$$\begin{aligned} Pr(\omega^O_{m,t} = \emptyset) &= (1 - \Pi^{draw}(D_{admeff}, t; \theta_{draw})) \\ &\quad + \sum_{s=pub,tra} \Pi^{draw}(D_{admeff}, t; \theta_{draw})) \\ &\quad \times \Pi^{Type}(\theta_{type}) \left(\int_{\tilde{\omega}} g_\omega(\tilde{\omega}|B)(1 - Pr(D^*(t, B, \tilde{\omega}) = 1; \theta_{pref}))d\tilde{\omega} \right) \end{aligned}$$

where all the functions are defined in the main text. This implies that observing a specific project $\omega^O_{m,t}$ is equal to:

$$Pr(\omega^O_{m,t}) = Pr(\omega^O_{m,t} = \emptyset)^{\mathbb{1}(\omega^O_{m,t} = \emptyset)} Pr(\omega^O_{m,t}|\omega^O_{m,t} \neq \emptyset)^{(1 - \mathbb{1}(\omega^O_{m,t} = \emptyset))} \tag{35}$$

Taking the logarithm of this equation and summing across time and municipalities delivers the result.

Supplementary Tables

Table 20
Number of projects across sectors

	Mean	St. Dev.	Median	Min	Max
Total Projects by Municipality	5.36	6.63	3	0	86
By sector:					
Clean water and basic sanitation	.38	.78	0	0	6
Agriculture	.2	.59	0	0	7
Environment and sustainable development	.06	.27	0	0	4
Business, Industry and Tourism	.03	.18	0	0	2
Communication	.01	.11	0	0	2
Culture	.08	.33	0	0	5
Children and Adolescents	.02	.13	0	0	2
Sports and recreation	.55	.83	0	0	5
Defence	0.03	0.06	0	0	1
Social Inclusion and reconciliation	.78	.34	0	0	5
Education	.40	.86	0	0	6
Urban equipment	.25	.53	0	0	5
Mines and Energy	.10	.36	0	0	3
Health and social protection	.12	.46	0	0	4
Transportation	1.33	1.22	1	0	7
Dwellings	.45	.74	0	0	5
Institutional strengthening	0.7	0.28	0	0	2
Labor	.002	.06	0	0	1
Law and justice	.01	.11	0	0	1
Risk management	.04	.21	0	0	2

Table 21

COST OF PROJECTS ACROSS SECTORS (IN 10000 USD (2014))

	No. Obs	**Mean**	**St. Dev.**	**Median**	**Min**	**Max**
All Projects	4290	46.75	107.65	20	0.07	4004
By sector:						
Clean water and basic sanitation	385	70.15	228.74	24.33	2	4004
Agriculture	210	27.94	64.9	14.45	.25	645.08
Environment and sustainable development	64	50.95	94.53	14.82	0.75	554.1
Business, Industry and Tourism	32	45.85	71.83	23.14	.7	350
Communication	12	51.48	94.93	18.49	2.49	344.38
Culture	82	28.12	44.82	12.25	.87	240.29
Children and Adolescents	16	36.83	39.55	29.95	1.94	146.09
Sports and recreation	547	35.49	56.21	20.86	1.3	317.3
Defence	4	112.45	139.65	62.4	13.3	311.75
Social Inclusion and reconciliation	78	47.98	123.5	15.4	1.25	705.6
Education	441	58.20	115.14	20.81	0.62	1027.96
Urban equipment	258	44.28	78.37	19.62	.85	535
Mines and Energy	106	52.91	126.6	19.24	.99	430.16
Health	124	57.52	115.87	20	0.007	889.8
Transportation	1348	50.65	84.29	24.32	1	1037
Dwellings	459	35.21	93.87	14.45	.16	1527.95
Institutional strengthening	81	14.91	13.46	12.38	.75	69.2
Labor	3	99.15	155.78	13.21	5.46	278.75
Law and justice	12	15.59	13.45	10.44	4.51	52.51
Risk management	46	28.2	35.14	14.77	.87	151.95

Table 22
REMAINING BUDGET BY PERIOD (IN 10000 2014 US)D

Period	Mean	Std. Dev.	Min.	Max.
Budget	164.834	423.515	1.298	7884.818
1	162.425	418.034	0	7884.818
2	155.425	401.471	0	7809.334
3	138.498	377.238	0	7809.334
4	121.625	322.184	0	5941.503
5	100.817	292.783	0	5941.503
6	76.658	240.864	0	4769.400
7	73.989	237.156	0	4769.400
8	62.727	193.502	0	3243.377
9	49.76	136.619	0	1720.112
10	37.816	109.509	0	1321.116
11	25.402	91.178	0	1054.7
12	7.874	45.267	0	660.551
N		1018		

Table 23

Difference statistic within some sub-samples

Sample	Statistic	Veto Period	No Veto Period	Difference
All	Mean	.75	.68	.06***
	St. Dev	(.38)	(.41)	(.02)
	N	542	454	
President and mayor same party	Mean	.77	.69	.07
	St. Dev	(.34)	(.44)	(.05)
	N	135	122	
President and mayor different parties	Mean	.75	.68	.07**
	St. Dev	(.39)	(.41)	(.02)
	N	135	122	
Governor and mayor same party	Mean	.80	.61	.19**
	St. Dev	(.34)	(.44)	(.08)
	N	52	44	
Governor and mayor different parties	Mean	.74	.69	.05**
	St. Dev	(.34)	(.44)	(.03)
	N	52	44	
No debt	Mean	.72	.67	.06
	St. Dev	(.40)	(.42)	(.04)
	N	253	223	
High debt	Mean	.77	.69	.08*
	St. Dev	(.37)	(.41)	(.03)
	N	157	123	
Far from a major city in the state	Mean	.75	.67	.08**
	St. Dev	(.38)	(.41)	(.02)
	N	265	201	
Close to a major city in the state	Mean	.75	.69	.05
	St. Dev	(.38)	(.42)	(.04)
	N	277	253	

Note: The difference statistic computed using unpaired samples, assuming the samples were independent with unequal variances.

Chapter 2 : Campaign Spending and Strategy in U.S. Congressional Elections*

2.1 Introduction

Candidates running for political office spend a vast and ever-growing sum of money. In the 2012 cycle, candidates running for seats in the U.S. Congress spent about $1.9 billion, representing an increase of almost 50% in real terms relative to 2000. Despite this significant sum of funds channeled to political campaigns, there seems to be no consensus among social scientists as to the impact of this money on political outcomes.[41] For example, Feldman and Jondrow [1984], Ragsdale and Cook [1987], and Levitt [1994] find no statistically significant effect of incumbent spending on outcomes - and perhaps even a negative effect - whereas Abramowitz [1988], Grier [1989], Moon [2006], and da Silveira and de Mello [2011] find a positive and statistically significant effect. Furthermore, most of this literature is unable to capture the heterogeneity of campaign spending effects across candidates. In this chapter, I propose a new empirical framework that explicitly models the heterogeneity in the use and effect of campaign funds. To this end, I use a model of campaigning that allows funds to be spent on different campaign strategies, which may affect election outcomes differentially. I argue that an understanding of the impact of campaign funds on elections is possible only when the heterogeneous effects of campaigning strategies is uncovered.

*This chapter is co-authored with Devin J. Reilly.

[41]See, for example, Stratmann [2005]

This approach enables us to investigate campaigning strategies employed by the candidates running for political office. In particular, I model and analyze the campaign tone (positivity or negativity) a candidate uses. In fact, understanding campaign strategies is of interest in and of itself. Evidence suggests that campaigns have become increasingly negative in tone since 2000. For instance, Fowler and Ridout [2013] point out that in the 2000 presidential election, approximately 60% of ads were negative.[42] In 2012, approximately 85% of the total ads were negative. This rapid increase in negativity has sparked wide and often critical commentary of such advertisements.[43] The particular channel I investigate is built on the anecdotal and empirical evidence that suggests negative advertising may discourage voter turnout. For instance, it is widely believed that heavy negative campaigning between the two major party candidates in the 2000 Minnesota Gubernatorial elections depressed their turnout, which opened the door for the third-party candidate Jesse Ventura to win the election.[44] There is also some concern that negative campaigning may contribute to polarization or voter fatigue.[45] This feature of campaign strategy is often overlooked in the empirical literature.[46] The political science literature has often found that not only does negative advertising differ from positive advertising in its overall effects on voters, but also the effects vary *across* different groups of potential voters.[47] For instance, negative advertising may have a demobilizing effect on ideological voters, while positive advertising may be more effective in attracting swing voters. To the extent that optimal campaign strategies differ systematically across different types of candidates and elections, it is important to understand the differential impact negativity can have on voting outcomes.

This chapter builds on the argument that campaign finance and strategies are heavily interrelated, and therefore should be analyzed together. To understand the true impact of a dollar on election outcomes, one must understand how that dollar will be spent in a campaign. This is because

[42] Negativity is measured as attack or contrasting advertising, which is typical in the political science literature.

[43] See, for instance, *Washington Post*, Feb. 20, 2012, "Study: Negative Campaign Ads Much More Frequent, Vicious Than In Primaries Past."

[44] See, for example, Lentz [2001].

[45] See, for instance, Ansolabehere et al. [1994], which in an experimental setting found evidence that negative advertisements "weakened confidence in the responsiveness of electoral institutions and public officials. As campaigns become more negative and cynical, so does the electorate."

[46] Often, any notion of strategy besides overall spending is overlooked. Three important exceptions are Stromberg [2008], which looks at presidential campaign stops, Nalebuff and Shachar [1999], which investigates the exertion of candidates' effort to increase participation, and Gordon and Hartmann [2013a], which analyzes the optimal allocation of advertising across states under the Electoral College. Even among these cases, none includes a choice of overall negativity of the campaign.

[47] See, for example, Ansolabehere and Iyengar [1997], and several others discussed in section 2.4.1.

elections differ in terms of their fundamentals, which ultimately determine how effective the campaign strategies employed will be. Thus, the effect of campaign funds depends on the effectiveness of the strategy that will be used in equilibrium, which will vary across candidates and elections. Unless one understands *how* funds will be allocated among different strategies, one cannot be able to uncover their true impact on the outcomes of elections. In addition, the campaign strategies chosen depend on the available funds to the candidate and to his opponent. For example, a candidate might be more likely to engage in negative campaigning when both he and his opponent have large budgets, but may tend to be more positive when he has a large money advantage. Hence, approaching these two questions in isolation could result in misleading answers. The theoretical and empirical strategy tries to avoid this by focusing on this very important interplay of campaign funds and strategies.

To this end, I develop a model featuring a game between candidates that decide their campaign strategies. In particular, candidates decide on how to allocate their total budgets between positive and negative campaigning. I denote a candidate's campaigning that includes information only about himself as a positive one. On the other hand, when the campaigning includes information about the opponent, it is a negative one.[48] Each constituency has three types of voters: the base (ideological) voters for each candidate and swing voters. I assume that negative campaigning is a demobilizing tool: it demobilizes the supporters of the opponent at the expense of possibly alienating some of the candidate's own. On the other hand, positive campaigning is used to attract swing voters to vote in favor of the candidate. More precisely, I assume that ideological voters decide only on their turnout. When they do, they vote for the specific candidate they support. Swing voters, on the other hand, always turnout to vote, but decide on whom to vote based on the (positive) campaigns of the candidates. Hence, in the model, a candidate campaigns negatively to reduce the turnout of the opponent's supporters at the cost of decreasing his own. Positive campaigning, on the other hand, increase the portion of swing voters who vote for her.

Elections differ from each other in the measure of voter groups. The measure of ideological voters for each candidate, and hence swing voters, vary across elections, which result in different equilibrium campaigning strategies for each election. In the empirical specification, these levels of support are drawn from a distribution depending on the election-specific observables. These draws are observed by the candidates while they are unobservable to the econometrician. Given initial support, candidates choose their allocations simultaneously, after which the election takes place and the

[48] This categorization is the norm in the literature when measuring negative advertising.

winners are realized.

To infer the overall campaign strategy of the candidates, I use data from from the Wisconsin Political Advertisement project that records each television advertisement aired by a candidate. This dataset also records the tone of the advertisement. Hence, I assume that the TV campaigning strategy is representative of the overall strategy of the player. Moreover, to estimate the distribution of voters, I use data from the American National Election Study. I calibrate the model to match patterns of campaign tone observed in the data, and then use the calibrated model to understand the effects of spending and strategies on voting outcomes.

To see why a model is necessary to understand the impact of campaign funds and strategies, consider the following. Negativity may be a useful strategy for candidates who are trailing, as it may lead the base of supporters of a front-runner to shrink. Conversely, positive campaigning may be relatively more effective for a front-runner. This is in line with the observation in the data that incumbents tend to be much more positive than challengers. However, if one uses ex-post vote measures to try and infer the effectiveness of advertising, one would tend to see negativity correlated with low vote shares and thereby conclude negative advertising is ineffective. Note that this goes beyond purely incumbent-challenger races, as even in open seat elections, ex-post vote margin is negatively correlated with campaigning negativity. Controlling for the endogenous decision of campaign negativity with respect to things like initial voter support is important in understanding the overall effectiveness of different campaign strategies.

The calibrated model suggests that campaign spending is relatively ineffective at increasing vote shares. For the average election, which has budgets of about $2.4 million, a 10% increase in one candidate's budget increases his expected vote differential by about 0.4 percentage points. This is roughly in line with results from Levitt [1994], among others. To understand the differential impact of campaign spending, I perform multiple exercises. For example, consider an election where the candidates have the same measure of supporters. When these candidates have no funds to campaign, they are expected to tie in the election. Now consider providing one candidate with a $2.1 million budget while the other $700,000 (which are approximately the 75th and 25th percentile of observed budgets, respectively). This yields a 2.5 percentage point improvement in the expected vote of the first candidate. While not insignificant in absolute terms, a $1.4 million advantage is approximately the same size as the median budget. Thus, 2.5 percentage points is arguably a relatively minor increases for such a sizable budget advantage. I employ other calculations to find that, albeit small, trailing candidates benefit from extra funds more than the leading ones. The model also implies that negative campaigning is relatively effective for candidates who face an opponent with a high level of initial support, while positive campaigning is relatively effective for candidates in elections where neither side has a particularly high initial support. Still, the differences in effectiveness are not large, especially given that overall effectiveness is low. Finally, the model implies slightly decreasing returns to spending. Both this feature and the relative effectiveness of negative campaigning for trailing candidates may explain why the previous literature tends to find challenger spending is relatively more effective than incumbent spending.

The rest of the chapter is organized as follows. Section 2.2 discusses the related literature while section 2.3 focuses on the underlying institutional framework. Section 2.4 describes the model and the proof of the existence of a unique equilibrium. Section 2.5 discusses the data, and section 2.6 describes the calibration of the model. Finally, section 2.7 provides the description and results from model simulations, and section 2.8 concludes.

2.2 Related Literature

The key feature of the model is the candidate's decision of how to allocate his budget between positive and negative campaigning. Several theoretical papers focus on this decision. One of the earliest examples is Harrington and Hess [1996], which studies negative campaigning in a spatial framework. In their model, the negativity of a campaign depends primarily on the personal attributes of the candidate. Later works focus on the signaling game associated with advertising when candidate qualifications are unknown to voters (see Bhattacharya [2012] and Hao and Li [2013], for instance).

For tractability, I abstract from the spatial framework and the signaling aspect of political advertising and focus on the direct effect by using an "influence function" Bhattacharya [2012] that affects voter support for each candidate. The theoretical framework I utilize for the campaign stage of the model is similar to Skaperdas and Grofman [1995]. They model a two-candidate competition where each candidate decides on positive and negative campaigning levels given fixed and equal budgets. Negative campaigning by a candidate depresses turnout (for both his own and his opponent's supporters) while positive campaigning influences undecided voters. Through this setup, they argue that they can broadly match some regularities of political competition - namely, that the front-runner chooses more positive advertising than his opponent and that negative campaigning is greater the stronger his opponent's support. I differ from their analysis in allowing for the possibility of asymmetric budgets and decreasing returns to negative campaigning. More importantly, there is no empirical component to their analysis, whereas I calibrate the model to match campaign tone data, allowing for rich heterogeneity in budget, advertising, and district-specific data.

In addition to theoretical work, there have been several structural models studying the determinants and effects of political advertising and campaigning. Gordon and Hartmann [2013a] focuses on the allocation of television advertising across markets in presidential elections. They use a BLP-type setup to understand how the Electoral College system distorts advertising decisions relative to a popular vote system. Shachar [2009] attempts to explain the finding from Nalebuff and Shachar [1999] that participation rates in U.S. presidential elections tend to be higher in states with narrower expected margins of victory. The author models campaign marketing activities in a two-candidate contest and estimates the model, finding that candidates advertise more in close states, which can drive higher turnout. Stromberg [2008] also estimates a model of the allocation of resources in U.S. presidential candidates under the Electoral College, with a focus on campaign visits rather than advertising.

Previous structural work surrounding campaign strategies differs from ours in several respects. To my knowledge, none differentiates between positive and negative campaigning, which previous empirical work has shown affect turnout and election outcomes in distinct ways.[49] Furthermore, most analyses use U.S. presidential election data and rules, in particular the Electoral College, whereas I use Senate and House elections with plurality voting systems. The focus in the campaign stage of the model is to understand the overall and relative effectiveness of positive and negative campaigning in winning elections, not on the allocation induced by electoral rules.

More broadly, the model sheds light on the overall impact of spending on elections. A few previous attempts have been made to estimate the overall impact of campaign spending on election outcomes. Palda and Palda [1998] uses regression analysis of French election data and finds a very small effect of campaign spending on vote shares. Levitt [1994] uses races with the same two candidates to estimate the effect of spending on outcomes and finds little to no effect, as well. However, Stratmann [2009] utilizes the same methodology but analyzes the effect of television advertising on vote shares, finding a significant impact. Additionally, da Silveira and de Mello [2011] uses a quasi-natural experiment in

[49] See, for instance, Ansolabehere and Iyengar [1997]

Brazil due to the two-round voting system and a rule that allocates TV advertising exogenously and differently in the first and second rounds. The authors find a large causal effect of TV advertising on election outcomes. Finally, in one of the few examples of structural analysis of campaign finance, Kawai and Sunada [2015] models fund-raising and spending in House elections. The authors find a relatively small effect of spending on election outcomes, slightly larger than that of Levitt [1994]. For other examples, Stratmann [2005] provides a thorough review of the literature. While these works shed light on the impact of campaign contributions and spending, I focus on the effectiveness of different campaign strategies, namely negative versus positive campaigning, which can help to explain the relative difference in effectiveness for challengers and incumbents, for example.

2.3 Legal Background

In the analysis, an important component of the model is the receipts of political campaigns, taken as exogenous. Therefore, it is important to note some of the legal background surrounding campaign finance laws. While there were some changes to such laws within the sample, the regulations regarding political action committees (PAC)[50] and individual contributions did remain constant over this period. The limits on PAC contributions to candidates is the same throughout the sample. Most changes due to the Bipartisan Campaign Reform Act in 2002 involved the use of "soft money" (i.e. nonfederal funds subject to less regulation prior to the reform) and independent expenditures. Both soft money and independent expenditures deal with spending by parties and outside groups, not by the candidates. In the model, I focus only on spending by candidates, so these changes are generally not directly relevant for the type of campaign spending I consider. While outside spending may affect voters and therefore act as well, it is also important to note that within the sample, outside spending is a relatively minor part of aggregate federal campaign spending. For instance, in 2008, total spending in all federal races was approximately $5.3 billion, while outside spending comprised only about 6% of this.[51] In 2004, outside spending was only about 4.7% of total federal election spending, and in 2000 the number was 1.8%. Thus, outside spending did not play a quantitatively

[50] Political action committees are groups "organized for the purpose of raising and spending money to elect and defeat candidates." They are an important component of campaign contributions, and face a legal limit on how much they can contribute in each cycle to each candidate.

[51] This data is from *Opensecrets.org*, which extracts data from FEC filings. It does not accurately break down total outside spending by race type due to limitations in how the reports are filed. This is why I only list total federal election spending, which includes the presidential race in each year.

significant role .

In a similar vein, I also do not consider the emergence of super PACs. This is because the *Citizens United v. Federal Election Commission* Supreme Court decision, which allowed for the existence of such organizations, occurred in 2010. Super PACs can engage in an unlimited amount of spending so long as their expenditures are made independently from campaigns. Since my sample ends in 2008, super PACs do not affect the analysis.

In addition to laws regarding campaign contributions, it is also important to understand the protocol for campaign advertising. Television advertisements in the model are how I measure campaign strategies. In the United States, the main regulation on television advertisements is that stations and cable networks must "treat legally qualified candidates equally in allocating airtime." That is, if a station provides airtime to one candidate, it must offer "the same amount of airtime with the same audience size to all other candidates at the same rate," though if the other candidates cannot afford this airtime the network is under no obligation to provide it at a lower price Karanicolas [2012]. Thus, conditional on both candidates having sufficient funds, this effectively guarantees symmetry in access for each candidate in a given election. There is also a "reasonable access" rule that ensures availability of advertising time to all candidates at the rates paid by their most favored advertisers. This also implies that rationing of advertising spots will not occur in most cases. Overall, the legal rules regarding television advertising govern the setup of the model for political campaigns.

2.4 Model

I consider a model of campaigning in which there are two candidates competing for votes in an election. There are E elections held in the sample, with one Democratic and one Republican candidate in each election. Each candidate is endowed with a budget consisting of individual and PAC contributions, which is common knowledge among the candidates. Given these budgets, candidates allocate their funds between negative and positive advertising to maximize their expected vote shares. Given the chosen strategies, expected vote shares for each candidate $i \in \{1, 2\}$ are realized as a function of campaign strategies. Finally, an election-day shock that is orthogonal to information at the time campaigning decisions are made is realized, and a winner is determined.

More specifically, the model of the campaign game between candidates is similar to Skaperdas and

Grofman (1995), although I deviate from it considerably. There are two candidates $i \in \{1,2\}$ in every election. For each election, there is a population of unit mass of potential voters. Within each of these populations, a share $r_i \in [0,1]$ initially supports candidate $i \in \{1,2\}$, which I sometimes denote as ideological voters. These shares are restricted to be such that $r_1 + r_2 \leq 1$. The remaining share $R = 1 - r_1 - r_2$ are considered swing voters. These shares are known to candidates prior to their campaigning decisions.

Candidate i is endowed with budget B_i, comprised of all political contributions. Candidates simultaneously select the share of their budget to spend on positive or negative campaigning. I assume that different campaign strategies have differential effects on each group of voters. Negativity by candidate i primarily demobilizes his opponents' initial supporters, though they may also have the cost of demobilizing his own support, as well. These demobilized supporters do not turnout to vote. Positivity attracts swing voters. More precisely, a larger level of positivity by candidate i will attract a larger share of the R mass of swing voters, all else equal. In this sense, positivity assumed to be persuasive to swing voters.

Formally, let x_i denote candidate i's *level* of negative spending and y_i denote his positive spending level. Given (y_1, y_2), candidate i will receive share $q^i(y_1, y_2) \equiv \frac{(1+y_i)^{1/\gamma}}{(1+y_1)^{1/\gamma}+(1+y_2)^{1/\gamma}}$ of the total mass R of swing voters, where $\gamma \geq 1$ is a parameter. These functions have the property that $q^i(y_1, y_2)$ is increasing and concave in y_i, and decreasing and convex in y_j for $j \neq i$. Furthermore, they assume that if neither side chooses any positive campaigning, swing voters will split evenly among the two candidates. Note that the smaller γ is, the more effective positive campaigning is in gaining swing voter support.

For negative campaigning levels (x_1, x_2), the total shares of support retained by candidates 1 and 2 are given by:

$$exp\{-\alpha_1 x_1 - \alpha_2 x_2\} \text{ for candidate 1}$$

$$exp\{-\alpha_1 x_2 - \alpha_2 x_1\} \text{ for candidate 2}$$

respectively, with parameters $\alpha_1 \geq 0$ and $\alpha_2 > \alpha_1$. That is, α_1 captures the idea that negative campaigning by a candidate will demobilize part of his own base, whereas α_2 reflects that negativity will demobilize part of his opponent's base.

Therefore, given campaigning choices (x_1, x_2, y_1, y_2), the expected share of support for candidate

$i \in \{1, 2\}$ (with $j \neq i$) is:

$$V^i(x_i, y_i; x_j, y_j) = r_i exp\{-\alpha_1 x_i - \alpha_2 x_j\} + R\left(\frac{(1+y_i)^{1/\gamma}}{(1+y_i)^{1/\gamma} + (1+y_j)^{1/\gamma}}\right).$$

The final component determining vote shares is a mean zero random shock ε with CDF $H(\varepsilon)$. I assume ε has full support. This is an exogenous popularity shock that is unknown to the candidates when they make their campaigning decisions. It encompasses all uncertainty that is realized on election day, and is orthogonal to information the candidates have at the time they make their campaigning decisions. A given $\varepsilon > 0$ corresponds to a net gain in support for candidate 2, while $\varepsilon < 0$ corresponds to a net gain in support for candidate 1. Therefore, candidate 1 wins if:

$$V^1(x_1, y_1; x_2, y_2) - V^2(x_2, y_2; x_1, y_1) \geq \varepsilon$$

which happens with probability:

$$H(V^1(x_1, y_1; x_2, y_2) - V^2(x_2, y_2; x_1, y_1))$$

The that probability that candidate 2 wins is thus:

$$1 - H(V^1(x_1, y_1; x_2, y_2) - V^2(x_2, y_2; x_1, y_1))$$

Given that $H(\cdot)$ is strictly increasing, the objective function is the expected vote share, $V^1(x_1, y_1; x_2, y_2) - V^2(x_2, y_2; x_1, y_1)$. Taking as given budgets, his opponent's spending decisions, and initial support, the problem of candidate 1 is:

$$\begin{aligned} \max_{x_1, y_1} \; & r_1 \exp\{-\alpha_1 x_1 - \alpha_2 x_2\} + R\left(\frac{(1+y_1)^{1/\gamma}}{(1+y_1)^{1/\gamma} + (1+y_2)^{1/\gamma}}\right) \\ & - r_2 \exp\{-\alpha_1 x_2 - \alpha_2 x_1\} - R\left(\frac{(1+y_2)^{1/\gamma}}{(1+y_1)^{1/\gamma} + (1+y_2)^{1/\gamma}}\right) \\ & \text{s.t. } x_1 + y_1 \leq B_1 \text{ and } x_1, y_1 \geq 0 \end{aligned} \tag{36}$$

Since the objective function is strictly increasing in both x_1 and y_1, the budget constraint binds with equality at the optimum.[52] Hence, I can rewrite the objective of the first candidate as:

$$\begin{aligned} \max_{x_1} \quad & r_1 \exp\{-\alpha_1 x_1 - \alpha_2 x_2\} - r_2 \exp\{-\alpha_1 x_2 - \alpha_2 x_1\} + \\ & R\left(\frac{(1+B_1-x_1)^{1/\gamma} - (1+B_2-x_2)^{1/\gamma}}{(1+B_1-x_1)^{1/\gamma} + (1+B_2-x_2)^{1/\gamma}}\right) \\ & \text{s.t. } x_1 \in [0, B_1] \end{aligned} \tag{37}$$

The problem of candidate 2 is analogous.

A strategy for candidate i is a function x_i, which maps budgets and initial support levels to a negative campaigning proportion.[53] Formally, a strategy of candidate i therefore $x_i : \mathbb{R}^2_+ \times \triangle^2 \to [0,1]$. That is, given budgets B_1 and B_2, and initial support levels r_1, r_2, and $R = 1-r_1-r_2$, $x_i(B_1, B_2, r_1, r_2, R)$ giving negativity *as a proportion of total budget* for candidate i.

The definition of equilibrium of the campaign game for a given election is as follows:

Definition 1. *Given initial support* (r_1, r_2, R)*, and budgets* (B_1, B_2)*, an equilibrium of this game is a pair of functions* $(\hat{x}_1(B_1, B_2, r_1, r_2, R), \hat{x}_2(B_1, B_2, r_1, r_2, R))$ *that give negative campaigning proportions for each candidate, such that for each level of initial support and budgets,* $\hat{x}_1(B_1, B_2, r_1, r_2, R)$ *solves candidate 1's problem given* $\hat{x}_2(B_1, B_2, r_1, r_2, R)$*, and vice-versa.*

2.4.1 Discussion of the Theoretical Setting

In order to make the model tractable for the empirical application, I have imposed some assumptions on the effect each type of campaigning has on voters. In this section I discuss some empirical work that supports these assumptions, as well as the potential shortcomings of the approach.

As described previously, positivity and negativity in the model differ in their effect on different types of voters. Negative campaigning suppresses turnout among ideological types, while positive campaigning affects which candidate a swing voter prefers, but not her decision to turn out. One piece of anecdotal evidence on how negative campaigning can suppress turnout is from the 1998 gu-

[52] I abstract from the savings/borrowing decision of the candidates across election cycles. In the sample, the median savings as a percentage of total receipts for Democrats, conditional on saving, is 1.3%, while for Republicans it is 1.4%. The median borrowing as a percentage of total receipts for Democrats, conditional on borrowing, is 2.1%, while for Republicans it is 4.4%.

[53] Note that here I denote the strategy as a *proportion* of the total budget. This will ease notation in the calibration section, but obviously is simply a normalization since $x_i B_i$ gives the level of negative spending. Also, trivially if $B_i = 0$, then I denote x_i as 0.

bernatorial elections in Minnesota, as described in the introduction. There is also empirical research that suggests negative campaigning can reduce turnout. One of the earliest studies documenting the effect of negative advertising on turnout is Ansolabehere et al. [1994]. In an experimental setting, the authors find a strong demobilization effect from negative advertising - exposure to negative ads decreased intentions to vote by 5%. They further support these findings using aggregate level data in a follow-up paper, Ansolabehere et al. [1999]. Recent work by Krupnikov [2011] also supports the potential demobilizing effect of negative advertisements on supporters of the advertisement's target. Note that while there are other studies that argue that negative advertising may not demobilize as much as Ansolabehere et al. [1994] claim (see, for instance, Finkel and Geer [1998], Ashworth and Clinton [2006], and Lau et al. [2007]), the model does allow for demobilizing effects of negative campaigning to be arbitrarily small.

Furthermore, negative campaigning in the model not only suppresses turnout for the target candidate, but may also demobilize supporters the sponsoring candidate. This has been referred to as the "boomerang" effect in the literature. Garramone [1985] provides some of the earliest evidence for this effect, as well as the general demobilizing effect of negative campaigning. Fridkin and Kenney [2004] and Fridkin and Kenney [2011] also find that negative advertisements can depress evaluations of the target and the sponsor. However, Krupnikov [2011] argues that negative advertising may not have a demobilizing effect on supporters of the advertisement's sponsor. In calibration, I do not restrict the "boomerang" effect to be strictly positive.

2.4.2 Theoretical Results

In this section I present results from the theoretical model. Proofs of the relevant propositions and lemmas are in Appendix 2.9.1. The main result is that an equilibrium of the campaign game exists. To ease exposition, I define the following functions for $i \in \{1,2\}$ and $i \neq j \in \{1,2\}$:

$$MB_n^i(x_1,x_2) = r_j\alpha_2 exp\{-\alpha_1 x_j - \alpha_2 x_i\} - r_i\alpha_1 exp\{-\alpha_1 x_i - \alpha_2 x_j\} \tag{38}$$

$$MB_p^i(x_1,x_2) = \frac{2R}{\gamma}\left[\frac{(1+B_i-x_i)^{1/\gamma-1}(1+B_j-x_j)^{1/\gamma}}{((1+B_i-x_i)^{1/\gamma}+(1+B_j-x_j)^{1/\gamma})^2}\right]. \tag{39}$$

$MB_n^i(x_1,x_2)$ is the marginal benefit of negative campaigning for candidate i computed at campaigning levels (x_1,x_2). $MB_p^i(x_1,x_2)$ is similarly defined to be the marginal benefit from positive campaigning for i. These equations are trivially obtained by taking the first-order conditions of the objective functions.

The following statements are useful in proving the existence of the equilibrium.

Proposition 2. *In the campaign stage of the model, the following statements are true:*

1. *If* $MB_n^i(x_1, x_2) > 0$*, then* $\left.\dfrac{\partial MB_n^i(x_j)}{\partial x_i}\right|_{(x_1,x_2)} < 0$.
2. $\dfrac{\partial MB_p^i}{\partial x_i} > 0$.
3. $\left.\dfrac{\partial MB_n^i(x_1,x_2)}{\partial x_i}\right|_{(x_1,x_2)} < 0$ *if and only if* $\left.\dfrac{\partial MB_n^j(x_1,x_2)}{\partial x_i}\right|_{(x_1,x_2)} > 0$ *and vice versa.*
4. $\left.\dfrac{\partial MB_p^i(x_1,x_2)}{\partial x_j}\right|_{(x_1,x_2)} < 0$ *if and only if* $\left.\dfrac{\partial MB_p^j(x_1,x_2)}{\partial x_i}\right|_{(x_1,x_2)} > 0$ *and vice versa.*

Item one of Proposition 2 state that the marginal benefit of campaigning is decreasing whenever it is positive. The first two items of Proposition 2 state that the marginal returns to both negative and positive campaigning are decreasing in the relevant range. This, as I show in the next lemma, will imply that the best response of the candidates will be singletons. Note that it is also trivial that if $MB_n^i(0, x_2) < 0$, then $MB_n^i(x_1, x_2) \leq 0$ for all $x_1 \in [0, B_1]$. If there was any x that violated this, it would be a contradiction to the fact that MB_n^i is continuous since it would have to discretely jump from 0 to a positive value by the first item in Proposition 2.

The last two items state that when negative (or positive) campaigning is locally a strategic substitute for one candidate, it is complementary for the other. This only holds locally, and suffices to provide us with uniqueness. In the next lemma, I show that the best responses are functions.

Lemma 2.4.1. *The best response correspondence for each player,* $BR_i(x_j)$*, is singleton, i.e.* $BR_i : [0, B_j] \to [0, B_i]$ *is a function. Moreover,* $BR_i(x_j)$ *is continuous in* x_j*.*

With these two previous results, I now prove existence of an equilibrium in the campaign stage of the model.

Theorem 1. *An equilibrium of the campaign game exists.*

For the remainder of the chapter, in each election I denote the Democrat as candidate 1 and the Republican as candidate 2.

2.4.3 Empirical model

To add flexibility to the model in matching the data, I introduce heterogeneity that is unobservable to the econometrician, but observable by both candidates prior to deciding their allocations. Given parameters and budgets, the levels of initial support (r_1, r_2, R) pin down the optimal campaigning decisions. While I observe some information regarding the levels of initial support in each district or state (e.g. demographics, surveys on party support, etc.), I do not have complete information on these variables. Thus, I assume that while I can observe the *mean* levels of initial support for each election conditional on parameters, only the candidates observe the specific realization of support.[54]

More specifically, let Z denote the demographic characteristics of the district or state in which a given election is held. I assume that the initial levels of support are drawn from a Dirichlet distribution,[55] but the exact draws are observed only by the candidates. That is, the random draw of initial supports $(\tilde{r}_1, \tilde{r}_2, \tilde{R}) \sim Dir(k\overline{r}_1, k\overline{r}_2, k\overline{R})$, where $Dir(\cdot)$ is the three-parameter Dirichlet distribution, $\overline{r}_i = r_i(Z) + \psi_s S_i + \psi_{dem} D_i + \psi_{inc} Inc_i$ for $i \in \{1, 2\}$, and $\overline{R} = 1 - \overline{r}_1 - \overline{r}_2$. The functions $r_1(Z)$ and $r_2(Z)$ are known functions mapping demographics to initial support. The construction of this functions is discussed in section 2.5. k is a parameter that does not affect the mean, but is inversely related to the variance of the distribution.[56] Parameters ψ_s, ψ_{dem}, and ψ_{inc} shift the mean of the distribution. D_i is an indicator taking a value of 1 if candidate i in election e is a Democrat and 0 otherwise,[57] and Inc_i is an indicator taking a value of 1 if candidate i is an incumbent, 0 if neither candidate is an incumbent (i.e. it is an open seat election), and -1 if his opponent is an incumbent.[58] Finally, S_i is an indicator function that takes a value of 1 if candidate i is skilled and 0 otherwise. I assume that candidates observe both his own and his opponent's skill realizations,

[54] There are very few elections in the original sample that feature a prominent third-party candidate, so I do not include these in the final sample.

[55] The Dirichlet distribution has support in the 2-dimensional simplex, making it ideal for initial support draws, since (r_1, r_2, R) necessarily must be in the 2-dimensional simplex.

[56] If $(x_1, x_2, x_3) \sim Dir(k\alpha_1, k\alpha_2, k\alpha_3)$ with $\sum_{i=1}^{3} \alpha_i = 1$, then $Var(x_i) = \frac{\alpha_i(\sum_{j \neq i} \alpha_j)}{k+1}$, while $\mathbb{E}[x_i] = \alpha_i$. Therefore, the variance of x_i is decreasing in k, but the mean is unaffected by changes in k.

[57] I allow for a shift in initial support for Democrats (i.e. ψ_{dem}) due to the fact that the measurement of Democratic support $r_1(Z)$ is systematically lower than that of Republican's. While this may be an accurate measurement of initial support, there may also be systematic undermeasurement of Democratic support. Including ψ_{dem} allows us to control for this possibility.

[58] Note that this particular structure assumes that if a candidate is an incumbent, the boost in his initial support ψ_{inc} is taken from what would be his opponent's initial support, all else equal. Results are robust to assuming that the shift in support comes from swing voters, i.e. $Inc_i = 1$ if i is an incumbent and 0 otherwise.

but the econometrician does not. I further assume S_i is Bernoulli distributed with:

$$Pr\{S_i = 1\} = \frac{\exp\{\beta_c\}}{1 + \exp\{\beta_c\}}.\text{[59]} \tag{40}$$

Given a realization of initial support $(\widehat{r}_1, \widehat{r}_2, \widehat{R})$ and candidates' budgets, i's problem is:

$$\begin{aligned} \max_{x_i} \;\; & \widehat{r}_i exp\{-\alpha_1 x_i - \alpha_2 x_j\} - \widehat{r}_j exp\{-\alpha_1 x_j - \alpha_2 x_i\} \\ & \qquad + \widehat{R}\left(\frac{(1 + B_i - x_i)^{1/\gamma} - (1 + B_j - x_j)^{1/\gamma}}{(1 + B_i - x_i)^{1/\gamma} + (1 + B_j - x_j)^{1/\gamma}}\right) \\ \text{s.t. } \; & x_i \in [0, B_i] \end{aligned} \tag{41}$$

Since I assume the candidates observe $(\widehat{r}_1, \widehat{r}_2, \widehat{R})$, there is no change in the information set of the players, and therefore the existence and uniqueness follows. The distribution of initial support will generate a distribution of negative campaigning for each candidate, and thus will generate a likelihood function.

2.5 Data

I implement the model by using data from 2000, 2004, and 2008 House of Representatives and Senate races. In order to infer campaign strategies, I use data on political advertising tone from the Wisconsin Advertising Project. I merge this data with contribution data from the Database on Ideology, Money in Politics, and Elections (DIME), individual level opinion data from American National Election Studies (ANES), publicly available House and Senate election results, and demographic data from the 2000 Census and the American Community Survey (ACS).[60] I also hand collect some relevant data, such as incumbency status, for each race.

[59] This functional form is strictly to keep the probability of being skilled between zero and one.

[60] These data are gathered from Bonica [2013], Goldstein et al. [2002], Goldstein and Rivlin [2007], Goldstein and Rivlin [2011], University of Michigan [2000], University of Michigan [2004], University of Michigan [2008], and U.S. Census Bureau [2002].

2.5.1 Advertising and Elections

WiscAds uses a technology that monitors the transmission of 35 national networks in the top Designated Market Areas (DMA). A DMA is a geographical region where individuals receive the same TV content and it is the smallest geographical unit in which a politician can buy air time. Every time there is a political advertisement in these markets, WiscAds captures it. A team of students research assistants then analyzes the storyboard of the advertisement to code it into the dataset. I therefore have detailed information on each advertisement: tone (i.e. whether it is positive or negative), exact date and time, station, and ad sponsor, among other things. It also importantly includes the candidate, party, or group for which the ad was aired in support. The dataset also contains an estimated cost variable. There are three ad tone types in the data: positive, contrast, and attack. I follow the convention in the literature and define negative ads as those classified as either contrast or attack ads.

The sample is limited to the geographical borders of WiscAds for each year. I merge the counties covered by WiscAds with the counties in each district.[61] Over the span of the three election cycles I consider, WiscAds should in principle cover 1,390 races. However, none of the candidates running in 814 of these elections purchased airtime and hence are not in the WiscAds dataset.[62] I therefore have no information on the campaign strategies of the candidates in those 814 elections. I also drop the 20 elections in which ads were purchased that were held in Louisiana, since this state employs a runoff system, and the 23 elections where a third-party candidate was a winner or a runner-up due to the method by which I estimate the supporters of each candidate. Finally, I drop 183 elections for which at least one candidate received positive contributions, but did not advertise, since I have no way to infer overall campaign strategies without observing advertisements. This leaves us with 361 elections over the three cycles.

Details of the type of elections covered in the final sample are given in Table 24. I have between 20 and 23 Senate elections for each year, and about 85 House elections for 2000 and 2004. For 2008, there are 126 elections included. Among the 814 elections in which neither candidate had a television advertisement, 758 were House elections. Since these elections tend to be less competitive, advertising in general is less common. Hence, I observe a large fraction of the House elections where at least one candidate purchased TV ads. Within the final sample, I have 200 Republican wins versus

[61] I do not observe this directly from WiscAds. I obtained the list of counties in each DMA and year from SRDS [2000, 2004, 2008].

[62] Among these elections, 56 are Senate and 758 are House races.

Table 24
SAMPLE ELECTIONS BY RACE TYPE AND YEAR

YEAR	RACES	
	Senate	House
2000	23	82
2004	20	90
2008	20	126
Total	63	298

161 Democrat wins. The average winning margin for both parties is almost 18%. The summary of election results is available in Table 25. Table 26 shows the distribution of incumbency status in the sample. There are 16 elections for an open seat in the Senate and 61 in the House. The remaining 47 Senate races and 237 House races involve an incumbent. Given that the sample period covers a relatively successful period for Republicans, there are 180 races with a Republican incumbent, and 104 with a Democratic incumbent.

Table 25
ELECTION DATA - ALL YEARS

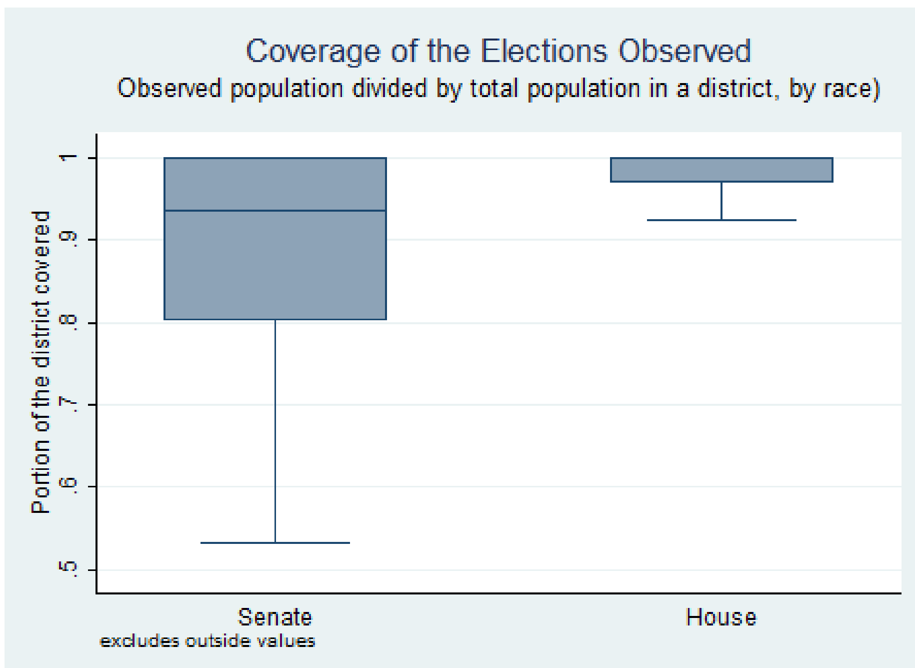
	Democrats	Republicans
Winning Margin	17.75	17.82
	(14.87)	(11.61)
Winner	161	200
Total Races	361	

Table 26
INCUMBENCY STATUS

TYPE	RACES	
	Senate	House
Open Seat	16	61
Democrat Incumbent	19	85
Republican Incumbent	28	152
Total	63	298

In 2000, WiscAds followed only the top 75 DMAs, in 2004 the top 100 DMAs, and in 2008 all of the 210 DMAs, hence why the 2008 sample includes more races than previous years. Note that in each case, the DMAs cover a very large portion of the U.S. population: in 2000, the top 75 DMAs accounted for 78% of the population. In 2008, nearly the entire population is covered. However, since the observed DMAs do not exactly cover the entire U.S. population, I only partially observe the campaigns for some elections – that is, there are some races where I observe political television advertisements only in *some* of the counties within the district or state in which the election is held. To quantify the degree to which this occurs, for each race, I compute the size of the population in the intersection of the DMAs I observe and the Congressional district (for House races) or the state (for Senate races) of the election, and divide by the district or state size. I find that, on average, the dataset contains 91% of the population in a district or state. The boxplot for this measure is displayed in Figure 16. For House races, the 75th percentile is above 90% whereas for the Senate, it is around 53%. That is, for 75% of House races, at least 90% of the population is in a DMAs I observe. The median coverage for Senate races is 93.5%, while for the House it is 100%.

Figure 16
Party Support Boxplot



Although the dataset covers a large portion of each race, the incompleteness of the data could still be problematic for empirical implementation. The potential issue is the fact that I expect the top DMAs to contain more populous urban areas which may be more Democratic than the rest of the country. Hence, the areas I observe might have a Democratic bias, which could potentially affect the strategies of the candidates. In carrying out the empirical analysis, I assume that the candidate has the same campaign strategy across the district.

To investigate the degree to which campaign strategies may differ across different populations, I analyze the variance in campaign strategy for elections in which advertising occurs in more than one DMA. In particular, let $n^d_{i,e}$ be the cost of all negative advertisements aired in DMA d by candidate i in election e, and let $t^d_{i,e}$ be the cost of all advertisements in d aired by this candidate. I denote the campaign strategy in this DMA for candidate i in election e as:

$$N^d_{i,e} = \frac{n^d_{i,e}}{t^d_{i,e}}.$$

Letting $N_{i,e} = \frac{n_{i,e}}{t_{i,e}}$ denote the campaign strategy for candidate i in election e across *all* DMAs. Finally, I compute for each DMA the absolute deviation from the mean, $|N^d_{i,e} - N_{i,e}|$. Since air time is purchased in bulk, I consider campaigns that placed more than 500 ads in at least two different DMAs. Among these campaigns, the median absolute deviation is 0.034 for Republicans and 0.021 for Democrats. The 75th percentile is 0.093 for Republicans and 0.088 for Democrats. While there may be systematic differences between DMAs in the sample and outside the sample, this evidence is suggestive of the idea that strategies do not change dramatically across different populations.

2.5.2 Estimates of Initial Voter Support

An important determinant of the equilibrium of the model is the measure of the voter types in each election, in particular $(r_1(Z), r_2(Z), R(Z))$, where Z is the distribution of demographics in the district or state in which a given election is held in a given year. In order to estimate those parameters, I use the joint distribution of the demographic characteristics for each district and state for the years 2000, 2004, and 2008. I construct this data using the 2000 census and the American Community Survey (ACS) for 2005 and 2008.[63] Due to data limitations I consider only race, gender, and income.

The next step is to estimate the probability an individual supports a party (or not) conditional on demographic characteristics. I use the ANES survey data to estimate the probability of identifying with a particular party conditional on demographic characteristics. In the ANES, each surveyed individual is asked about his or her relevant demographic characteristics of race, gender, and income. Furthermore, to identify party support, each individual is asked the following question:

[63] I use the 2005 ACS for the 2004 elections since there is no 2004 ACS. Also note that, for 2008, the ACS is the three-year estimates, which allows analysis at a smaller geographic area.

Generally speaking, do you think of yourself as a Republican, a Democrat, or an Independent? Would you call yourself a strong Democrat/Republican or a not very strong Democrat/Republican? Do you think of yourself as closer to the Republican Party or to the Democratic party?

I consider an individual to be an ideological voter if he answers this question with a strong partisan preference. The summary statistics for the ANES data are given in Tables 27, 28, and 29 for years 2000, 2004, and 2008 respectively. These data are given for the entire ANES samples, as well as broken down by party identification.

Table 27
ANES 2000 - Summary Statistics
(By party identification)

	Overall		Democrat		Swing		Republican	
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
Black	0.119	0.324	0.236	0.425	0.077	0.267	0.023	0.149
Female	0.551	0.498	0.603	0.490	0.538	0.499	0.499	0.501
$LowInc$	0.518	0.500	0.575	0.495	0.513	0.500	0.443	0.497
$MidInc$	0.412	0.492	0.378	0.485	0.430	0.496	0.427	0.495
$HighInc$	0.071	0.256	0.047	0.212	0.057	0.232	0.129	0.336
# of people	1577		552		635		390	

Note: All variables dummies. $LowInc$ is $[0, 50K]$, $MidInc$ is $[50K, 75K]$, $HighInc$ is $[75K, \infty)$

Table 28
ANES 2004 - Summary Statistics
(By party identification)

	Overall		Democrat		Swing		Republican	
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
Black	0.156	0.363	0.307	0.462	0.143	0.351	0.007	0.084
Female	0.516	0.500	0.607	0.489	0.463	0.499	0.486	0.501
$LowInc$	0.507	0.500	0.546	0.499	0.536	0.499	0.428	0.496
$MidInc$	0.404	0.491	0.396	0.490	0.385	0.487	0.438	0.497
$HighInc$	0.089	0.284	0.057	0.233	0.079	0.271	0.135	0.342
# of people	1577		552		635		390	

Note: All variables dummies. $LowInc$ is $[0, 50K]$, $MidInc$ is $[50K, 75K]$, $HighInc$ is $[75K, \infty)$

Table 29
ANES 2008 - Summary Statistics
(By party identification)

	Overall		Democrat		Swing		Republican	
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
Black	0.121	0.326	0.245	0.430	0.085	0.280	0.014	0.117
Female	0.545	0.498	0.620	0.486	0.488	0.500	0.535	0.499
$LowInc$	0.482	0.500	0.533	0.499	0.516	0.500	0.361	0.481
$MidInc$	0.401	0.490	0.403	0.491	0.392	0.488	0.414	0.493
$HighInc$	0.117	0.322	0.065	0.246	0.092	0.289	0.225	0.418
# of people	1577		552		635		390	

Note: All variables dummies. $LowInc$ is $[0, 50K]$, $MidInc$ is $[50K, 75K]$, $HighInc$ is $[75K, \infty)$

I estimate the probability that an individual is an ideological voter for a specific party or a swing voter using a multinomial logistic regression. I use the following variables in the estimation. $ID \in \{0, 1, 2\}$ is the party identification where 0 indicates that an individual is a swing voter, 1 indicates that the individual is an ideological Democrat and 2 a Republican. The explanatory variables for individual i in vector z_i are $(black_i, female_i, inc_i^0, inc_i^1, inc_i^2)$. The indicators $inc_0 = 1$ if the individual's income is less than \$50,000, $inc_1 = 1$ if the individual's income is between \$50,000 and \$75,000, and $inc_2 = 1$ if the income is greater than \$75,000. Setting the base outcome as being a swing voter, I estimate the vector of coefficients $\{\beta_k\}_{k=1}^2$ and get the following probabilities for each individual:

$$Pr(ID = k|z_i) = \frac{exp(\beta_k z_i)}{1 + \sum_{i=1}^2 exp(\beta_k z_i)} \quad \text{for} \quad k \in \{1, 2\} \tag{42}$$

and $Pr(ID = 0|z_i) = 1 - Pr(ID = 1|z_i) - Pr(ID = 2|z_i)$. I separately estimate coefficients for each year. These estimation results are given in Tables 30, 31, and 32 for years 2000, 2004, and 2008 respectively.

Table 30
MULTINOMIAL LOGIT RESULTS - 2000

	Democrat		Republican	
	β_D	St. Dev	β_R	St. Dev
constant	-.51*	.261	.281	.204
black	1.3***	.178	-1.22***	.38
female	.03*	.125	-.06	.136
$0 < Inc < 50K$	-.075	.275	-.88***	.24
$50K < Inc < 75K$	-.512	.275	-.79***	.240
Observations	1577			
Psuedo R^2	.0429			
LR-χ^2	145.06			

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Base outcome is Independent.

Table 31
MULTINOMIAL LOGIT RESULTS - 2004

	Democrat		Republican	
	β_D	St. Dev	β_R	St. Dev
constant	-.725	.289	.248	.234
black	1.03***	.186	-3.07***	.68
female	.674***	.157	-.20	.159
$0 < Inc < 50K$	-.125	.312	-.700**	.266
$50K < Inc < 75K$	-.064	.312	-.326	.212
Observations	1088			
Psuedo R^2	.0712			
LR-χ^2	169.26			

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Base outcome is Independent.

Table 32

Multinomial Logit Results - 2008

	Democrat		Republican	
	β_D	St. Dev	β_R	St. Dev
constant	-.808***	.191	.376	.148
black	1.25***	.151	-1.78***	.385
female	.556***	.108	-.41***	.116
$0 < Inc < 50K$	.08	.206	-1.350***	.266
$50K < Inc < 75K$	.229	.191	-.891***	.170
Observations	2136			
Psuedo R^2	.0637			
LR-χ^2	294.75			

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Base outcome is Independent.

Finally, let Z denote the empirical joint distribution of demographics in a given election. Each element Z^i is the probability a random individual in the district has set of demographic characteristics i, where i is some combination of included characteristics. This distribution estimated from the Census and the ACS for the relevant year. Let $\overline{N} = 12$ denote the total number of possible demographic groupings. Then, I define

$$r_k(Z) = \sum_{i=1}^{\overline{N}} Pr(ID = k|i)Z^i \tag{43}$$

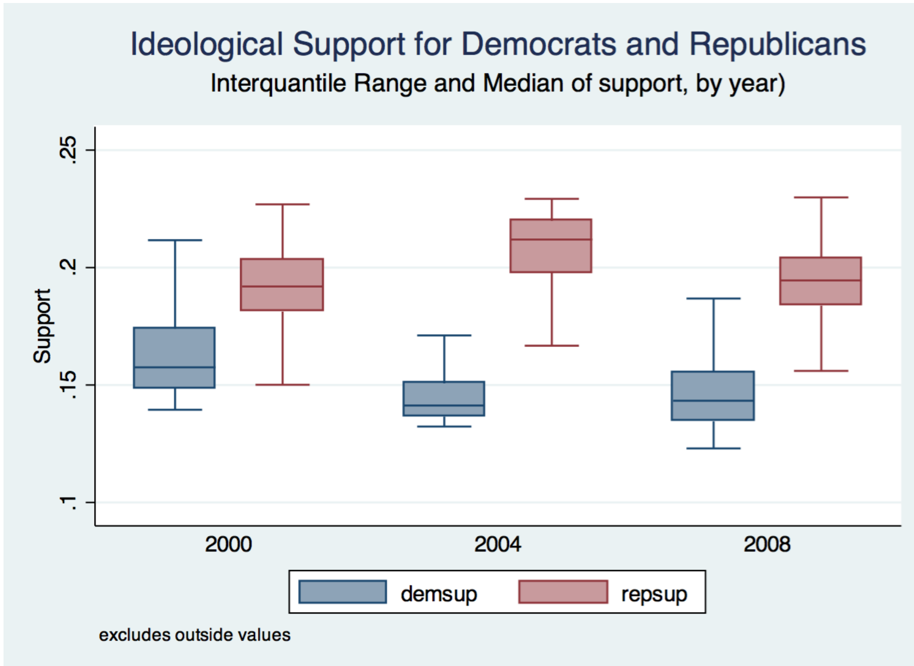
for $k \in \{1, 2\}$. Lastly, $R(Z) = 1 - r_1(Z) - r_2(Z)$ corresponds to swing voters in the model. Details about this variable can be found in Table 33, while Figure 17 provides the boxplot.

Table 33

Ideological Support for Parties - All Years

	Democratic Support		Republican Support	
	Mean	St. Dev	Mean	St. Dev
2000	0.291	0.030	0.342	0.025
2004	0.272	0.045	0.370	0.047
2008	0.259	0.038	0.385	0.037

Figure 17
ESTIMATED PARTY SUPPORT BOXPLOT



2.5.3 Candidate Budgets

I measure budgets as the total real receipts of a candidate over the campaign cycle.[64] This data comes from DIME, which extracts the receipts from Federal Election Commission filings. In the model, each candidate is endowed with a budget to allocate between positive and negative spending. While I only observe positivity and negativity for television advertising, I use this to infer overall campaign strategy. As Gordon and Hartmann [2013b] note, television advertising comprises the largest component of media spending for political campaigns. Furthermore, for both parties, television ads generally constitute a considerable element of candidate budgets, as well. Table 34 shows the total receipts by party, and the average proportion of budgets devoted to television ads in the sample is 46.0% for Democrats and 37.9% for Republicans. I therefore use receipts as the measure of candidate budgets and the breakdown of television advertising tone as the measure of campaign

[64] I note that, while in principle candidates can borrow or save campaign funds, in the sample saving and borrowing constitute a small fraction of total receipts. Among campaigns whose receipts exceed disbursements, the median savings rate, which is $\frac{receipts-disbursements}{receipts}$, for Republicans is 1.3% and for Democrats is 1.4%. Furthermore, among campaigns whose disbursements exceed receipts, the median savings rate for Republicans is -2.2%, and for Democrats is -4.4%. Note also that since budgets may be spent on items other than advertising, I am assuming that the tone of advertisements reflects the overall negativity of the campaign.

strategy.

Table 34
TOTAL RECEIPTS BY PARTY

	Democrats		Republicans	
	Mean	St. Dev	Mean	St. Dev
Total receipts	2,633,155	5,193,242	2,582,212	3,646,026
Ads as % of receipts	46.0%	0.403	37.9%	0.327

Note: Totals in nominal U.S. dollars.

2.5.4 Summary Statistics

I now document several of the main summary statistics and regularities in the data. Table 35 breaks down the total advertisements and advertisement tone by party. Democratic candidates placed, on average, 2,151 ads in a race, while Republicans placed about 1,963. The average number of Democratic negative ads in a race is 1,465, while for Republican candidates it is 1,274. On average, a Democratic (Republican) candidate's negative ads amount to 58.1% (54.6%) of his total ads. There is not a significant difference either in the total number of ads aired or their average negativity across parties. I observe the same pattern for the estimated costs, as seen in Table 36.

Table 35
NUMBER OF ADS AND AD TYPES BY PARTY - ALL YEARS

	Democrats		Republicans	
	Mean	St. Dev	Mean	St. Dev
Total Ads	2150.8	3440.3	1962.6	3467.3
Positive Ads	685.8	1052.7	688.9	1305.4
Contrast Ads	592.2	1092.1	443.0	925.7
Attack Ads	872.5	1979.1	830.6	1913.2
Negative Ads	1464.7	2893.3	1273.7	2536.4
% of Neg Ads	58.1%	.361	54.6%	.353

Note: "% Neg Ads" only for those with positive amount of advertising.

Table 36
Ad Costs by Party

	Democrats		Republicans	
	Mean	St. Dev	Mean	St. Dev
Total cost	1,165,178	1,920,977	1,050,046	1,745,814
Neg Ad cost	798,735	1,412,325	693,226	1,253,378
% cost of Neg Ads	58.1%	.361	54.6%	.353
Campaigns with no spending	18		13	

Note: Totals in nominal U.S. dollars.

While the broad strategies of candidates do not vary significantly across parties, the strategies do differ between incumbents and challengers, and in close races versus landslides. Table 37 provides the total ads and ad types by incumbents and challengers, whereas Table 38 does the same for estimated costs. Incumbents, on average, place about 350 more advertisements in each race and spend $250,000 more on television advertising than challengers. This stark difference is caused in part by fewer funds received by challengers. Table 39 shows that incumbents in the sample receive on average $1.1 million more than challengers. The data also show that incumbents allocate most of their air time to positive advertisements: 38.8% of incumbents' total advertising spending goes to negative ads, whereas for challengers this number is 71.5%.

Table 37
Number of Ads and Ad Types by Incumbency - All Years

	Incumbents		Challengers	
	Mean	St. Dev	Mean	St. Dev
Total Ads	2015.2	3545.9	1657.3	3353.7
Positive Ads	897.1	1433.2	348.0	682.3
Contrast Ads	368.2	840.3	536.9	1083.8
Attack Ads	749.5	2017.6	772.3	2035.6
Negative Ads	1117.7	2592.8	1309.2	2939.1
% of Neg Ads	38.6%	.352	71.5%	.306

Note: "% Neg Ads" only for those with positive amount of advertising.

Table 38
AD COSTS BY INCUMBENCY

	Incumbents		Challengers	
	Mean	St. Dev	Mean	St. Dev
Total cost	1,075,472	1,620,455	826,419	1,483,449
Neg Ad cost	609,754	1,187,704	662,548	1,280,989
% cost of Neg Ads	38.8%	0.354	71.5%	0.307
# in the data	284		284	
# with no spending	0		30	

Note: Totals in nominal U.S. dollars.

Table 39
TOTAL RECEIPTS BY INCUMBENCY

	Incumbents		Challengers	
	Mean	St. Dev	Mean	St. Dev
Total receipts	2,830,316	3,535,365	1,733,865	2,720,863
Ads as % of receipts	34.3%	0.283	41.8%	0.380

Note: Totals in nominal U.S. dollars.

Next, I classify the elections according to the ex-post vote margins and analyze the differences in advertising choices and budgets. I consider an election to be close if the winning margin is less than 5 percentage points and a blowout if the margin is larger than 20 percentage points. Then, I look at the difference between the sum of the total ads (Table 40), money spent (Table 41), and receipts (Table 42) by both campaigns. In landslide elections, of which there are 130 observations, the mean number of ads by both candidates is 1,475. In the 61 close elections I observe, the mean number of ads is 8,385, around 5.5 times as much as in landslides. Furthermore, campaigns tend to be much more negative in close elections. Around 74% of all ads aired in such elections were negative, compared to 26.5% in landslides. Similarly stark differences remain when comparing negativity in terms of money spent. Finally, as expected, total receipts in close races are much larger than in landslide elections, with around \$8.1 million in the former as compared to \$3.3 million in the latter.

Table 40

NUMBER OF ADS AND AD TYPES BY CLOSENESS OF ELECTION

	Close (<5% margin)		Landslide (>20% margin)	
	Mean	St. Dev	Mean	St. Dev
Total Ads	8385.0	9148.0	1474.5	2185.3
Positive Ads	2252.5	2726.5	891.1	1352.1
Contrast Ads	2365.3	2787.3	303.6	682.8
Attack Ads	3767.2	4606.2	279.8	663.2
Negative Ads	6132.5	6720.3	583.5	1189.2
% of Neg Ads	74.3%	0.169	26.5%	0.285
Total Elections	61		130	

Note: Totals are for both candidates in nominal U.S. dollars. "% of Neg Ads" only for those who had positive amount of advertising.

Table 41

AD COSTS BY CLOSENESS OF ELECTION

	Close (<5% margin)		Landslide (>20% margin)	
	Mean	St. Dev	Mean	St. Dev
Total cost	4,502,892	4,694,231	708,389	1,227,679
Neg Ad cost	3,387,087	3,402,736	262,022	566,929
% cost of Neg Ads	75.5%	0.171	24.9%	0.284
# in the data	61		130	

Note: Totals are for both candidates in nominal U.S. dollars.

Table 42

TOTAL RECEIPTS BY CLOSENESS OF ELECTION

	Close (<5% margin)		Landslide (>20% margin)	
	Mean	St. Dev	Mean	St. Dev
Total receipts	8,114,928	10,921,429	3,331,386	5,090,576
Ads as % of receipts	65.7%	0.362	19.4%	0.190
# in the data	61		130	

Note: Totals are for both candidates in nominal U.S. dollars.

Another interesting feature of the data is the presence of corner solutions. There are many elections where a candidate's strategy is to fill his airtime solely with positive or negative advertisements. Detailed information about this, broken down by the party, is given in Table 43. In 176 of the elections, I observe both candidates allocating their air time to both positive and negative ads. For the rest, there are either no ads by one politician, or at least one candidate chooses a corner strategy. Table 44 breaks down the selected strategies by incumbency (among those elections involving an incumbent). While 97 out of the 284 incumbents in the sample chose only positive ads, 70 challengers chose exclusively negative, again reflecting the relative propensity of a challenger to campaign negatively. Only 16 incumbents went entirely negative, and only 20 challengers went entirely positive.

Table 43
Distribution of Ad Strategies by Party

		Republicans				
		All positive	Interior	All Negative	No Ads	**Total**
Democrats	All positive	20	21	11	12	64
	Interior	14	176	21	1	212
	All negative	17	40	10	0	67
	No Ads	17	1	0	0	18
	Total	68	238	42	13	361

Table 44
Distribution of Ad Strategies by Incumbency

		Challengers				
		All positive	Interior	All Negative	No Ads	**Total**
Incumbents	All positive	14	29	26	28	97
	Interior	4	126	39	2	171
	All negative	2	9	5	0	16
	No Ads	0	0	0	0	0
	Total	20	164	70	30	284

2.6 Calibration and Fit

Given parameter values, for each election I can simulate draws from the initial support distribution and solve for the campaigning equilibrium. The full set of parameters is:

$$\Theta = \{\gamma, \alpha_1, \alpha_2, \beta_c, \psi_s, \psi_{dem}, \psi_{inc}, k\}$$

To select parameters, I calibrate the model to roughly match the various conditional means of negative campaigning, in particular in the aggregate and by party. I also roughly match the distributions of observed campaigning strategies. In order to calibrate the model, first consider how the various parameters differentially affect the observed distribution of outcomes. First, as the parameter ψ_{inc} increases, we will tend to observe higher initial support for incumbents and lower initial support for challengers. In turn, this will tend to generate more negativity from challengers and less negativity from incumbents. However, it will not affect negativity in open seat races. The parameter ψ_{dem} has a similar effect, except with a larger value of ψ_{dem} generating more negativity by Democrats and less by Republicans.

Next, consider the parameters affecting the marginal productivity of negative and positive campaigning, $(\alpha_1, \alpha_2, \gamma)$. Note that a proportional increase in α_1 and α_2 tends to make negative campaigning more productive. At first glance, it appears that γ could simultaneously be adjusted to keep the *relative* productivities of negativity and positivity the same, and thus not change the equilibria. Given the structure of the game, changes in γ have a differential effect on outcomes depending on the budget sizes. As an illustration, consider elections in which only one candidate has a positive budget. Without loss of generality, let candidate 1 have the positive budget. The mass of corner solutions at exclusively positive campaigning in these elections is given by the measure of $\tilde{r}_1$ and $\tilde{r}_2$ such that the marginal benefit of negative campaigning less than the marginal benefit of positive campaigning at $x_1 = 0$:

$$Pr\left\{\tilde{r}_2 < \frac{\tilde{r}_1\alpha_1}{\alpha_2} + \frac{(1-\tilde{r}_1-\tilde{r}_2)}{2\gamma\alpha_2}\frac{(1+B_1)^{1/\gamma-1}}{(1+(1+B_1)^{1/\gamma})^2}\right\}.$$

Note that given the distribution of $(\tilde{r}_1, \tilde{r}_2, \tilde{R})$ is Dirichlet, which has full support on the two-dimensional simplex, this mass will be strictly positive. In the data, I observe, even among elections with only one positive budget, a wide range in values of B_1 (or B_2), as well as variation in the estimates of $r_1(Z_e)$, $r_2(Z_e)$, and $R(Z_e)$, incumbency status. For instance, for these elections the minimum budget is $\$269,000$ (in real 2000 dollars) while the maximum is \$3.3 million. The lowest budget is in the 6.7th percentile among all positive budgets, while the highest is in the 85th percentile. Furthermore, $r_1(Z_e)$ in elections with one budget ranges from 0.123 to 0.218, while $r_2(Z_e)$ ranges from 0.091 to 0.223. In the full sample, $r_1(Z_e)$ ranges between 0.123 and 0.282, while $r_2(Z_e)$ ranges between 0.087 and 0.230.

Since the sample of one-budget elections features wide variation in B_1 and demographics. These elections will have the same probability distribution for initial support. Consider B_1 approaching 0. The above mass of corner solutions in those elections is approximately

$$Pr\left\{\tilde{r}_2 < \frac{\tilde{r}_1\alpha_1}{\alpha_2} + \frac{1-\tilde{r}_1-\tilde{r}_2}{8\gamma\alpha_2}\right\}.$$

Therefore, any other combination of parameters α_1, α_2, and γ that generate the same mass should have $\alpha_1' = \kappa\alpha_1$, $\alpha_2' = \kappa\alpha_2$, and $\gamma' = \frac{1}{\kappa}\gamma$.

Now consider a similar race but with a large budget B_1'. The mass of corner solutions at $x_1 = 0$ in that election is given by

$$Pr\left\{\tilde{r}_2 < \frac{\tilde{r}_1\alpha_1}{\alpha_2} + \frac{1-\tilde{r}_1-\tilde{r}_2}{2\gamma\alpha_2}\frac{(1+B_1')^{1/\gamma-1}}{(1+(1+B_1')^{1/\gamma})^2}\right\}$$

Now, evaluated at the above defined α_1', α_2', and γ', I have the new mass to be

$$Pr\left\{\tilde{r}_2 < \frac{\tilde{r}_1\alpha_1}{\alpha_2} + \frac{1-\tilde{r}_1-\tilde{r}_2}{2\gamma\alpha_2}\frac{(1+B_1')^{\kappa/\gamma-1}}{(1+(1+B_1')^{\kappa/\gamma})^2}\right\}.$$

The last term, reflecting the marginal benefit of positive campaigning evaluated at the corner, is now more affected by the change in the γ parameter than in the low budget case, and therefore the probability mass of corner solutions will be different. Therefore, changes in the of values of $(\alpha_1, \alpha_2, \gamma)$ have a differential effect on the proportion of elections with only positive campaigning, which helps in calibrating the parameters.

Finally, consider parameters k, β_c, and ψ_s. These parameters all govern the spread in the distribution of initial support. As discussed in the empirical model section, a lower value of k corresponds to a higher variance in initial support, which generates wider variation in campaigning choices. Holding fixed incumbency status, one can think of drawing initial support from a mixture distribution, where k governs the variance of all the underlying distributions, ψ_s governs the relative means of the underlying distributions, and β_c governs the probability of drawing from each distribution.

I argue that these parameters affect the variation of initial support, and therefore campaigning strategies, in different ways. First note that, given parameters, the draw for initial support is a mixture of four Dirichlet distributions. In particular, denoting the probability of being skilled as $\tilde{\beta}_c = \frac{e^{\beta_c}}{1+e^{\beta_c}} \in [0, 1]$, and ignoring ψ_{dem} and ψ_{inc} for notational simplicity:

- With probability $\tilde{\beta}_c^2$, initial support is drawn from $Dir(k(r_1(Z_e)+\psi_s), k(r_2(Z_e)+\psi_s), k(R(Z_e)-2\psi_s))$;
- W.p. $\tilde{\beta}_c(1-\tilde{\beta}_c)$, initial support is drawn from $Dir(kr_1(Z_e), k(r_2(Z_e)+\psi_s), k(R(Z_e)-\psi_s))$;
- W.p. $\tilde{\beta}_c(1-\tilde{\beta}_c)$, initial support is drawn from $Dir(k(r_1(Z_e)+\psi_s), kr_2(Z_e), k(R(Z_e)-\psi_s))$;
- W.p. $(1-\tilde{\beta}_c)^2$, initial support is drawn from $Dir(kr_1(Z_e), kr_2(Z_e), kR(Z_e))$.

Given parameters $(k, \tilde{\beta}_c, \psi_s)$, note that the mean of r_1 under the mixture distribution is given by:

$$\tilde{\beta}_c(r_1(Z_e)+\psi_s)+(1-\tilde{\beta}_c)r_1(Z_e) = r_1(Z_e)+\tilde{\beta}_c\psi_s, \tag{44}$$

and the variance of r_1 is given by:

$$\tilde{\beta}_c\left(\frac{(r_1(Z_e)+\psi_s)(1-r_1(Z_e)-\psi_s)}{k+1}\right)+(1-\tilde{\beta}_c)\left(\frac{r_1(Z_e)(1-r_1(Z_e))}{k+1}\right)+\tilde{\beta}_c(1-\tilde{\beta}_c)\psi_s^2. \tag{45}$$

To show that k, $\tilde{\beta}_c$, and ψ_s affect the distribution in different ways, I show that for $\psi_s \neq 0$, any two different sets of parameters $(k, \tilde{\beta}_c, \psi_s)$ and $(k', \tilde{\beta}_c', \psi_s')$ generate a different distribution for initial support.[65] I prove this by contradiction. Consider two different parameter values and let $\tilde{\beta}_c' = a\tilde{\beta}_c$. For the mean to be the same under both distributions, it must be that $\psi_s' = \frac{\psi_s}{a}$, as follows from (44). Additionally, let the variance under parameters $(k, \tilde{\beta}_c, \psi_s)$ as V_1 (given in (45)), and the variance under the alternative parameters be given by V_1', or:

$$a\tilde{\beta}_c\left(\frac{(r_1(Z_e)+\frac{\psi_s}{a})(1-r_1(Z_e)-\frac{\psi_s}{a})}{k'+1}\right)+(1-a\tilde{\beta}_c)\left(\frac{r_1(Z_e)(1-r_1(Z_e))}{k'+1}\right)+a\tilde{\beta}_c(1-a\tilde{\beta}_c)\frac{\psi_s^2}{a^2},$$

which is non-linear in both a and k. In order for $V_1 = V_1'$, I can rearrange and solve for k', which yields:

$$k' = \frac{a\tilde{\beta}_c(r_1(Z_e)+\frac{\psi_s}{a})(1-r_1(Z_e)-\frac{\psi_s}{a})+(1-a\tilde{\beta}_c)r_1(Z_e)(1-r_1(Z_e))}{V_1-a\tilde{\beta}_c(1-a\tilde{\beta}_c)\frac{\psi_s^2}{a^2}}-1. \tag{46}$$

Thus, it must be that if two different parameter vectors have the same distribution of initial support, with $\tilde{\beta}_c' = a\tilde{\beta}_c$, then it is a necessary condition that $\psi_s' = \frac{\psi_s}{a}$, and k' must satisfy (46).

[65] I proceed assuming $\psi_s > 0$, since if $\psi_s = 0$, then the value of $\tilde{\beta}_s$ is irrelevant as being skilled would not affect anything.

However, the joint density of initial support under parameters $(k, \tilde{\beta}_c, \psi_s)$ can be written as:

$$g(r_1, r_2, R; k, \tilde{\beta}_c, \psi_s) = \sum_{i,j=0}^{1} \tilde{\beta}_c^{i+j}(1-\tilde{\beta}_c)^{2-i-j} \left(r_1^{k(r_1(Z_e)+\psi_s i)} r_2^{k(r_2(Z_e)+\psi_s j)} R^{k(R(Z_e)-\psi_s(i+j))} \right)$$
$$\times \frac{1}{B(k(r_1(Z_e)+\psi_s i), k(r_2(Z_e)+\psi_s j), k(1-r_1(Z_e)-r_2(Z_e)-\psi_s(i+j)))}.$$

where $B(a,b,c)$ is the beta function, and with $0 < r_1 + r_2 < 1$ and $R = 1 - r_1 - r_2$. Alternatively, under $(k', \tilde{\beta}'_c, \psi'_s)$ as specified above, the density is given by:

$$g(r_1, r_2, R; k', \tilde{\beta}'_c, \psi'_s) = \sum_{i,j=0}^{1} (a\tilde{\beta}_c)^{i+j}(1-a\tilde{\beta}_c)^{2-i-j} \times$$
$$\left(r_1^{k'(r_1(Z_e)+\frac{\psi_s}{a} i)} r_2^{k'(r_2(Z_e)+\frac{\psi_s}{a} j)} R^{k'(R(Z_e)-\frac{\psi_s}{a}(i+j))} \right) \times$$
$$\frac{1}{B(k'(r_1(Z_e)+\frac{\psi_s}{a} i), k'(r_2(Z_e)+\frac{\psi_s}{a} j), k'(1-r_1(Z_e)-r_2(Z_e)-\frac{\psi_s}{a}(i+j)))}.$$

Clearly, this is different from the density under the original parameters. This contradicts the supposition that the mixture distribution was identical under both sets of parameters. Therefore, changes in these parameters will affect the distributions of initial support, and therefore campaigning strategies, differentially.

Presented in Table 45 are the calibrated parameters. I note that the calibrated model implies essentially no "boomerang" effect from negative campaigning on a candidate's own supporters. I also let $\widehat{\psi}_{inc} = 0.022$, which corresponds approximately to a 4.4% advantage in initial support.[66] Additionally, the shifter $\widehat{\psi}_{inc} = 0.1157$ indicates an 11.6% increase in initial support, conditional on being skilled. The probability of being skilled, which is governed by β_c, is approximately 97%. Finally, the calibrated value of $\widehat{\psi}_{dem} = 0.045$ implies that the estimates for Democratic support based solely on demographics, $r_1(Z_e)$, are persistently low. Incidentally, the mean of $r_1(Z_e) + \widehat{\psi}_{dem}$ across all elections in the sample is 0.1987, which is approximately equal to the mean of $r_2(Z_e)$, given by 0.1958.

To investigate the fit of the calibrated model, Table 46 shows the average proportion of negative campaigning in the data as compared to simulations, while Figure 18 shows the distributions of negativity. The top two graphs are for Democrats and the bottom two are for Republicans, while within each pair the top presents the distribution in the data and the bottom presents the simulated

[66] This is because, by assumption, if candidate i is an incumbent and candidate j is the challenger, then mean initial support for candidate i increases by $\widehat{\psi}_{inc}$ and for candidate j decreases by $\widehat{\psi}_{inc}$.

Table 45
Calibrated Parameter

Parameter	Value
γ	1.0297
α_1	0.0001
α_2	0.5922
β_c	3.1837
ψ_s	0.1157
ψ_{inc}	0.0215
ψ_{dem}	0.0446
k	162.65

distribution. Overall, the model captures several important features of the data, both quantitatively and qualitatively. In the data, overall a bit more than half (56.7%) of advertisements are negative, with Democrats performing more negative campaigning than Republicans, by about 3.8 percentage points. The model captures these features as well, only slightly predicting both overall negativity and negativity by party by approximately 4 percentage points. In addition, the data shows that challengers tend to go significantly more negative than incumbents by a margin of 71.5% to 38.8%. The model overpredicts the negativity of incumbents by only 3.7 percentage points, but more significantly underpredicts the negativity of challengers, by 8.3%. Still, the model does broadly capture the significant differences between the two groups. Finally, while the model predicts candidates in open seat elections will campaign negatively about 51.1% of the time, in the data they do so about 65.2% of the time.

Table 46
Mean Negativity, Data vs. Simulated

	Data	Simulated
Overall	0.567	0.520
Democrats	0.586	0.536
Republicans	0.548	0.505
Incumbents	0.388	0.425
Challengers	0.715	0.632
Open Seats	0.652	0.511

Note: As % of Total Budget

Figure 18

Histograms of Negative Campaigning Proportions, True vs. Simulated

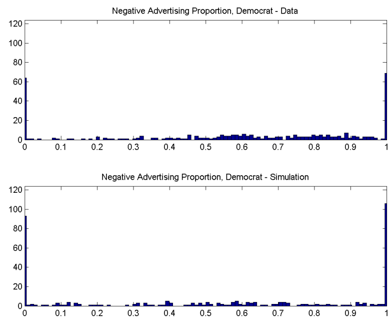


Figure 19

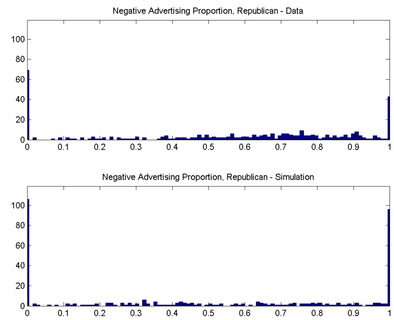


Figure 20

2.7 Results

Having shown the model can capture the salient features regarding campaign strategies, I move to analyzing the model's implications for the effectiveness of campaign strategies and spending. To interpret the remaining coefficients, I perform several exercises. As a first measure of the overall effectiveness of money, consider an open-seat election in which both candidates have the mean value of $\overline{r}_i^e$ and budgets B_i. In the sample, this implies (conditional on being skilled) values of $\overline{r}_1^e = 0.3144$, $\overline{r}_2^e = 0.3114$, $B_1 = \$2.393$ million, and $B_2 = \$2.338$ million. I then compute the change in expected vote share resulting from a 10% increase in one candidate's budgets, which corresponds to about $240,000 dollars. Note that for this exercise, I recompute the equilibrium under the new budgets. This exercise implies that, for both Democrats and Republicans, the increase in the expected vote differential in response to the increase is approximately 0.4 percentage points. This order of magnitude is consistent with Levitt [1994], which estimates that, in 1990 U.S. elections, a $100,000 increase in spending by a candidate increases his vote share by less than 0.2 percentage points for incumbents, and by between 0.19 and 0.42 percentage points for challengers. Converting $100,000 in 1990 to 2000 dollars, the model implies that such an increase in spending increases the expected vote differential by about 0.22 percentage points in an open seat elections. Unlike Levitt [1994], I find very little difference in ad effectiveness if I vary incumbency status, though at this point I hold both initial support and budgets fixed and approximately equal. Differences in marginal effectiveness across incumbents and challengers may be largely explained by systematic differences in initial support, average budget sizes, and diminishing returns of campaign spending, which I explore below.

To analyze the model in an alternative fashion, I note that Congressional districts based on the 2000 Census contain on average 521,759 voting age individuals.[67] Under the calibrated parameter, a budget increase of $1 per voting age individual (approximately $521,759) implies an increase in expected vote difference of about 0.87 percentage points in a representative election. This is in line, albeit a bit smaller, than estimates from Palda and Palda [1998], which find using French data that "incumbent candidates can at best expect to win 1.01% of the popular vote for each extra Franc they spend per registered voter in their district."[68]

In order to more richly characterize the implications of the model, I also investigate the overall effectiveness of spending for different combinations of initial support, incumbency, and budgets. Table 47 shows results when candidate 1 is an incumbent, while Table 48 shows results for an open seat election. The tables are constructed as follows. Fixing $\overline{r}_1^e + \widehat{\psi}_{inc}$ and $\overline{r}_2^e - \widehat{\psi}_{inc}$ (which vary by column), I compute the expected vote share if there was no spending, which is given by $\overline{r}_1^e - \overline{r}_2^e + 2\widehat{\psi}_{inc}$. Then, given B_1 and B_2 (which vary by row), I compute the equilibrium of the campaign game and the resulting expected vote difference. The numbers in the tables then reflect the pre-spending expected vote difference minus the post-spending expected vote difference – that is, a negative number indicates that, after spending, candidate 2 is relatively better off. I note that B_i low is selected to be approximately the 25th percentile of all budgets, B_i mid is approximately the median budget, and B_i high is the 75th percentile, while the initial supports are analogously defined. Note also that I keep things perfectly symmetric between the two sides, except for incumbency, to control for party-specific factors.

[67]This estimate comes from the 2008 ACS 3 year estimates of total population over 18 by Congressional district. Note that I do not have data on voter registration data by district, so this is an upper bound.

[68]Palda and Palda [1998] used data from 1993 French elections. Converting a 1993 French Franc to 2000 U.S. dollars implies that one French Franc from 1993 is worth about $0.22 in 2000 U.S. dollars. Under the calibrated model, an increase of $0.22 per voting age individual in spending increases the expected vote difference (in the representative election) by approximately 0.2 percentage points.

Table 47

CHANGE IN EXPECTED VOTE DIFFERENTIAL FOR CANDIDATE 1, FROM NO SPENDING TO EQUILIBRIUM WITH (B_1, B_2) - 1 IS INCUMBENT

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ high	$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid	$\overline{r}_1^e$ high/$\overline{r}_2^e$ low
B_1 low, B_2 high	-2.497	-2.555	-2.761
B_1 mid, B_2 mid	0.001	-0.008	-0.131
B_1 high, B_2 low	2.528	2.497	2.452

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326, plus/minus $\widehat{\psi}_{inc} = 0.0215$ for candidates 1/2. Results in percentage points.

Table 48

CHANGE IN EXPECTED VOTE DIFFERENTIAL FOR CANDIDATE 1, FROM NO SPENDING TO EQUILIBRIUM WITH (B_1, B_2) - OPEN SEAT ELECTION

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ high	$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid	$\overline{r}_1^e$ high/$\overline{r}_2^e$ low
B_1 low, B_2 high	-2.495	-2.501	-2.619
B_1 mid, B_2 mid	0.038	0	-0.038
B_1 high, B_2 low	2.619	2.501	2.495

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326. Results in percentage points.

Table 47 illustrates several important features of the model's implications for campaign spending effectiveness.[69] First, fixing a given row, note that the percentage change decreases as $\overline{r}_1^e$ increases and $\overline{r}_2^e$ decreases. This indicates that when candidate 1 is behind, his spending is relatively more productive, consistent with previous evidence (see Levitt [1994] and Palda and Palda [1998]). This is largely driven by the fact that, when one's opponent has a high level of initial support, negative campaigning is particularly effective. This is illustrated most clearly when B_1 mid, B_2 mid and $\overline{r}_1^e$ high, $\overline{r}_2^e$ low. Here, even though both candidates have identical budgets, candidate 2 benefits relatively more from his spending, albeit a minor amount of 0.131 percentage points. More generally, table also reflects the relative ineffectiveness of spending. For instance, even when the incumbent has a high budget and the challenger has a low budget – which corresponds to a $1.4 million advantage – in net, the incumbent's spending increases his expected margin by only 2.5%. The same dollar advantage is only slightly more effective for the challenger, yielding an increase of 2.8% for the challenger.

To investigate the relative effectiveness of positive versus negative campaigning, I consider how a large increase in either all positive or all negative campaigning affects the expected vote share.

[69]Table 48 shows the same figures, except for an open seat election, with similar implications.

In particular, for the same combinations of budgets and initial support as above, I compute the equilibrium. Then, I compute the expected vote difference due to a sizable increase ($237,000, or 10% of the average budget in the sample) in exclusively positive or exclusively negative campaigning for one candidate, without allowing for a response from the opponent.

Table 49

Change in Expected Vote Differential for Candidate 1, from $237,000 Increase in $B1$ - 1 is Incumbent

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ high		$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid		$\overline{r}_1^e$ high/$\overline{r}_2^e$ low	
	All Pos.	All Neg.	All Pos.	All Neg.	All Pos.	All Neg.
B_1 low, B_2 high	0.4432	0.4231	0.4442	0.3744	0.4445	0.3257
B_1 mid, B_2 mid	0.4200	0.4214	0.4174	0.3744	0.4166	0.3257
B_1 high, B_2 low	0.4094	0.4109	0.3919	0.3744	0.3904	0.3257

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326, plus/minus $\widehat{\psi}_{inc} = 0.0215$ for candidates 1/2. Results in percentage points

Table 50

Change in Expected Vote Differential for Candidate 1, from $237,000 Increase in $B1$ - Open Seat Election

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ high		$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid		$\overline{r}_1^e$ high/$\overline{r}_2^e$ low	
	All Pos.	All Neg.	All Pos.	All Neg.	All Pos.	All Neg.
B_1 low, B_2 high	0.4481	0.4496	0.4435	0.4042	0.4444	0.3555
B_1 mid, B_2 mid	0.4374	0.4390	0.4175	0.4042	0.4172	0.3555
B_1 high, B_2 low	0.4265	0.4281	0.3984	0.3998	0.3917	0.3555

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326.

Table 51

Change in Expected Vote Differential for Candidate 1, from $237,000 Increase in $B1$ - 1 is Incumbent, fixed $\overline{r}_2^e$

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ mid		$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid		$\overline{r}_1^e$ high/$\overline{r}_2^e$ mid	
	All Pos.	All Neg.	All Pos.	All Neg.	All Pos.	All Neg.
B_1 low, B_2 high	0.4803	0.3744	0.4442	0.3744	0.4070	0.3744
B_1 mid, B_2 mid	0.4524	0.3744	0.4174	0.3744	0.3808	0.3744
B_1 high, B_2 low	0.4247	0.3744	0.3919	0.3744	0.3667	0.3681

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326, plus/minus $\widehat{\psi}_{inc} = 0.0215$ for candidates 1/2. Results in percentage points.

Table 49 shows the results of this exercise. Note that Table 50 shows the same figures for an open seat elections, with similar implications. These results illustrate some important points. First, note that the overall vote share increases are relatively small, between 0.33 and 0.44 percentage *points*.

This reinforces the notion that campaign spending is relatively ineffective at increasing vote shares. Second, note that, for fixed budget levels, the effectiveness of negative campaigning decreases noticeably as $\overline{r}_2^e$ decreases, while the effectiveness of positive campaigning remains relatively constant.[70] This is largely due to the fact that, for Table 49, I simultaneously change the initial supports for both candidates so as to keep the measure of swing voters, $\overline{R}^e$, constant. I can also increase $\overline{r}_1^e$ while keeping $\overline{r}_2^e$ fixed, which necessarily decreases the measure of swing voters. In Table 51, I show results to illustrate this. In this case, as $\overline{r}_1^e$ increases, and thus $\overline{R}$ decreases, the effectiveness of positive campaigning decreases, while for negative campaigning it remains essentially constant. This illustrates that the level of $\overline{r}_1^e$ is not the key factor for the relative ad effectiveness for candidate 1, but rather the levels of $\overline{r}_2^e$ and $\overline{R}^e$. This is particularly true since, in the calibrated model, the value of the "boomerang" effect is minor, implying that the level of own initial support is not directly important for the optimal strategy.

Finally, holding fixed a level of initial support, as B_1 increases (and B_2 decreases), I see a decline in the relative effectiveness of additional campaign spending.[71] While it is more noticeable for positive campaigning (decreasing by about 0.05 percentage points from $B_1 = \$700,000$ to $B_1 = \$2.1$ million), it is still apparent in negative campaigning. This is due to diminishing returns from campaign spending that, while not strong, are present.

2.8 Conclusion

The effect of money on election outcomes is a widely discussed topic in economics and political science. A key factor that determines the effectiveness of money and its differential impact across candidates is campaign negativity, which is often overlooked by other studies. In particular, given that different candidate-types (e.g. incumbents versus challengers) use campaign funds in systematically different ways, recovering the true impact of money on election outcomes requires an understanding how effective alternative strategies are. To this end, I develop a structural model featuring a game between candidates who choose a level of negativity. Positive and negative campaigning affect different groups of voters in different ways: positivity is persuasive to swing voters deciding for whom to vote, whereas negativity affects polarized voters' decision of whether or not to turnout. Using

[70] I note that even for negative campaigning, the degree to which ad effectiveness changes as $\overline{r}_2^e$ changes is small, with changes of at most 0.1 percentage points.

[71] I also performed this exercise holding fixed B_2 as B_1 changes, and the result is essentially identical.

data on levels of negativity from television advertising, candidate budgets, and other candidate- and district-specific observables, I calibrate the model, which provides implications for the overall and relative effectiveness of campaign strategies.

The calibrated model suggests that campaign spending is mostly ineffective at increasing vote shares. For the average election, which has budgets of about $2.4 million, a 10% increase in one candidate's budget increases his expected vote differential by about 0.4 percentage points. This is roughly in line with results from Levitt [1994], among others. In alternative terms, in an election where both candidates have similar levels of initial support, if one candidate has a $2.1 million budget while the other $700,000, this yields a 2.5 percentage point improvement in the expected vote differential for the first candidate. I employ other calculations to find that, albeit small, the trailing candidates benefit from extra funds more than the leading ones. I also find that negative campaigning is relatively effective for candidates who face an opponent with a high level of initial support, while positive campaigning is relatively effective for candidates in elections where neither side has a particularly high initial support. Finally, the model implies slightly decreasing returns to spending. This may, in part, explain why the previous literature tends to find challenger spending is relatively more effective than incumbent spending, as incumbents typically have large budget advantages.

2.9 Appendix

2.9.1 Proofs

Proof of Proposition 2.

1. Suppose $MB_n^i(x_1,x_2) > 0$. Then definition 38 implies that $\frac{r_j\alpha_2}{r_i\alpha_1} > exp\{(\alpha_2-\alpha_1)(x_i-x_j)\}$. Now note that

$$\frac{\partial MB_n^i(x_j)}{\partial x_i} = r_j\alpha_2^2 exp\{-\alpha_1 x_j - \alpha_2 x_i\} - r_i\alpha_1^2 exp\{-\alpha_1 x_i - \alpha_2 x_j\}$$

which is negative if and only if $\frac{\alpha_2}{\alpha_1}\frac{r_j\alpha_2}{r_i\alpha_1} exp\{(\alpha_2-\alpha_1)(x_i-x_j)\} > \frac{r_j\alpha_2}{r_i\alpha_1} exp\{(\alpha_2-\alpha_1)(x_i-x_j)\} > 0$. But the first inequlity is satisfied due to the modeling assumption $\alpha_2 > \alpha_1$. The second inequality is obtained by the previous fact stated. Hence the statement is correct.

2. This is trivial since one can immediately see that $\frac{\partial MB_p^i(x_i)}{\partial x_i} > 0$ once the derivative is taken:

$$\begin{aligned}\frac{\partial MB_p^i(x_i)}{\partial x_i} &= \frac{2R}{\gamma K^4}\left[(1-1/\gamma)(1+B_i-x_i)^{1/\gamma-2}(1+B_j-x_j)^{1/\gamma}K^2\right.\\ &\quad\left.+2K/\gamma(1+B_i-x_i)^{2/\gamma-2}(1+B_j-x_j)^{1/\gamma}\right]\end{aligned}$$

where $K = (1+B_i-x_i)^{1/\gamma} + (1+B_j-x_j)^{1/\gamma}$.

3. Note that

$$\frac{\partial MB_n^i(x_j)}{\partial x_j} = r_i\alpha_1\alpha_2 exp\{-\alpha_1 x_j - \alpha_2 x_i\} - r_j\alpha_1\alpha_2 exp\{-\alpha_1 x_i - \alpha_2 x_j\}$$

Hence

$$\frac{\partial MB_n^i(x_j)}{\partial x_j}\begin{cases} <0 & \text{if } \frac{r_j}{r_i} > exp\{(\alpha_2-\alpha_1)(x_i-x_j)\}\\ =0 & \text{if } \frac{r_j}{r_i} = exp\{(\alpha_2-\alpha_1)(x_i-x_j)\}\\ >0 & \text{if } \frac{r_j}{r_i} < exp\{(\alpha_2-\alpha_1)(x_i-x_j)\}\end{cases} \tag{47}$$

The statement follows directly. To see it, suppose $\frac{\partial MB_n^i(x_j)}{\partial x_j} < 0$, that is $\frac{r_j}{r_i} > exp\{(\alpha_2-\alpha_1)(x_i-x_j)\}$. Taking the inverse of both sides immediately implies $\frac{r_i}{r_j} < exp\{(\alpha_2-\alpha_1)(x_j-x_i)\}$ which means $\frac{\partial MB_n^j(x_i)}{\partial x_i} > 0$. All other directions are similar.

4. Taking the appropriate derivatives, one can show that

$$sgn\left(\frac{\partial MB_p^i(x_1,x_2)}{\partial x_j}\right) = sgn\Big(-1/\gamma(1+B_j-x_j)^{1/\gamma-1}(1+B_i-x_i)^{1/\gamma-1}K^2 + 1/\gamma 2K(1+B_j-x_j)^{2/\gamma-1}(1+B_i-x_i)^{1/\gamma-1}\Big) = sgn\left(B_j-x_j-(B_i-x_i)\right)$$

where K is as defined above. The result follows immediately.

□

Proof of Lemma 2.4.1. Take $\tilde{x}_j \in [0, B_j]$. First, notice that if a corner $\{0, B_i\}$ is a best response, it is the unique one. To see this, note that $0 \in BR_i(\tilde{x}_j)$ if $MB_n^i(0,\tilde{x}_j) < MB_p^i(0,\tilde{x}_j)$. But since the marginal benefit of positive ads is increasing in x_1 and that of negative ads is decreasing in x_1, $MB_n^i(x_i,\tilde{x}_j) < MB_p^i(x_i,\tilde{x}_j)$ for all x_i, which implies 0 is the unique best response. The same idea in the opposite direction applies for B_i.

On the other hand, if $x_i \in (0, B_i)$ (an interior action) is in the best response, it is the unique one. To see this note that $MB_p^i(x_1,x_2) > 0 \ \ \forall\ x_k \in [0, B_k], \ \ k \in \{1,2\}$. Since for any interior best response it must be that $MB_n^i(x_i,\tilde{x}_j) = MB_p^i(x_i,\tilde{x}_j) > 0$. Recall that when $MB_n^i(x_1,x_2) > 0$, then $\left.\frac{\partial MB_n^i(x_j)}{\partial x_i}\right|_{(x_1,x_2)} < 0$. Also since $\frac{\partial MB_p^i}{\partial x_i} > 0$, the LHS is decreasing in x_i whereas the LHS is increasing. Hence there can be only one x_i that satisfies the condition $MB_n^i(x_i,\tilde{x_j}) = MB_p^i(x_i,\tilde{x_j})$.

Therefore, the best response is a function and is given by

$$BR_i(x_j) = \begin{cases} \tilde{x}_i & \text{if } MB_n^i(\tilde{x}_i,x_j) = MB_p^i(\tilde{x}_i,x_j) \\ 0 & \text{if } MB_n^i(0,x_j) \leq MB_p^i(0,x_j) \\ B_i & \text{if } MB_n^i(B_i,x_j) \geq MB_p^i(B_i,x_j) \end{cases} \tag{48}$$

Functions $MB_k^i(x_i,x_j)$, $k \in \{p,n\}$ are continuous in both x_i and x_j. Moreover, operations = and > preserve continuity. Hence, BR_i must be continuous. □

[**Proof of Theorem 1**] Define the function $f : [0, B_i] \to [0, B_i]$, $f(x) = BR_1(BR_2(x))$. Obviously, a strategy profile $(x_1^*, BR_2(x_1^*))$ is an equilibrium if and only if $f(x_1^*) = x_1^*$.

By Lemma 2.4.1, both BR_1 and BR_2 are continuous, which implies that f is also continuous. Since it also maps a compact set to itself, by Brouwer's fixed point theorem, there exists $x^* \in [0, B_1]$ such that $f(x^*) = x^*$. Hence, $(x^*, BR_2(x^*))$ is an equilibrium. This completes the proof.

Bibliography

Colombia's General System of Royalties: Opportunities and challenges behind the reform. Inner mimeo, DNP, July 2013.

A. I. Abramowitz. Explaining senate election outcomes. *American Political Science Review*, 82(2): 385–403, June 1988.

D. Acemoglu, C. Garca-Jimeno, and J. A. Robinson. State Capacity and Economic Development: A Network Approach. *American Economic Review*, 105(8):2364–2409, August 2015. URL https://ideas.repec.org/a/aea/aecrev/v105y2015i8p2364-2409.html.

J. Akin, P. Hutchinson, and K. Strumpf. Decentralisation and government provision of public goods: The public health sector in Uganda. *Journal of Development Studies*, 41(8):1417–1443, 2005. URL https://ideas.repec.org/a/taf/jdevst/v41y2005i8p1417-1443.html.

J. Albrecht, A. Anderson, and S. Vroman. Search by committee. *Journal of Economic Theory*, 145(4):1386–1407, July 2010. URL https://ideas.repec.org/a/eee/jetheo/v145y2010i4p1386-1407.html.

S. Ansolabehere and S. Iyengar. *Going Negative: How Political Advertising Divides and Shrinks the American Electorate.* New York: Free Press, 1997.

S. Ansolabehere, S. Iyengar, A. Simon, and N. Valentino. Does attack advertising demobilize the electorate. *American Political Science Review*, 88(4):829–838, 1994.

S. Ansolabehere, S. Iyengar, and A. Simon. Replicating experiments using aggregate and survey data: The case of negative advertising and turnout. *American Political Science Review*, 93(4): 901–909, 1999.

M. C. Araujo, F. H. Ferreira, P. Lanjouw, and B. Ozler. Local inequality and project choice : theory and evidence from Ecuador. Policy Research Working Paper Series 3997, The World Bank, Aug. 2006. URL https://ideas.repec.org/p/wbk/wbrwps/3997.html.

S. Ashworth and J. D. Clinton. Does advertising exposure affect turnout? *Quarterly Journal of Political Science*, 2(1):27–41, 2006.

P. Bardhan and D. Mookherjee. Pro-poor targeting and accountability of local governments in West Bengal. *Journal of Development Economics*, 79(2):303–327, April 2006. URL https://ideas.repec.org/a/eee/deveco/v79y2006i2p303-327.html.

P. Bardhan and D. Mookherjee. Political Clientelism and Capture: Theory and Evidence from West Bengal, India. Technical report, 2013.

T. Besley and S. Coate. Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of Public Economics*, 87(12):2611–2637, December 2003. URL https://ideas.repec.org/a/eee/pubeco/v87y2003i12p2611-2637.html.

T. Besley, R. Pande, and V. Rao. Just Rewards? Local Politics and Public Resource Allocation in South India. *World Bank Economic Review*, 26(2):191–216, 2012. URL https://ideas.repec.org/a/oup/wbecrv/v26y2012i2p191-216.html.

S. Bhattacharya. Campaign rhetoric and the hide-&-seek game, 2012. Working Paper.

R. M. Bird. Fiscal Decentralization in Colombia: A Work (Still) in Progress. International Center for Public Policy Working Paper Series, at AYSPS, GSU paper1223, International Center for Public Policy, Andrew Young School of Policy Studies, Georgia State University, Apr. 2012. URL https://ideas.repec.org/p/ays/ispwps/paper1223.html.

A. Bonica. Database on ideology, money in politics, and elections, 2013. URL http://data.stanford.edu/dime/. Public version 1.0 [Computer file].

A. Channa and J.-P. Faguet. Decentralization of Health and Education in Developing Countries: A Quality-Adjusted Review of the Empirical Literature. STICERD - Economic Organisation and Public Policy Discussion Papers Series 038, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE, July 2012. URL https://ideas.repec.org/p/cep/stieop/038.html.

O. Compte and P. Jehiel. Bargaining and Majority Rules: A Collective Search Perspective. *Journal* *of Political Economy*, 118(2):189–221, 04 2010. URL https://ideas.repec.org/a/ucp/jpolec/v118y2010i2p189-221.html.

B. S. da Silveira and J. M. de Mello. Campaign advertising and election outcomes: Quasi-natural experiment evidence from gubernatorial elections in brazil. *Review of Economic Studies*, 78(2):590–612, April 2011.

F. Ellis and N. Mdoe. Livelihoods and Rural Poverty Reduction in Tanzania. *World Development*, 31(8):1367–1384, August 2003. URL https://ideas.repec.org/a/eee/wdevel/v31y2003i8p1367-1384.html.

G. S. Eskeland and D. Filmer. Autonomy, participation, and learning in Argentine schools - findings andtheir implications for decentralization. Policy Research Working Paper Series 2766, The World Bank, Jan. 2002. URL https://ideas.repec.org/p/wbk/wbrwps/2766.html.

J.-P. Faguet and F. Snchez. Decentralization's Effects on Educational Outcomes in Bolivia and Colombia. *World Development*, 36(7):1294–1316, July 2008. URL https://ideas.repec.org/a/eee/wdevel/v36y2008i7p1294-1316.html.

J.-P. Faguet and F. Snchez. Decentralization and access to social services in Colombia. *Public Choice*, 160(1):227–249, July 2014. URL https://ideas.repec.org/a/kap/pubcho/v160y2014i1p227-249.html.

P. Feldman and J. Jondrow. Congressional elections and local federal spending. *American Journal of Political Science*, 28(1):147–164, February 1984.

S. E. Finkel and J. G. Geer. A spot check: Casting doubt on the demobilizing effect of attack advertising. *American Journal of Political Science*, 42(2):573–595, 1998.

A. Fiszbein. The Emergence of local capacity: Lessons from Colombia. *World Development*, 25(7):1029–1043, July 1997. URL https://ideas.repec.org/a/eee/wdevel/v25y1997i7p1029-1043.html.

A. Foster and M. Rosenzweig. Democratization, decentralization and the distribution of local public goods in a poor rural economy. *PIER Working Paper*, 2001. URL http://siteresources.worldbank.org/INTPUBSERV/Resources/paper_foster-rosenzweig.pdf.

E. F. Fowler and T. N. Ridout. Negative, angry, and ubiquitous: Political advertising in 2012. *The Forum*, 10(4), Feb 2013.

K. Fridkin and P. J. Kenney. Do negative messages work? the impact of negativity on citizens' evaluations of candidates. *American Politics Research*, 32(5):570–605, 2004.

K. L. Fridkin and P. J. Kenney. Variability in citizens' reactions to different types of negative campaigns. *American Journal of Political Science*, 55(2):307–325, 2011.

E. Galasso and M. Ravallion. Decentralized targeting of an antipoverty program. *Journal of Public Economics*, 89(4):705–727, April 2005. URL https://ideas.repec.org/a/eee/pubeco/v89y2005i4p705-727.html.

G. M. Garramone. Effects of negative political advertising: The roles of sponsor and rebuttal. *Journal of Broadcasting and Electronic Media*, 29(2):147–159, 1985.

K. Goldstein and J. Rivlin. Congressional and gubernatorial advertising, 2003-2004, 2007. Combined File [dataset]. Final release.

K. Goldstein and J. Rivlin. Congressional and gubernatorial advertising, 2007-2008, 2011. Combined File [dataset]. Final release.

K. Goldstein, M. Franz, and T. Ridout. Political advertising in 2000, 2002. Combined File [dataset]. Final release.

B. R. Gordon and W. R. Hartmann. Advertising competition in presidential elections, 2013a. Stanford University Graduate School of Business Research Paper No. 2131.

B. R. Gordon and W. R. Hartmann. Advertising effects in presidential elections. *Marketing Science*, 32(1):19–35, June 2013b.

K. Grier. Campaign spending and senate elections, 1978-84. *Public Choice*, 63(3):201–219, December 1989.

L. Hao and W. Li. Misinformation. *International Economic Review*, 54(1):253–277, 02 2013. URL http://ideas.repec.org/a/wly/iecrev/v54y2013i1p253-277.html.

J. J. Harrington and G. D. Hess. A spatial theory of positive and negative campaigning. *Games and Economic Behavior*, 17(2):209–229, December 1996. URL http://ideas.repec.org/a/eee/gamebe/v17y1996i2p209-229.html.

M. Karanicolas. Regulation of paid political advertising: A survey, 2012. Centre for Law and Democracy.

K. Kawai and T. Sunada. Campaign finance in u.s. house elections, September 2015. Manuscript.

M. P. Keane and K. I. Wolpin. The Career Decisions of Young Men. *Journal of Political Economy*, 105(3):473–522, June 1997. URL https://ideas.repec.org/a/ucp/jpolec/v105y1997i3p473-522.html.

Y. Krupnikov. When does negativity demobilize? tracing the conditional effect of negative campaigning on voter turnout. *American Journal of Political Science*, 55(4):796–812, 2011.

A. M. Larson. Natural Resources and Decentralization in Nicaragua: Are Local Governments Up to the Job? *World Development*, 30(1):17–31, January 2002. URL https://ideas.repec.org/a/eee/wdevel/v30y2002i1p17-31.html.

R. R. Lau, L. Sigelman, and I. B. Rovner. The effects of negative political campaigns: A meta-analytic reassessment. *The Journal of Politics*, 69(4):1176–1209, 2007.

J. Lentz. *Electing Jesse Ventura*. Lynne Rienner Pub, 2001.

S. D. Levitt. Using repeat challengers to estimate the effect of campaign spending on election outcomes in the u.s. house. *Journal of Political Economy*, 102(4):777–98, August 1994. URL http://ideas.repec.org/a/ucp/jpolec/v102y1994i4p777-98.html.

N. V. Loayza, J. Rigolini, and O. Calvo-Gonzlez. More Than You Can Handle: Decentralization and Spending Ability of Peruvian Municipalities. *Economics and Politics*, 26(1):56–78, 03 2014. URL https://ideas.repec.org/a/bla/ecopol/v26y2014i1p56-78.html.

G. Mansuri and V. Rao. *Localizing Development : Does Participation Work?* Number 11859 in World Bank Publications. The World Bank, 2013. URL https://ideas.repec.org/b/wbk/wbpubs/11859.html.

P. S. Maro. The impact of decentralization on spatial equity and rural development in Tanzania. *World Development*, 18(5):673–693, May 1990. URL https://ideas.repec.org/a/eee/wdevel/v18y1990i5p673-693.html.

A. Merlo and A. de Paula. Identification and Estimation of Preference Distributions When Voters Are Ideological. PIER Working Paper Archive 11-001, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania, Dec. 2010. URL https://ideas.repec.org/p/pen/papers/11-001.html.

W. Moon. The paradox of less effective incumbent spending: Theory and tests. *British Journal of Political Science*, 36(4):705–721, October 2006.

R. Myerson. Local Agency Costs of Political Centralization. Working paper series, University of Chicago, April 2015. URL http://home.uchicago.edu/~rmyerson/research/localagency.pdf.

B. Nalebuff and R. Shachar. Follow the leader: Theory and evidence on political participation. *American Economic Review*, 89(3):525–547, June 1999. URL http://ideas.repec.org/a/aea/aecrev/v89y1999i3p525-547.html.

W. E. Oates. *Fiscal federalism*. Harcourt Brace Jovanovich, New York, 1972.

F. Palda and K. Palda. The impact of campaign expenditures on political competition in the french legislative elections of 1993. *Public Choice*, 94(1-2):157–74, January 1998. URL http://ideas.repec.org/a/kap/pubcho/v94y1998i1-2p157-74.html.

N. Persico and A. Lizzeri. The Provision of Public Goods under Alternative Electoral Incentives. *American Economic Review*, 91(1):225–239, March 2001. URL https://ideas.repec.org/a/aea/aecrev/v91y2001i1p225-239.html.

L. Ragsdale and T. E. Cook. Representatives' actions and challengers' reactions: Limits to candidate connections in the house. *American Journal of Political Science*, 31(1):45–81, February 1987.

J. M. Ramirez, Y. Diaz, and J. G. Bedoya. Decentralization in Colombia: Searching for social equity in a bumpy economic geography. Working Papers 337, ECINEQ, Society for the Study of Economic Inequality, Aug. 2014. URL https://ideas.repec.org/p/inq/inqwps/ecineq2014-337.html.

R. Shachar. The political participation puzzle and marketing. *Journal of Marketing Research*, 46 (6):798–815, December 2009.

S. Skaperdas and B. Grofman. Modeling negative campaigning. *American Political Science Review*, 89(1):49–61, 1995.

SRDS. *The Lifestyle Market Analyst*. Wilmette, IL : Standard Rate and Data Service, 2000, 2004, 2008.

T. Stratmann. Some talk: Money in politics. a (partial) review of the literature. *Public Choice*, 124

(1):135–156, July 2005. URL http://ideas.repec.org/a/kap/pubcho/v124y2005i1p135-156.html.

T. Stratmann. How prices matter in politics: the returns to campaign advertising. *Public Choice*, 140(3):357–377, September 2009. URL http://ideas.repec.org/a/kap/pubcho/v140y2009i3p357-377.html.

D. Stromberg. How the electoral college influences campaigns and policy: The probability of being florida. *American Economic Review*, 98(3):769–807, 2008.

M. Tommasi and F. Weinschelbaum. Centralization vs. Decentralization: A Principal-Agent Analysis. *Journal of Public Economic Theory*, 9(2):369–389, 04 2007. URL https://ideas.repec.org/a/bla/jpbect/v9y2007i2p369-389.html.

C. f. P. S. University of Michigan. American national elections time series study, 2000. American National Elections Study.

C. f. P. S. University of Michigan. American national elections time series study, 2004. URL http://www.electionstudies.org/. American National Elections Study.

C. f. P. S. University of Michigan. American national elections time series study, 2008. URL http://www.electionstudies.org/. American National Elections Study.

U.S. Census Bureau. Census 2000. American FactFinder, U.S. Census Bureau, 2002. URL <http://factfinder.census.gov/>.