Università Commerciale "Luigi Bocconi" Milano

"XVII" Cycle

Ph.D. Dissertation in Economics

"Essays in Political Economics"

P.h.D. Dissertation by :

"Leandro Machado de Magalhaes"

Matr. n°: "18226"

Anno Accademico 2005-2006

To Lia and Francisco

ACKNOWLEDGEMENTS

I am grateful to Guido Tabellini for his support and guidance; Antonio Merlo, Carlo Favero, and Petra Todd for their invaluable advice. I would specially like to acknowledge the contribution of Lucas Ferrero to the completion of this work.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis combines three essays in Political Economics. In the first Lucas Ferrero and I study how institutions and politics affect economic variables. We study the effect of veto power and political alignment between the executive and the legislative on the overall tax level of the American states. We define budgetary separation of powers and present a theoretical model of budgetary bargaining in the American States. Specifically, if the state allows the governor to *line item veto* the budget and if there is a *divided government* in place, there is budgetary separation of powers and taxes are lower. We use regression discontinuity design to establish a causal relation between a divided government and lower tax rates in states with line item veto. The discontinuity arises at the point in which the governor gains control of both chambers. We estimate the jump at the discontinuity semiparametrically and compare the results to a fixed effects estimation.

In Chapter 2 we allow for a dynamic model in which the tax level presents persistence over time. We account for the endogeneity of such model and show that our empirical results in Chapter 1 are robust to this modification. We go further and analyze other implications of our theoretical model such as the interaction between another institutional feature, supermajority requirements for a tax increase, and budgetary separation of powers; how line item veto should have a disproportionate impact on expenditures that may be easily target to specific constituencies; and, finally, how changes in political preferences affect budgetary separation of powers.

In my third and last Chapter I present an analysis of one of the most important puzzles for political economics: individual decisions to participate in the political process or to abstain. I bypass the analysis of pivotal-voter models that have little significance in mass elections and look back at the spatial model and the role played by information. I combine the empirical spatial analysis of Poole and Rosenthal (1984) and the uncertain voter model of Degan and Merlo (2006). The cost of voting is modelled as the cost of voting for the candidate that may turnout to be the farthest from the voter. Only uninformed individual may abstain, informed ones vote. I show indicative results that statistics identifying the political position of individuals and their level of information go along way in helping predict individual participation behavior.

The chapters are organized as independent self-contained papers. The Appendix and the Reference section are common to all three.

Chapter 2

Budgetary Separation of Powers in the American States: A Regression Discontinuity Analysis

2.1 Introduction

A presidential regime is characterized by the separation of powers: the presence of an independently elected executive which does not depend on a vote of confidence by the parliament[1]. Persson et al. (2000), have linked the concept of separation of powers with the size of government. They predict a lower tax level in presidential regimes than in parliamentary regimes.

[1]These features are shared in the definitions by Lijphart (1999) in 'Patterns of Democracy' and Shugart and Carey (1992) in 'Presidents and Assemblies'. Lijphart also requires a one person executive, and Shugart and Carey include in the definition some law making power to the executive.

In this paper we argue that the concept of separation of powers must be narrowed down when studying the institutional mechanisms that influence the size of government. For this purpose we define the concept of *budgetary separation of powers*. It is present in a regime when the political group controlling the tax level is not the residual claimant of a tax increase, that is, the extra resources from the tax increase can not be appropriated by that group. This feature may or may not be present in parliamentary or presidential regimes [2]. We study the case of the American states. These are defined as presidential regimes: the executive power is separate and independently elected. Yet, budgetary separation of powers will only be present when the government is divided and the governor has the line item veto. Our empirical results support this theoretical prediction.

In the American states, line item veto is mostly seen as a tool to cut down the pork and trim the budget. It allows governors to cut specific appropriation items, language, or trim values down. Most states have had this feature since the end of the 19^{th} century. Today 45 states have this institutional feature[3]. At the Federal level its adoption has been controversial. Many Presidents urged Congress to give

[2] In Latin America presidential regimes, for example, most executives may initiate tax increasing bills, write the budget, have decree power, and even have veto rights with amendment powers. Budgetary separation of power, as we define here, will hardly be present. In an empirical study by Persson and Tabellini looking for the effects of presidentialism on the tax level, 'The Economic Effects of Constitutions'(2003)Persson and Tabellini (2003) table 6.2, the IV result on a panel of countries depended on the exclusion of a Latin America dummy. In the OLS Latin America is the only continental dummy that is statistically significant

[3] Only two states adopt the line item veto within our data set, at the very end of the sample. For all purposes we consider line item veto time invariant. For a study on the reasons for its adoption see de Figueiredo Jr. (2002)

this power to them. During the Reagan and Bush years, a Democrat controlled Congress refused to yield. When Republicans became the majority under Clinton, they approved it only to see it judged unconstitutional by the Supreme Court in a 6-3 decision[4]. To this day the President can only block veto the budget proposed by Congress, which would imply a government shut down. Government employees would stay home and government provided services stop except for limited essential areas.

Many other institutional features aimed at limiting the size of the budget and the tax rate have been adopted across states. In the seventies, tax and expenditure limitations were introduced by many. Recently, supermajority requirements to raise taxes have been adopted as well[5]. Moreover, all states except Vermont have some form of balanced budget requirement and no-carry-over deficit rules. All these rules have been adopted with the objective of improving fiscal perfomance and keeping taxes under control.

A large empirical and theoretical literature has studied these institutions and their effects on state's finances, theoretically and empirically. Bohn and Inman (1996) work with a panel on 47 states from 1970 to 1991. Since line item veto is time invariant, they regress the fixed effects on the institutional features. They find that states with line item veto and no-deficit rules have lower deficits. Alt and Lowry (1994) and Poterba (1994) are interested in how governments respond to recession

[4]For a more detailed account of the Supreme Court ruling see Urofsky and Finkelman (2002)

[5]Knight (2000a) has found a significant negative effect of supermajority requirements on the tax level controlling for the endogeneity arising from self selection.

driven deficits. Their findings are that unified governments tend to respond faster, specially if they have strong no-carry-over rules; and, more interesting for our own work, the adjustment under unified governments is relatively more dependent on tax increases. For a comprehensive review see Besley and Case (2003).

On whether the line item veto affects the tax rate, we start by mentioning two works based on cross section estimates. Abrams and Dougan (1986) find no effect of line item veto on the tax level. Alm and Evers (1991) find a a positive relationship between the veto itself and the tax level, and a negative relation between the tax level and an interaction between line item veto and an indicator for divided government.

Closer to our work, Holtz-Eakin (1988) studies a panel from 1966 to 1983. He runs a fixed effect model interacting the time invariant line item veto with partisan variables that indicate different levels of control of state institutions. He finds a negative impact on spending but a *positive* impact on the overall taxation. This is not seen as unexpected by Holtz-Eakin (1988). He had no prior on the direction the line item veto would affect the tax level. In his model the governor represents the preferences of the median voter in the state and the chambers represent the preferences of the median legislator. Line item veto brings the outcome closer to the governor's preferred point. Since the governor's preferred point is unknown, the direction of the line item veto effect on tax and expenditure is not predicted.

The most recent empirical work to our knowledge on the effects of line item veto is Besley and Case (2003). They present no model but argue that the line item veto should improve the bargaining power of the governor. They have a longer data set and interact line item veto with a dummy for divided government. In their estimates a divided government in a state with line item veto has a negative effect on the tax level.

To infer causality, the variable of interest, a dummy for *divided government* interacted with *line item veto*, must be considered as a treatment that is assigned randomly across all state-years. This is not the case since line item veto is mostly time invariant and a divided government is the result of elections. To infer causality we must account for potential endogeneity. Omitted variables are of particular concern, such as changes in political climate or of preferences over the tax level across states and over time. Another issue is serially correlated outcomes, which are common in the diffs-in-diffs literature, and may result in downward biased standards errors[6].

We use a panel of 47 states across 38 years. Our left hand side variables is the average tax rate over potential GDP, *ttax_gdpp*. Our variable of interest is the interaction between line item veto and a dummy for divided government. First we present full sample estimates with state and year fixed effects. Our results are similar to those in Besley and Case (2003). We find a significant negative correlation between the tax level and a divided government in a state with line item veto. We then add variables that are proxies for omitted variables such as idiosyncratic political preferences: state level turnout and election results for lower political offices.

[6] Bertrand et al. (2004) study this problems with simulations in a diff-in-diffs context.

Finally, we present our regression discontinuity design estimates, which deal with omitted variables by comparing state-years around the discontinuity. For this purpose we define the variable *gov_strength*. We assume that the legislative and the executive are aligned only when *both* chambers are controlled by the same party as the governor's. Therefore, we define *gov_strength* as the seat share of the governor's party in the chamber where this share is the least. The function describing the relation between the average tax rate and *gov_strength* is assumed to be continuous except for a discontinuity at $gov\_strength = 0.5$. Above it both chambers are aligned with the governor; below it at least one chamber is controlled by the opposition.

The semiparametric approach allow us to estimate the discontinuity without having to assume a particular functional form for the relation between *ttax_gdp* and *gov_strength*; the shape of the function is retrieved nonparametrically. Our results imply a jump at the discontinuity of 0.3 in the average state tax level, which is 5% of GDP. Moving from a unified to a divided government in a state with line item veto decreases the average tax level from 5 to 4.7% of GDP.

Before going through the details of our estimation strategy in Section 3 we develop a model in the next section to make clear our prediction that a divided government brings taxes down only in states with line item veto. Our model is a variation of the separation of powers model in Persson et al. (2000) that accommodates institutional features of the American states[7]. It delivers a clear prediction for the tax level and

[7]This is a familiar model of conflicting transfer provisions in different politico-institutional settings. It has been used in different applications, for example, Grossman and Helpman (2005).

transfers in equilibrium, which depend on the institutional and political setting of that state and year. We show that line item veto works in keeping taxes low because it allows a minority governor to prevent the majority controlled legislature, that has agenda setting powers over both taxation and allocation, from being the residual claimant of a tax increase.

2.2 Budgetary Separation of Powers in the States

In the American states, by either constitutional or statutory requirements, the power to initiate tax increasing bills and to approve the budget lies with the legislature. Even if the budget is written by the governor or by independent agencies, it can be amended and rewritten at will once it reaches the House and Senate[8]. This leaves the legislature with all effective agenda setting power. They propose a tax rate and how to allocate revenues [9]. We focus our discussion on two institutional players, the legislative L and the executive E, and on two constituencies, the democrats D and republicans R.

2.2.1 Setup

A state is made of two groups of voters, Republicans, R; and Democrats, D. Each group is composed by a continuum of voters of mass one. Individuals in either group

[8]For detail information on states budget procedures see NCSL (2005).

[9]We will abstract away from deficits and veto overrides in the following model. Most States have stringent no-deficit-carry-over rules and override requirements are usually two-thirds.

are identical in every aspect except for their preferences for direct transfers, f, they receive from the government. Republicans can only derive utility from f^R transfers; Democrats only from f^D.

There are two offices in each state, the legislative L and the executive E. Each group appoints one politician to run for each office in every period. Since the groups are of the same size, election results are decided by the flip of a coin. The randomness of elections is what we try to reproduce in the empirical part with the regression discontinuity design by focusing on close parliamentary elections[10].

The role of the legislative is to make a budget proposal, which consists of a lump sum tax rate, τ; an amount for each transfer, f^D and f^R; and rents, r^L and r^E, for the politician in the legislative and executive offices.

The role of the executive is to veto the proposal. If the executive only has the power to block veto the proposal, a status quo is triggered: $f^E = f^L = 0$ and $r^E = r^L = \overline{r}$, exogenously given. Line item veto implies the executive may cut down f^R, f^D, r^L, and r^E separately; or trigger the status quo as well. The resources from the cuts go to lower taxes. The block veto remains an option also in the states with line item veto; it may be used if $r < \overline{r}$. We make the following parametric assumption: $1 - \delta < \overline{r} < \frac{2}{3}$.

Politicians and individuals are infinitely lived. The intertemporal utility of a

[10]The assumption of two groups is not essential. Similar results for the tax rate could be generated with three districts, each with a representative, and an executive elected by all. We would, however, have an undetermined transfer allocation. What we need in both models, is that two districts be partisan and total party alignment between chambers and governor (at least around the discontinuity).

voter in group i at period s is given by:

$$u_s^i = \mathbf{E}_s \left[\sum_{t=s}^{\infty} \delta^{t-s} w^i(q_t) \right],$$

where δ is a discount factor and q_t is a vector of policies $q_t = [\tau_t, f_t^D, f_t^R, r_t^L, r_t^E]$. The utility function in each period for a given voter in group i is given by:

$$w^i = c^i = y - \tau + f^i.$$

Voters want as much transfer and as little taxes as possible. All policy variables are constrained to be nonnegative. Individual income y is normalized to 1.

Politicians want to appropriate rents, r. Each politician $l = L, E$ maximizes her own rents:

$$W_s^l = \sum_{t=s}^{\infty} \delta^{t-s} r_t^l D_t^l,$$

where D_t^l is one if in office in period t and zero otherwise.

General expenditures cannot be financed with deficits, only through revenues derived from a state wide lump sum tax. The resources are used to pay for politicians rents and group specific transfers. When choosing policy, politicians face the following government budget constraint:

$$2\tau = f^R + f^D + r^L + r^D,$$

the total amount of taxes is 2τ since each group has size 1.

As a benchmark, consider what a benevolent central planner would do. She would maximize the sum of voters utilities by setting rents to zero and share transfers equally. If taxation were somewhat distortionary, transfers would be set to zero.

Taxes would be just high enough to pay for the transfers.

In the following, policy choices are delegated to politicians. This implies three sources of conflict: between different types of voters, between voters and their representatives, and among the politicians themselves. We first discuss the case of one party controlling both offices; neither of the veto types plays a role. We move to the case of divided governments. In the states with block veto budgetary separation of power is not present even if the government is divided. Taxes are maximum. In a state with line item veto, a minority governor prevents the agenda setter from being the residual claimant of a tax increase. When this is the case budgetary separation of powers is present and the tax level will constrained.

2.2.2 Timing

Voters hold the incumbents that belong to their group accountable with the following backward looking strategy: I vote for the incumbent candidate if my utility is above a certain threshold ω_i; otherwise I vote for *another* politician of my own group. This rule is used for both the executive and the legislator. The voters who do not belong to the same group as the incumbent vote for a candidate of their own group. In equilibrium no politician is punished and the same politicians randomly alternate in power.

The legislative games starts with two incumbents. The timing of the game is as follow:

1. *Nature* decides the outcome of the elections for the legislature L and the executive E.

2. Voters of both types set their reservation utilities, ω^i, simultaneously and taking into account the subsequent stages of the game.

3. L makes a proposal for the allocation of resources and for the tax level: q_L.

4. The executive may veto the budget. The cuts go towards lower taxes. Depending on the state either block veto or line item veto is available.

5. Appointments are made and elections are held.

We look for sequential equilibria. We define here equilibrium in the block veto case and leave to the appendix the definition of equilibrium in the line item veto case:

1. for any given vector of reservation utilities at period t: $w_t = (w_t^R, w_t^D)$; at the veto stage, the executive prefers $q_t^B(w_t)$ to the status quo outcome;

2. for any given w_t, the legislator L prefers $q_t^B(w_t)$ to any other policy satisfying the condition above;

3. the reservation utilities w_t^{iB} are optimal for the voters of each type i, when one takes into account that policies in the current period are set according to

$q_t^B(w_t)$; and takes as given the reservation utilities of the individuals of type $j \neq i$, the identity of the legislator, and of the executive.

2.2.3 One-party rule

In the case of a one-party rule, the block veto and line item veto cases are identical. Both deliver a tax rate that is maximum. For taxes to be lower a divided government must be in place and the governor must be able to line item veto the budget, as we shall see in the next sections.

We have a one-party rule government when both the legislator L and the executive E belong to the same group, call it group i. Call the group of voters whose politicians are out of both offices j. The veto in this case, be it block or line item veto, only matters for how rents are divided among politicians. The voters of the politicians in office set their reservation utilities and both politicians are held accountable to the voters of group i.

PROPOSITION 1 In a one-party rule government of group i, there is a unique stationary equilibrium that satisfies the following conditions:

$$\tau^* = 1;$$

$$r^L = 2 - \overline{r} - \delta\frac{2-\overline{r}}{2-\delta}; \quad r^E = \overline{r} - \delta\frac{\overline{r}}{2-\delta};$$

$$f^{i*} = \delta\frac{2-\overline{r}}{2-\delta} + \delta\frac{\overline{r}}{2-\delta}; \quad f^{j*} = 0;$$

$$\omega^i = f^{i*}; \quad \omega^j = 0;$$

and all politicians are reappointed to run in the next election.

Proof of Proposition 1.

The first step is to determine the outside option for politicians. Suppose politicians decide to forego their political career. The optimal strategy for the agenda setter is to set the tax level to the maximum, $\tau = 1$, and buy off the executive not to have her proposal vetoed. In this case L's payoff is: $2 - \overline{r}$, where $\overline{r}$ is the minimum the executive accepts; any less and the veto would be used.

In equilibrium voters must make politicians at least indifferent between running away with everything and continuing their political careers. They subtract from the rents above the discounted continuation value of being a politician; making politicians indifferent between running away and delivering the transfers. Call W^E the continuation value of being a politician running for the executive and W^L for the legislative. Voters allow enough resources for L to appropriate $2 - \overline{r} - \delta W^L$ and for E to appropriate $\overline{r} - \delta W^E$. Summing up we have the total rents in equilibrium: $r^* = 2 - \delta W^L - \delta W^E$.

LEMMA 1. There are zero transfers for the group of voters, j, whose politicians are not in office.

Proof. Suppose there is an equilibrium in which the voters of group j are receiving positive transfers: $f^j > 0$. It is optimal for voters of group i to deviate and set their reservation utilities such that any resources are shifted away from f^j to f^i . Politicians comply and are reelected. QED.

Voters of group i take the rents into their budget constraint and maximize:

$$Max_{fi,\tau}\ w^i = 1 - \tau + f^i$$

$$s.t.\ f^i + 2 - \delta W^E - \delta W^L \leq 2\tau,$$

which yields:

$$\tau^* = 1,$$

$$f^{i*} = 2 - r^L + r^E = \delta W^L + \delta W^E.$$

Finally, to retrieve the results in Proposition 1 we define W^E and W^L. The continuation value of being the legislator is given by the probability of being elected to office each period, $\frac{1}{2}$, and receiving r^L:

$$W^L = \frac{1}{2}(2 - \overline{r} - \delta W^L) + \delta W^L = \frac{2 - \overline{r}}{2 - \delta}.$$

The continuation value of being the presidential candidate is given by:

$$W^E = \frac{1}{2}(\overline{r} - \delta W^E) + \delta W^E = \frac{\overline{r}}{2 - \delta}.$$

QED.

In the case of one-party rule all the decision power is in the hands of group i. The optimal taxes are maximum because voters of group i only incur half the cost of taxation but receive all its marginal benefits in the form of direct transfers f^i. Group i is the residual claimant of the tax increase and controls the agenda. There is no incentive to veto when powers are aligned.

Let's move on to the case of a divided government. As we shall see, block veto does not prevent the legislative voters from being the residual claimants of a tax increase and taxes are maximum. Only with line item veto shall taxes be restrained.

2.2.4 Divided Government and Block Veto

We have a divided government when at least one of the chambers is controlled by the party opposed to the governor's. Each politician is held accountable to one of the two groups of voters, R and D. To simplify the exposition we shall hereafter identify the each group with the position their representative holds, L or E.

PROPOSITION 2. In a state in which the executive has block veto power and the government is divided, there is a unique stationary equilibrium satisfies the following conditions:

$$\tau^* = 1;$$

$$r^{L*} = 2 - \overline{r} - \delta\frac{2-\overline{r}}{2-\delta}; \quad r^{E*} = \overline{r} - \delta\frac{\overline{r}}{2-\delta};$$

$$f^{L*} = \delta\frac{2-\overline{r}}{2-\delta}; \quad f^{E*} = \delta\frac{\overline{r}}{2-\delta};$$

$$\omega^L = f^{L*}; \quad \omega^E = f^{E*};$$

and all politicians are reappointed to run in the next election.

Proof.

The first step is to note that the equilibrium rents are the same as in the one-party rule case: $r^* = r^{L*} + r^{E*} = 2 - \delta W^E + \delta W^L$.

Here we need to consider the symmetric maximization problem of both groups. Let's look at legislative voters L. They maximize their utility taking as given the transfers to the other group:

$$Max_{f^L, \tau}\ w^i = 1 - \tau + f^L$$

$$s.t.\ f^L + f^E + 2 - \delta W^E - \delta W^L \leq 2\tau,$$

which yields:

$$\tau^* = 1,$$

$$f^{L*} = 2 - r^* - f^E.$$

By symmetry of the problem, we have:

$$f^{E*} = 2 - r^* - f^L.$$

Both groups wish to maximize their own transfers taking into account the transfers the other group is asking for. In equilibrium, however, the transfers to the executive voters are restrained.

LEMMA 2. $f^{E*} = \delta W^E$.

Proof.

Suppose there is an equilibrium with $f^E > \delta W^E$. It is optimal for the legislator not to deliver f^E, the cost is greater than the gain from paying a lesser bribe to the executive: $\overline{r} - \delta W^E$ instead of $\overline{r}$. The executive voters optimal deviation is to ask for $f^E = \delta W^{E}$[11].

Suppose there is an equilibrium with $f^E < \delta W^E$. The optimal deviation for the executive voters is to ask for $f'^E = \delta W^E - \epsilon$. This is true for whatever reservation utility voters in group L have asked for, and independently on whether the legislator will be reappointed or not. The legislators chooses the cheapest of the following alternative: delivering f'^E and paying the politician $\overline{r} - \delta W^E$ or paying $\overline{r}$. The deviation is such that the first is preferred. QED.

From the maximization above and Lemma 2 we have that $f^{L*} = \delta W^L$. Substituting $\delta W^E = \delta \frac{\overline{r}}{2-\delta}$ and $\delta W^L = \delta \frac{2-\overline{r}}{2-\delta}$ we have the results in Proposition 1 **QED**.

The main intuition of the proof is that the executive can always be bought. The budget proposed be the legislature is a take-it-or-leave-it offer. The outside option for the executive is either $\overline{r} - \delta W^E$, when she is being reelected; or $\overline{r}$, when she is being ousted. The executive voters use this difference to demand positive transfers

[11]The assumption that $\delta + \overline{r} \geq 1$ guarantees that the status quo outcome for the executive voters, $1 - 2\overline{r}$, is not preferred to $f^E = \delta W^E = \delta \frac{\overline{r}}{1-\delta}$.

in equilibrium. Lemma 2 is key to the uniqueness of the result in this and the next section. Taxes are maximal because the legislative voters control the agenda and are the residual claimants of any tax increase once r^*, and f^{E*} have been provided for. Any extra dollar goes to f^L. This will no longer be true, however, when line item veto is available to the executive.

2.2.5 Divided Government and Line Item Veto

Now the voters that control the executive have a credible threat in order to keep taxes low. They may ask for excess transfers to the other group to be trimmed. In doing so they reduce taxes and improve their lot. This is possible because budgetary separation of powers is present. The legislative and its voters no longer are the residual claimants of a tax increase.

PROPOSITION 3. In a state in which the executive has line item veto power and the government is divided, the unique stationary equilibrium satisfies the following conditions:

$$\tau^* = 1 - \frac{\delta}{2}\frac{2-\overline{r}}{2-\delta} + \delta\frac{\overline{r}}{2-\delta} < 1;$$

$$r^{L*} = 2 - \overline{r} - \delta\frac{2-\overline{r}}{2-\delta}; \quad r^{E*} = \overline{r} - \delta\frac{\overline{r}}{2-\delta};$$

$$f^{L*} = 2\delta\frac{\overline{r}}{2-\delta}; \quad f^{E*} = \delta\frac{\overline{r}}{2-\delta};$$

$$\omega^L = 1 - \tau^* + f^{L*}; \quad \omega^E = 1 - \tau^* + f^{E*};$$

and all politicians are reappointed to run in the next election[12].

[12] The assumption that $\overline{r} < \frac{2}{3}$, guarantees that $\tau^* < 1$

Proof.

Rents in equilibrium are the same as in the above sections. In the case the legislator decides to deviate, she sets taxes at maximum and buys off the executive with $\overline{r}$. Voters discount the continuation value of being in office and include the rents in their budget constraint. Rents are not cut below the level that makes politicians indifferent between delivering the transfers and foregoing their careers. At the veto stage the executive takes as given whatever value was assigned to her by the legislator. At that stage it is possible to cut down f^L or r^L in order to reach the executive voters reservation utilities through the correspondent tax decrease. This, of course, is not possible when only block veto is available.

LEMMA 3. $f^{L*} \leq 2\delta W^E$ and $f^{E*} = \delta W^E$.

Proof.

Suppose we are in an equilibrium as in the block veto case with $f^E = \delta W^E$, $f^W = \delta W^L$ and $\tau = 1$. Is there an optimal deviation for the executive voters? Yes, to set their reservation utilities at $\omega^E = \frac{\delta W^L}{2}$[13]. If the legislator tries to deliver f^L it will be cut, the cut is enough to reach the reservation utility of the executive voters. The same is true if the legislator tries to appropriate f^L as extra rent. The legislator is sure to loose reappointment. Taxes are set to maximum, $\tau = 1$,

[13] This is optimal if $\frac{\delta W^L}{2} > \delta W^E$; which is true by assumption since $\overline{r} < \frac{2}{3}$.

$f^E = \frac{\delta W^L}{2}$, and the rest goes to extra rents to the legislator. The amount requested by the executive voters is limited, any higher and the executive can not guarantee reelection by cutting rents or transfers. This deviation is optimal for the executive voters whenever $\frac{f^L}{2} \geq \delta W^E$.

Note that the results of Lemma 2 are also valid here and $f^{E*} = \delta W^E$.

QED.

The legislative voters maximize their utility as before:

$$Max_{\tau}\ w^L = 1 - \tau + f^L$$

$$\begin{aligned} s.t. \qquad & f^L + f^E + 2 - \delta W^E - \delta W^L \leq 2\tau, \\ & f^L \leq 2\delta W^E, \\ & f^{E*} = \delta W^E, \end{aligned}$$

which yields:

$$f^{L*} = 2\delta W^E,$$

and

$$2\tau^* = r^E + r^L + f^E + f^{L*},$$

that is,

$$\tau^* = 1 - \frac{\delta W^L}{2} + \delta W^{E*}.$$

Substituting the values for W^E and W^L we have the results in Proposition 3.

QED.

The line item veto allows executive voters to prevent the legislative voters from being the residual claimants of a tax increase once $\tau > \tau^*$. It allows for optimal deviations by the executive voters in out-of-equilibrium paths which, when only block veto is available, are not feasible. The main intuition is that when the line item veto is available, the budget is not a take-it-or-leave-it offer. Particularly, excessive transfers to the legislative constituency may be vetoed, bringing taxes down.

The model makes clear the mechanism through which budgetary separation of powers works. The addition of a common public good, a third district, as in Persson et al. (2000), or an executive elected by all districts would complicate the characterization of the equilibria but the main intuition would remain. When the state has the line item veto, the budget is no longer a take-it-or-leave-it offer to the executive, it can now cut down transfers to competing political groups. The key political assumption is alignment of interests between the governor and her party representatives. Without some degree of party alignment, the conflict over the tax rate would be between the governor and the legislature, and the tax rate should not be influenced by variations between divided and unified governments. But as we find in our empirical exercise, it is.

Another important assumption in the model is the randomness of elections. It buys us an equilibrium in which a divided and an unified government alternate. The regression discontinuity strategy comes the closest to this setup. It attempts to recreate the randomness of elections by focusing government that are divided or unified by a small margin.

2.3 Empirical Analysis

For there to be budgetary separation of powers in the American states we need an institutional feature: line item veto; and a political feature: divided government. We expect a divided government to have a negative effect on the tax level in the states with line item veto and no effect in the states with block veto. We bring this prediction to the data.

First we present a strategy that takes cares of all possible endogeneity that are time invariant or state specific. We find a negative partial correlation between *divided* government and the tax level in the states with line item veto. We have to be cautious not to draw conclusions of causality from these results. Omitted variables that vary across states and years such as preferences over tax rates and political mood remain a possible source of endogeneity.

We address this issue with two strategies. The first is to add variables that proxy for the omitted variable we are most concerned about, idiosyncratic political preferences. The second is the regression discontinuity design, which is closer to our model in so far as it comes close to recreating the condition of a random election.

2.3.1 Data

We use a sample of 47 US states for the period 1960-98[14]. Most political, fiscal, and population variables are the same as in Besley and Case (2003). We add the political variable on election results in states' lower offices, gathered by Ansolabehere and Snyder (2002) as a measure of political competition[15]. Some institutional and procedural variables, instead, have been collected from the National Association of State Budget Offices (NASBO) and the National Conference of State Legislatures (NCSL). We also conducted three e-mail surveys directed to state budget officers and legislature public officials to clarify ambiguous information and a few inconsistencies in the data.

The outcome variable we are interested in explaining is tax revenues over GDP. We call the variable *ttax_gdpp* and it is defined as the sum of state sales, corporate, and income taxes over potential GDP in 1982 dollars[16]. We use a Hodrik-Prescott filter to separate the cycle from the potential component of GDP. The average tax burden of an American state is around 5% of GDP. Socio-economic controls such as state population, state population square, proportion of aged (over 65) and kids (5 to 17) in the state are always included in the regressions. We also include the cycle

[14] There isn't enough data to include Alaska and Hawaii; Nebraska is excluded for being the only unicameral state.

[15] We are thankful to Stephen Ansolabehere, Timothy Besley, Anne Case, and James Snyder for making their data sets available to us.

[16] Alaska, Florida, South Dakota, Texas, Washington, and Wyoming do not have a state income tax. All results including the semiparametric RDD hold with their exclusion. On this note, we are bypassing a discussion on how tax rates are set in federal units taking into account the central government tax policy, see Klor (2005).

component over the trend, *cycle_trend*, to control for variations in tax revenue due to the cyclical behavior of GDP. To control for the presence of tax and expenditure limitations we use, *restrict*, which takes value 1 if such a limitation is present but it is advisory or may be overruled by a simple majority, and 2 if such a limitation can not be so easily overruled. We also include an indicator for the presence of super-majority requirements for a tax increase, *supmaj*; and indicators for the political identity of the governor: *demgov* and *indgov*.

Our variable of interest is divided government. To classify a government as *divided*, we first define a measure of the governor's strength in the legislature: the share of legislators with the same party identity as the governor. Since proposals must pass both chambers, our measure of strength is defined as:

$$gov\_strength = \min\{gov\_strengthHouse,\ gov\_strengthSenate\},$$

that is, the share of legislators belonging to the same party as the governor in the chamber where their numbers are the smallest. The variable *gov_strength* ranges from 0 to 1. Its conditional relation with the tax level is assumed to be continuous except at 0.5. Above it the governor's party has control of the agenda and of the veto; below it the agenda is at least partly controlled by the other party. If line item veto is present this should make a difference. The dummy *divided* takes the value 1 if $gov\_strength < 0.5$ and zero otherwise.

Throughout, we allow for the residuals of our regressions to have different variances across states and to be serially correlated. We show conservative heteroskedastic-robust standard errors in parenthesis and when the point estimate is also significant with the cluster-robust standard error, the standard errors are in boldface. Bertrand et al. (2004) show that when the sample of states is large, the use of clustered errors fairs well in face of intra-state serial correlation[17].

2.3.2 Fixed Effects

Full Sample

Our first step is to compare our results to those in Besley and Case (2003). Their explanatory variable is taxes per capita and ours is taxes over potential GDP, *ttax_gdpp*, as explained in the above section; they control for state income per capita and we for the cyclical component of GDP: *cycle_trend*.

The estimating equation is given by:

$$ttax\_gdpp_{st} = \zeta_s + \delta_t + \beta' X_{st} + \lambda LIV divided_{st} + \varepsilon_{st},$$

where ζ_s is a state fixed effect that allows us to control for time invariant state characteristics that can be correlated with institutional variables; δ_t is a year dummy

[17]We cluster by state even though we are in a limiting case since the number of clusters are less than the number of regressors if we include year and state dummies. Bertrand et al. (2004) run monte carlo experiments and find that the fully (cluster) robust estimator works well even when the cross-sectional is not much larger than the time series dimension. Theoretically, the use of the cluster-robust estimator is only justified as the number of clusters, states, is going to infinity.

capturing common shocks and trends; X is a matrix of controls, including socioeconomic and demographic characteristics, as well as other fiscal institutions; $LIVdivided_{st}$ is the interaction between *divided* and a dummy for state-years with line item veto. The results can be seen in **Table 1**.

Column 1 illustrates that contrary to the results in Besley and Case (2003) the interaction term *LIVdivided* is not significant, but *divided* by itself is. There are two states that adopt the line item veto right at the end of our sample. The variable *LIV* therefore is not completely time invariant and we add it in the regressions. It is, however, never significant.

A simple interaction, however, does not capture how the variable *LIV* may interact with all other variables. When we separate the sample in the following columns we see that *divided* has its negative sign only in the states with line item veto, and it significant also with clustered standard errors. The point estimate is around 0.14. A switch from a unified to a divided government is correlated with a decrease in the average tax rate of a state from 5% to 4.86% of GDP. The additional of political and institutional controls in column 3 does not change the results.

The political identity of the governor, a democratic (*demgov*) versus a republican, is not significant. The dummy for an independent governor, *indgov*, is positive and statistically significant. We point out that there are only 9 state-years with an independent governor in our sample. The dummy for a restrictive cap on the the tax rate, *restrict*, is significant but positive. Its sign may be related to the endogeneity

Table 2.1: Dependent Variable: ttax_gdpp–Fixed Effects

	(1) All States	(2) LIV States	(3) LIV States	(4) BV States	(5) BV States
divided	-0.12	-0.13	-0.15	0.08	-0.07
	(0.05)**	**(0.03)***	**(0.03)***	(0.07)	(0.03)
LIVdivided	0.01				
	(0.06)				
LIV	-0.13				
	(0.13)				
demgov	-0.02		-0.05		0.13
	(0.03)		(0.03)		(0.06)**
indgov	0.31		-0.09		0.43
	(0.16)		(0.14)		**(0.17)**
restrict	0.15		0.18		0.12
	(0.03)***		**(0.03)***		(0.13)
supmaj	-0.42		-0.39		0.59
	(0.06)*		**(0.06)***		(0.30)**
Observations	1833	1537	1537	296	296
R-squared	0.84	0.83	0.84	0.91	0.91
state and year dummies, population controls, cycle_trend included					

Huber-White robust standard errors in parentheses, standard errors in boldface are also significant with clustering by state. Number of clustered groups 47. The states of Nebraska, Alaska, and Hawaii are excluded in all regressions. The data set goes from 1960 to 1998. The symbol $*$ is significant at 10%; $**$ significant at 5%; $***$ significant at 1%. Population controls include state population, state population squared, percentage of aged, and of kids.

of its adoption by states with a high tax rate. The dummy for a supermajority requirement for a tax increase, *supmaj*, is significant and negative. For a detailed study on the supermajority requirement and its effects on the tax level controlling for the endogeneity of its adoption see Knight (2000b).

Restricting the Sample

Here we restrict the sample in the panel estimation to make it comparable to the sample we use in the regression discontinuity strategy: state-years with line item veto, a supermajority requirement for a veto override, and no supermajority requirement for a tax increase. This restrictions are in line with our model and allow all the state-years included to have comparable institutional features.

Our model assumes that the line item veto sticks once used. In states where the veto override requirement is a simple majority, in the absence of a override cost for the chambers, it is as if there was no veto. We restrict our comparison to state in which we are sure the veto sticks, at least around $gov\_strength = 0.5$.

The adoption of supermajority requirements makes the discontinuity at $gov\_strength = 0.5$ no longer relevant. A two-third majority is needed for a tax increase. We drop the 222 observations with this feature from the sample.

In Table 3 we only look at our restricted sample of 1159 observations. In column 2 we add the same political and institutional controls as in Table 1. In column 3 we include the amount of federal grants, *grant*, to the state in that year. It is significant

Table 2.2: Dependent Variable: ttax_gdpp–Restricted Sample

	(1) LIV_2/3 States	(2) LIV_2/3 States	(3) LIV_2/3 States	(4) LIV_2/3 States
divided	-0.10	-0.13	-0.13	-0.15
	(0.04)***	(0.04)***	(0.04)***	**(0.04)*****
demgov		-0.09	-0.08	-0.10
		(0.04)**	(0.04)**	(0.04)***
indgov		0.23	-0.37	-0.25
		(0.15)	**(0.15)****	(0.15)*
restrict		0.13	0.15	0.16
		(0.03)***	(0.03)***	(0.03)***
trend				0.01
				(0.00)***
grant			0.01	
			(0.00)***	
Observations	1159	1159	1070	1159
R-squared	0.85	0.85	0.86	0.85
state and year dummies, population controls, cycle_trend included				

Only observations with line item veto, override requirements of two-thirds, and no supermajority requirements for a tax increase are included in the regression. Huber-White robust standard errors in parentheses, standard errors in boldface are also significant with clustering by state. Number of clustered groups 47. The states of Nebraska, Alaska, and Hawaii are excluded in all regressions. The data set goes from 1960 to 1998. The symbol $*$ is significant at 10%; $**$ significant at 5%; $***$ significant at 1%. Population controls include state population, state population squared, percentage of aged, and of kids.

and positively correlated with the average tax revenues. An increase of 100 dollars per capita in grants is correlated with an increase in the average tax level from 5% to 5.01% of GDP. There is missing data for this variable and we choose to show results without it.

In column 4 we add the variable *trend*. If the economy of a state is growing tax revenues should increase even if there is no nominal increase of the tax rate. We decide to keep it out because *ttax_gdpp* is itself constructed as taxes revenues over *trend*, the muliticollinearity is high. The significance of *divided* is robust to its inclusion and to the inclusion of *grants*.

Endogeneity

Our main concern, whether or not we restrict our sample, is the possibility of omitted variables such as: idiosyncratic preferences over the tax level, candidates with particular tax profiles, and so on. To infer causality from the above results we would have to assume that *divided* is randomly assigned across state-years. Since *divided* is the result of elections and *LIV* is mostly time invariant we may expect our estimates of causal effects to be biased. If, for example, voters tend to vote a divided government to correct for a tax rate that is already high, we may expect our estimates in Tables 1 and 2 to be downward biased in absolute terms.

In order to make claims of a causal effect between a divided government and a lower tax rate in states with line item veto we move on to two strategies. First we present two variables that proxy for the omitted variables such as idiosyncratic political preferences. One is non partisan measure of political competition in elections for lower offices in the American states. Ansolabehere and Snyder (2002) collected election results for a number of directly elected state offices other than governorship and state assemblies; they include Attorney General, Lieutenant Governor, Auditor, etc.[18]:

$$pclow_{st} = -abs\left(VoteShareDemocratsatLowOffices_{st} - 0.5\right),$$

that is, the absolute value of the difference between the vote share democrats received in all lower level elections that year and 0.5. If the difference is zero elections are highly competitive between the two parties. The second variable is voter *turnout* in each election. In Table 3, columns 1 and 2, we show that these variables do no alter the significance of the correlation between *divided* and *ttax_gdpp*.

In columns 3 and 4 we hint at one future step of this research and include the *ttax_gdpp* lagged by one year. The tax rate is a highly persistence variable and this is seen by the point estimate of 0.84 of its lag. Estimating a model with a lagged dependent variable requires taking care of the endogeneity brought about the serial correlation. Here all we intend to show is that the point estimate of *divided* is highly significant after its inclusion albeit with a smaller point estimate. The large sample of almost 40 years give us some confidence that bias should be small, and by

[18] As in the data for election results for the state chambers, variation on lower level elections results only occur sporadically, varying by state every, on average every 3 or 4 years.

assuming that T goes to infinity we know that the estimate is consistent[19].

2.3.3 Discontinuity design

Regression Discontinuity is a quasi-experimental design with the defining characteristic that the probability of receiving treatment changes discontinuously as a function of one or more underlying variables[20]. The treatment, call it t, is known to depend in a deterministic way on some observable variable g, $t = f(g)$, where g takes on a continuum of values, and there exists a known point g_0 where the function $f(g)$ is discontinuous.[21] Around g_0 control and treated observations should be similar in observable characteristics and their unobservable characteristics are assumed not to differ systematically. Any discontinuity in the outcome of interest is attributed to the treatment status at g_0, since the uncontrolled factors are likely to behave similarly.

In our context, the tax rate (conditional on observables) is assumed to be a continuous function of the variable *gov_strength*, and we test for a discontinuity at 0.5. The approach consists in estimating the shape of the function of $E[ttax\_gdpp|X]$ on *gov_strength*. If we correctly identify the shape of the function we are able to estimate its jump.

Our model treats both the gubernatorial and parliamentary election as random.

[19] Arellano (2003) recalls this result in pg.84.

[20] For a detailed review of the regression discontinuity design and an application to election results see Lee (2005)

[21] More formally, the limits $t^+ \equiv \lim_{g \to g_0^+} \mathbf{E}[t|g]$ and $t^- \equiv \lim_{g \to g_0^-} \mathbf{E}[t|g]$ *exist* and $t^+ \neq t^-$. It is also assumed that the density of g is positive in the neighborhood of g_0. There are two types of discontinuity design, fuzzy and sharp designs. In the sharp design the treatment is known to depend in a deterministic way on some observed variables whereas in the fuzzy design there are also unmeasure factors that affect selection into treatment. Our case fits the sharp design.

Table 2.3: Dependent Variable: ttax_gdpp–Preferences

	(1) LIV_2/3 States	(2) LIV_2/3 States	(3) LIV_2/3 States	(4) LIV_2/3 States
divided	-0.12	-0.09	-0.05	-0.05
	(0.04)***	(0.04)**	**(0.02)***	**(0.02)***
demgov	-0.10	-0.08	0.00	-0.00
	(0.04)***	(0.04)**	(0.02)	(0.02)
indgov	-0.26	-0.19	-0.56	-0.54
	(0.14)*	(0.14)	**(0.08)***	**(0.08)***
restrict	0.12	0.08	0.05	0.04
	(0.03)***	(0.03)**	**(0.02)***	(0.02)**
turnout	-2.00	-0.84		-0.17
	(0.40)*	(0.43)*		(0.22)
pol_comp_low		1.06		-0.24
		(0.37)*		(0.21)
lag_ttax_gdpp			0.84	0.83
			(0.02)*	**(0.02)***
Observations	1159	1094	1127	1070
R-squared	0.85	0.86	0.96	0.96
state and year dummies, population controls, cycle_trend included				

Only observations with line item veto, override requirements of two-thirds, and no supermajority requirements for a tax increase are included in the regression. Huber-White robust standard errors in parentheses, standard errors in boldface are also significant with clustering by state. Number of clustered groups 47. The states of Nebraska, Alaska, and Hawaii are excluded in all regressions. The data set goes from 1960 to 1998. The symbol $*$ is significant at 10%; $**$ significant at 5%; $***$ significant at 1%. Population controls include state population, state population squared, percentage of aged, and of kids.

The regression discontinuity design comes close to recreating the randomness of elections by only looking at a small sample around the 0.5 discontinuity. There should not be significant differences in the observables and unobsevables (conditional on state and time effects, social, economic, and political controls) for state-years close to the discontinuity. We show results for the sample as in Table 2 and Table 3: state-years with line item veto, override requirements of two-thirds, and no super majority requirements for a tax increase.

One caveat remains. The ideal experiment, in accordance to our theoretical model, would be to look at simultaneous elections in which both the gubernatorial and the results in the chambers are close to 50%. Our sample size does not allow us to follow this strategy. We assume throughout that close elections are comparable whether they are midterm or simultaneous.

2.3.4 Semiparametric Regression Discontinuity Design

Semiparametric Procedure

We implement a semiparametric estimation as presented in Robinson (1988)[22]. In his procedure, one of the covariates enters the model nonlinearly. The procedure estimates the model without making parametric assumptions on the shape of the nonlinear relation. We are interested in the shape of $E[ttax\_gdpp \mid X]$ on $gov\_strength$. We reproduce his procedure except that, at the stage in which the the function is

[22]For a summary of the procedure and applications see Ichimura and Todd (2006).

estimated non parametrically, we allow for a discontinuity at $gov\_strength = 0.5$.

The model we are estimating is:

$$ttax\_gdpp_{st} = X'\beta + f(g_{st}) + \epsilon_{st},$$

where X is the matrix with state and year dummies, population, economic, and political controls. The function $f(g_{st})$ is the non-linear part of the model, and $g = gov\_strength$.

The first step is to estimate the correlation between g and all the other variables. We estimate each correlation non parametrically with a local linear regression. The definition of local linear regression can be found in the appendix.

The βs are estimated by the following OLS regression:

$$\hat{\beta} = (\overline{X}'\overline{X})^{-1}\overline{X}'\overline{\tau},$$

where each column of the matrix $\overline{X}$ is the fitted errors of the local linear regression of each column of X on g. The vector $\overline{\tau}$ is the fitted errors of a local linear regression of $ttax\_gdpp$ on g. If the density of g is zero or close to zero at any point, the estimator is unreliable and we solve this by trimming 4% of lowest density points of g. Our choice of bandwidth is $h = 0.05$, which is slightly lower than the rule-of-thumb bandwidth. We discuss its choice, the choice of kernel, and of the local linear regression method in the Appendix.

Once we have the $\hat{\beta}$s we retrieve the fitted errors:

$$\overline{ttax\_gdpp} = ttax\_gdpp - X'\hat{\beta}.$$

The shape of $f(g)$ is identified by running another local linear regression of $\overline{ttax\_gdpp}$ on $gov\_strength$. But since we are allowing for a discontinuity, we estimate one for $gov\_strength < 0.5$, and one for $gov\_strength \geq 0.5$. The result can be seen in **Figure**2.1.

The bandwidth and Kernel choice are illustrative of the intuition of this result. A bandwidth of $h = 0.05$ with the kernel of our choice implies that for the estimation of the local linear fitted value of $\overline{ttax\_gdpp}$ for a given value of $gov\_strength = g_o$, only data in the interval $[g_o - 0.055, g_0 + 0.05]$ is used, and more weight is given to the observations closer to g_o. In the next subsection we allow for a discontinuity. The closest the point being estimated is to the discontinuity the less data is used for its estimation. At the point of discontinuity itself on the right side only observations with $gov\_strength \in [0.5, 0.55]$ are included; on the left side only observations with $gov\_strength \in [0.45, 0.5)$ are used. In **Figure**2.2 we zoom into the discontinuity we have estimated.

The graphs have been produced estimating the model as in column 2 in Table 2. The sample is restricted to states with line item veto, override requirements of two-thirds, and no supermajority requirements for a tax increase. The following controls are added: state and year dummies, state population, state population squared, percentage of aged, of kids, *cycle_trend*, *restrict*, and *demgov*.

In Table 4 we show bootstrapped standard errors of the estimated discontinuity with different control choices. We bootstrap the residuals of our model 100 times.

Figure 2.1: Non Parametric Discontinuity - LIV_2/3

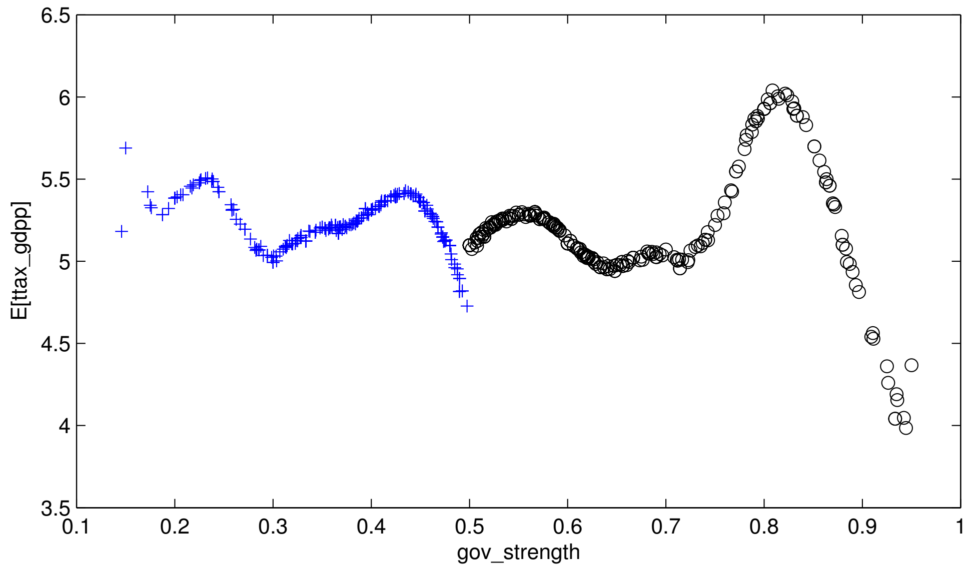
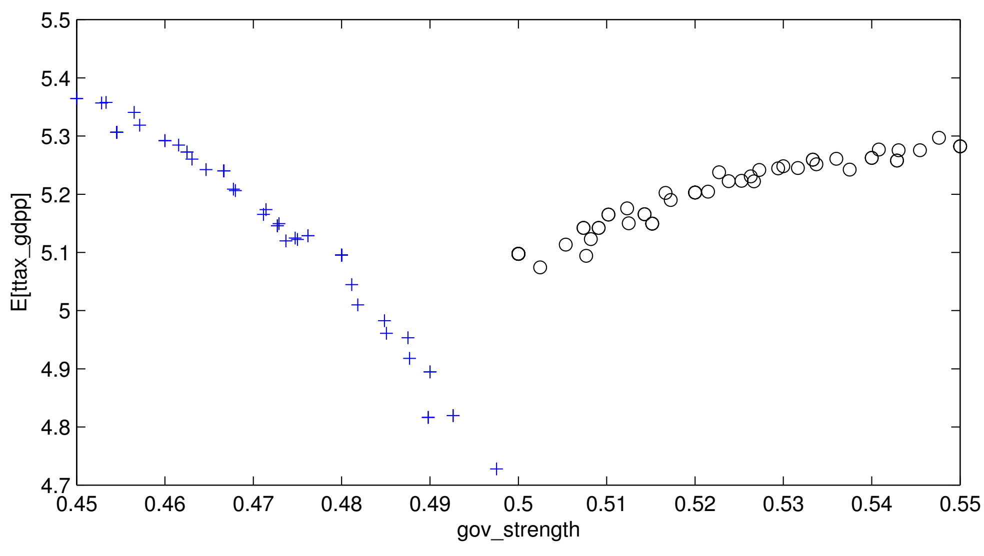


Figure 2.2: Non Parametric Discontinuity - LIV_2/3 - Zoom



Each bootstrap estimation consists in re-estimating the model and the jump at the discontinuity by adding a different sample of the bootstrapped residuals to the fitted dependent variable of the original sample estimate. As we can see, the jump at the discontinuity is highly significant and the different specifications do not alter the estimated discontinuity by much.

Our point estimate at the discontinuity is in the order of 0.3, two to three times higher than the point estimate in the fixed effects model. The switch from unified to divided government in a state with line item veto brings down the tax level from an average of 5% to 4.7%.

A word on the choice of bandwidth is in place. The efficiency of the estimation depends much more on the bandwidth than on the kernel selection. Too large a bandwidth and we may be oversmoothing our function of interest; too narrow and we may be subject to local outliers. Below we show results for three different bandwidths around our rule-of-thumb bandwidth of $h = 0.057$. The choice of an optimal bandwidth, however, still is an open question in the literature, specially for

Table 2.4: Nonparametric Estimation of $E[ttax\_gdpp \mid X]$ on *gov_strength*–Bootstrap

Controls	**Discont. at 0.5**	**Bootstp Mean**	**(Std. Err.)**
baseline	-0.3826**	-0.3776	(0.1752)
plus *demgov* and *restrict*	-0.3808**	-0.3761	(0.1573)
plus *turnout*	-0.3759**	-0.3657	(0.1655)
plus *pol_comp_low*	-0.2710*	-0.2241	(0.1596)
baseline plus *lag_ttax_gdpp*	-0.2605**	-0.2702	(0.1079)

Bootstrapped standard errors were retrieved resampling the residuals 100 times with replacement. Baseline controls are population controls, *demgov*, *restrict*, *cycle_trend*, state and year dummies.

semiparametric estimates where each individual regression may call for a different bandwidth. For a detailed account of this literature see Ichimura and Todd (2006).

Block Veto States

The graphic results from the line item veto states can be compared to the those in the states with block veto, see **Figure** 2.3. The result is not reliable because the sample is much smaller. It seems, however, to point in the direction of a reverse result and common pool problem. Taxes seem to be higher when the control of the legislature is shared by both parties.

Testing for Spurious Discontinuities

In this section we rerun the procedure but try to estimate a discontinuity where there should be none. We show the results graphically for $gov\_strength = 0.49$, $gov\_strength = 0.51$, $gov\_strength = 0.45$, and $gov\_strength = 0.55$. The results can be seen in the **Figures** 2.4 to 2.7 below.

Table 2.5: Nonparametric Estimation of $E[ttax\_gdpp \mid X]$ on *gov_strength*–Bandwidth

Bandwidth	**Discont. at 0.5**	**Bootstp Mean**	**(Std. Err.)**
h=0.05	-0.3808**	-0.3761	(0.1573)
h=0.45	-0.3754*	-0.3785	(0.2258)
h=0.6	-0.3747**	-0.3023	(0.1551)

Bootstrapped standard errors were retrieved resampling the residuals 100 times with replacement. Baseline controls are population controls, *demgov*, *restrict*, *cycle_trend*, state and year dummies.

Figure 2.3: Discontinuity Block Veto

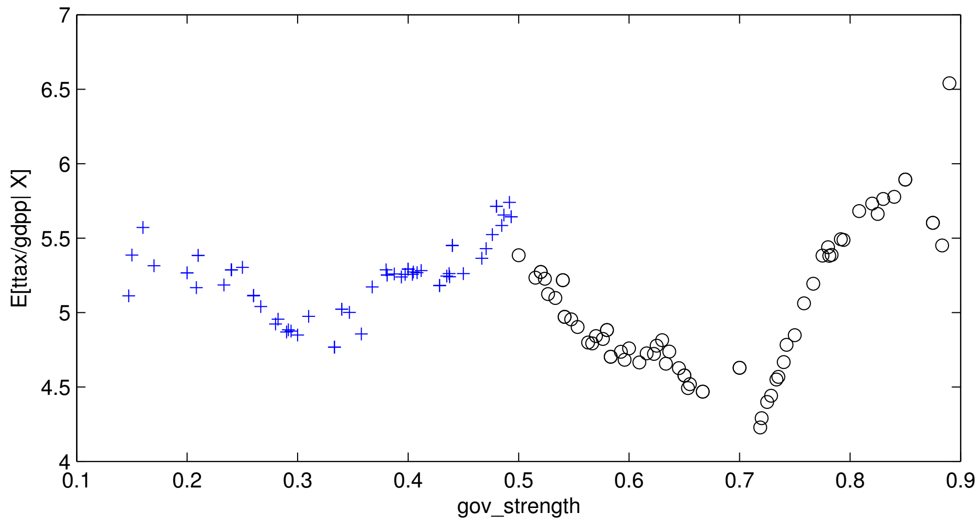


Figure 2.4: Discontinuity at 0.49

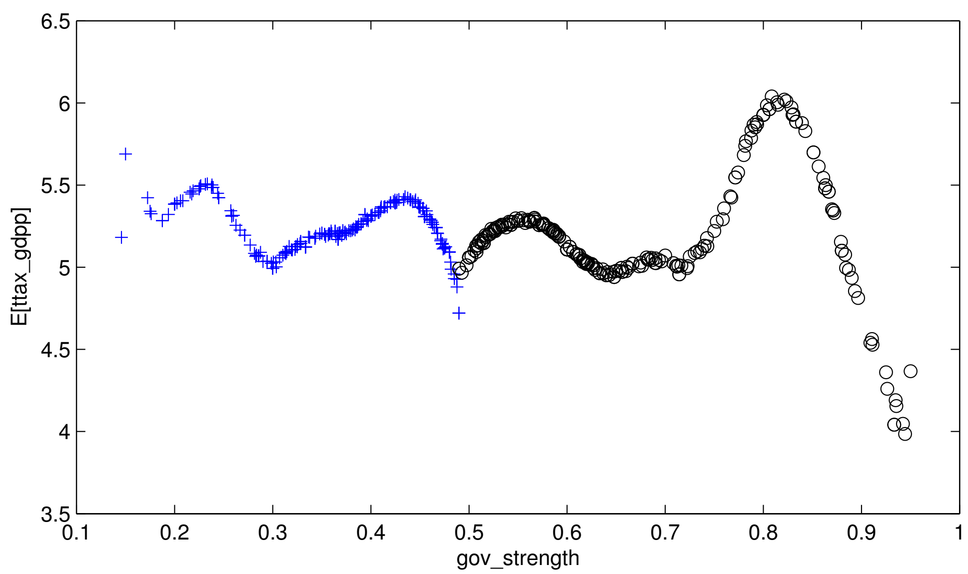


Figure 2.5: Discontinuity at 0.51

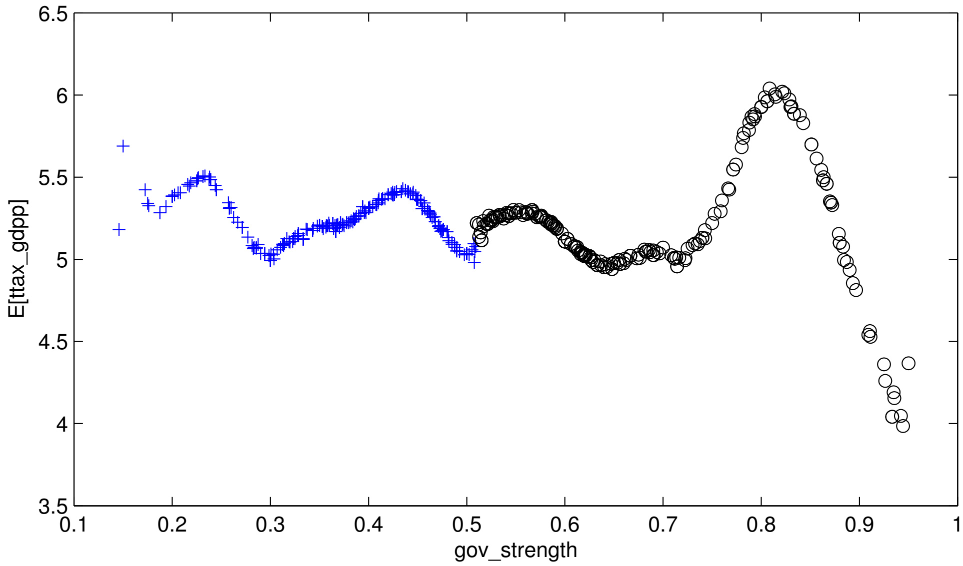


Figure 2.6: Discontinuity at 0.45

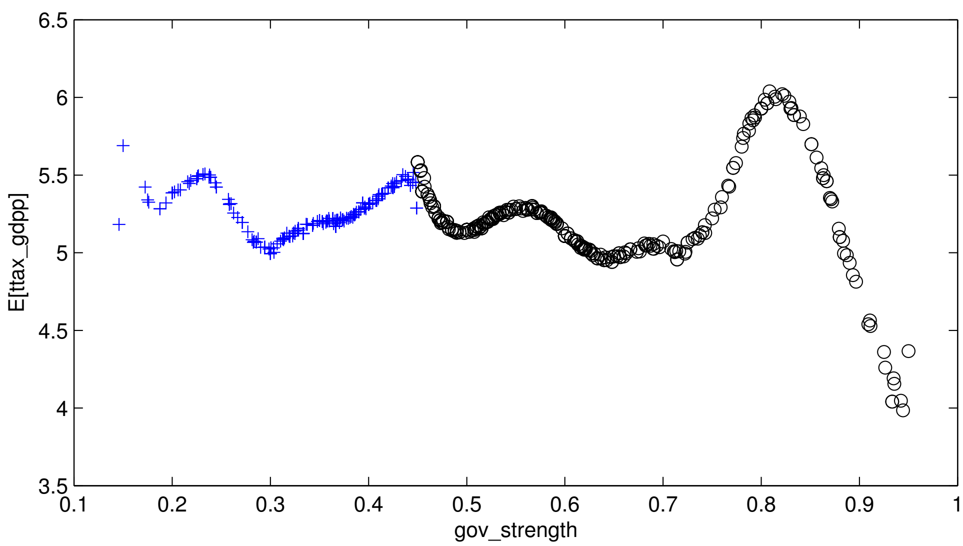
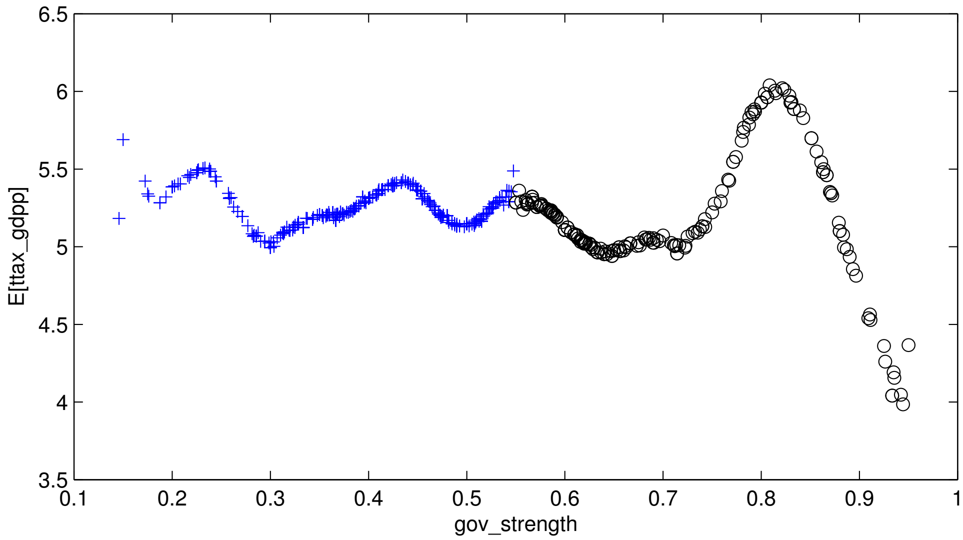


Figure 2.7: Discontinuity at 0.55



2.4 Concluding Remarks

Under the identifying assumptions of the regression discontinuity design we have established a causal relation between the type of political power, divided or unified government, and the tax level in the states that have the institution of line item veto. Moving from a marginally unified government to a divided government decreases the average tax rate from say 5% to 4.7% of GDP. Similar results for the states with block veto are not found, confirming what was predicted by the model.

The model identified budgetary separation of powers in the American states only when the government is divided in a state with line item veto. Even if the American states are classified as presidential systems with clear separation of powers, *budgetary separation of powers* is only present when certain institutional and political conditions are met. We look forward to further work on trying to identify budgetary separation of powers in other cases, be it presidential or parliamentary systems.

Regression discontinuity design will always be attractive for political economics, since the control of the agenda usually changes hands at 50%, and its use will certainly increase. The use of a semiparametric method to estimate the conditional relation between the dependent variable and the non-linear function of interest with its discontinuity has the appeal of not relying on functional form assumptions.

Chapter 3

Dynamics and Variable Treatment Effects of Budgetary Separation of Powers

3.1 Introduction

In Ferrero and Magalhães (2007), we define budgetary separation in the American states as the intersection of two events: the presence of the institution of line-item veto in hands of the executive/governor, and the presence of a divided government, meaning that the party controlling the legislature differs from the party identity of that of the executive. We model and test the prediction that only under budgetary separation of powers should we expect a lower size of government, measured as the average tax rate. The estimation strategy is grounded on microeconometric tools, mainly, regression discontinuity. We exploit the parallels of the budgetary separation problem with those of the the program evaluation literature: a binary, treatment-like, variable of interest, constructed upon another variable with a exogenous switching point, and potentially important self-selection/endogeneity problems. Our empirical analysis provides quasi-experimental evidence of important negative effects of budgetary separation on the average tax level.

However, it holds that the econometric theory developed in the context of micro panels is somewhat inappropriate for macro applications. Estimators are typically constructed for samples which have a small time series (T) and large cross section (N). Therefore the properties of estimators are derived exploiting asymptotics in the cross section. In macro panels, typically, neither N nor T are large. In addition, macroeconomic variables such as tax revenues, have important dynamic properties that may affect the validity of the estimation results and inference. Furthermore, in our case, past values of the average tax rates may influence voters evaluation of government performance, thus affecting voting decision.[1] It is straightforward to argue that voters, at least less ideological ones, reward/punish according to government performance.

[1]Another crucial problem in macro data is dynamic heterogeneity, which in this case may reflect different political, budgetary procedures, or regulations. We consider that the time span for each unit, 39 years, is short to implement SUR and panel VAR models in which time varying coefficients are examined.

In this paper we introduce a dynamic approach to studying the effects of budgetary separation, and we take on a more extensive look at the effects of budgetary separation of power as interacted with features of fiscal institutions and politics. The introduction of dynamic framework not only allow us to control for potential feedbacks to voting decision, but also to explore the time series properties of the variables of interest—serving as robustness checks of our previous results. We present the alternative estimation methods in this set up, conditional on the time span available. The result that budgetary separation of powers matters, negatively affecting the average tax level, goes through. Moreover, in this setup we can distinguish the effect on impact and the long run effect, which is similar in levels to the regression discontinuity point estimate in Ferrero and Magalhães (2007).

The notion of budgetary separation of powers has important, additional implications, when interacted with politics and other fiscal policy institutions. Firstly, we consider the potential interaction effects of two different types of fiscal institutions: one imposing constraints on outcomes, as that of *tax and expenditure limitations,* and another one imposing constraints on the budgetary process, as that of *supermajority requierments* for tax increases. We find that states with formal ceilings on taxes and expenditures are, in short, self-selected as the estimates are positive and significant in all specifications. More interestingly, based on the idea of budgetary separation, we argue that supermajority requirements should not be expected to have a direct effect on the average tax level *per se*, but an indirect effect as it broadens the extent of budgetary separation. With a supermajority requirement it is no longer enough for a party (group) to control one-half of the legislators in order to raise taxes and appropriate the residual proceeds; it takes at least a two-third majority to do so, making it more difficult to raise taxes along the budgetary process. We present robustness checks, including additional controls for political preferences and turnout; we also treat the variables of interest as predetermined and endogenous (We follow standard simulation results (Canova (2007);Arellano (2003)) when choosing the lag structure of the set of instruments).

Secondly, we extend the framework to the analysis of budget composition—instead of size. The budgetary bargaining model in Ferrero and Magalhães (2007) implied that the governor's incentive was to line-item veto spending programs should that are targeted to opposing political groups. General spending programs (perfect public goods) should not be as sensible to budgetary separation as specific transfers. We present preliminary evidence that this is the case. Finally, we show that the effects of budgetary separation of powers are stronger under a republican governors facing a democrat controlled legislature.

Another key modelling assumption in Ferrero and Magalhães (2007) is perfectly party alignment: the constituency groups for legislators and for the governor perfectly overlap when they share party affiliation. There is no within party conflict. This stringent assumption is justified in the context where the legislature has two parties of similar size fighting for control. We relax it here and look for a measure of degree of party alignment. In a different context—credibility of policy platforms for governor elections—, Grossman and Helpman (2005) argue that the higher the share of independents in the voting population, the higher the within party conflict as target constituencies diverge.[2] By analogy, in our case, the higher the share of independent voters, the less stark is party alignment, and the bite of budgetary separation of powers. As legislators try to target programs to their own constituency, there is room for active line-item vetoing even when the legislature is controlled by the same party as the party identity of the governor. We find empirical evidence that less party alignment lowers the budgetary separation effect in a significant and robust way.

3.2 Dynamics and Feedbacks

Panel data—or Time Series-Cross Sectional data—, is now widely used to estimate dynamic econometric models. While controlling for time invariant unobserved heterogeneity, it provides sufficient information for dynamic relations to be investigated. The introduction of dynamic framework to study the effects of budgetary separation in the American States, not only allow us to control for potential feedbacks on voting decisions, but also to explore the time series properties of the variables of interest, which has a time series nature. Besides, neglecting dynamic information can be

[2]In their case, this conflict is anticipated by voters, and the governor candidate when he sets his platform. Still, optimal platforms diverge both from the candidate's bliss policy (determined by her constituencies' preferences), but also from the ex-post implemented budget if she wins the election. Credibility is an issue, since the lower party alignment, the less credible platforms are.

costly biases and inference validity—serving as robustness checks of our previous results. Finally, the introduction of dynamics permits to distinguish between impact and long run average effects of budgetary separation and, thus, to better evaluate the potential effects of shifts in government composition on fiscal policy along the business cycle.

In this section we introduce the estimation of a single equation, autoregressive - distributed lag model to the panel of the American States, the same used in Ferrero and Magalhães (2007), with the number of states (N) still greater than the time span (T), but fixed N. This middle-ground feature requires at least some discussion on the estimation method chosen, and comparisons with other alternatives. Consequently, we present the alternative estimation methods in this set up, conditional on the time span available. The result that budgetary separation of powers matters, negatively affecting the average tax level, goes through. The point estimate on impact is lower than that of the regression discontinuity, but the long run effect is similar in level.

3.2.1 Specification and Estimation

The average tax rate is as a highly persistent variable. Moreover, strict exogeneity assumptions maintained under *within* panels specifications, rule out an important feedback effect: variations in taxes in $t-j$ affect voters's decisions at time t, either changing the size of a governor's support in the legislature, or changing the party identity of a governor for a given composition of the legislature. This a clear violation that can bias our estimates systematically, and one major concern. If that were the case, omitting the autoregressive component results in biased estimates of the budgetary separation effect, as it is uncorrelated with the error component in the present and future periods, but it may be correlated with the error component in previous periods through the feedback, as it is clear from the feedback equation (3.1). This would result in contemporaneous correlations for the estimating equation in first differences.

$$LIVdivided_{st} = \xi' \mathbf{z}_{st} + \sum \lambda_j \tau_{s,t-1} + v_{st}. \tag{3.1}$$

with the vector of controls $\mathbf{z}_{st}$ including time and state fixed effects, $LIVdivided$ is the binary treatment for budgetary separation of powers equals to 1 when a given state has a divided government and line-item veto; τ is the average tax rate: the sum of tax revenues divided by potential GDP obtained with a HP filter. The same source of bias affects ρ in (3.2), the dynamic equation of interest, as the autoregressive component is weakly exogenous. Both sources of biases can be addressed combining dynamic panels and instrumental variables estimates for our treatment variables.

$$\tau_{st} = \beta' \mathbf{x}_{st} + \delta_I \cdot LIVdivided_{st} + \sum \rho_j \tau_{s,t-j} + \varepsilon_{st}. \tag{3.2}$$

with the vector of controls $\mathbf{x}_{st}$ including time and state fixed effects, and the cyclical component of state GDP, divided by the trend, to control for fluctuations on average tax rates due to business cycle. δ_I, the coefficient of interest, has now an impact/short run interpretation. In steady state, the multiplier $\widehat{m} = \frac{1}{1-\sum \widehat{\rho}_j}$ can be used to retrieve the long run average effect.

We follow with a discussion of the alternative estimating methods for (3.2), and compare their results. In this section, we consider *LIVdivided* as strictly exogenous conditional on $\left(\mathbf{x}_{st}, (\tau_{s,t-j})_j\right)$. In section 4 we allow it to be correlated with past and contemporaneous error realizations.

Alternative estimation methods The models we consider still borrow from the micro panel literature in the sense that the specifications do not allow for lagged interdependencies across units. We consider three alternative strategies to estimate the above specification[3]: OLS with lags of the dependent variable, including time and fixed effects; and two increasingly popular related methods: the Arellano and Bond (1991) and Arellano and Bover (1995)/Blundell and Bond (1998) dynamic panel estimators. The latter are general estimators designed for situations with (1) relatively small T and large N panels, meaning fewer time periods related to units of observation in the cross section; (2) a linear functional relationship; (3) a single left-hand-side variable that is dynamic, depending on its own past realizations; (4) independent variables that may not be strictly exogenous, meaning possibly correlated with past (weakly exogenous or predetermined) and eventually current realizations of the error (endogenous); (5) fixed individual effects; and (6) heteroskedasticity and autocorrelation within individuals, but not across them.[4]

[3]We consider that the time span for each unit, 39 years, is short to implement SUR and panel VAR models with time varying coefficients are examined.

[4]Arellano and Bond (1991) (AB) estimation starts by transforming all regressors, usually by differencing, and uses the Generalized Method of Moments, and so is called “difference GMM”.

First differencing the above equation removes the state fixed effect and produces an equation that can be estimated using instrumental variables. Arellano and Bond (1991) derive a generalized method-of-moments estimator using lagged levels of the dependent variable and the predetermined variables and differences of the strictly exogenous variables. that instruments the differenced variables that are not strictly exogenous with all their available lags in levels. (Strictly exogenous variables are uncorrelated with current and past errors.) Arellano and Bond (1991) also develop an appropriate test for autocorrelation, which, if present, can render some lags invalid as instruments. This method assumes that there is no second-order autocorrelation for $\Delta\varepsilon$, in the estimating first-differenced equation—equivalently, no first order autocorrelation for the error in (3.2). We include Arellano and Bond (1991) autocorrelation tests for all specifications.

A problem with the original Arellano and Bond (1991) estimator is that lagged levels are poor instruments for first differences if the variables are close to a random walk. Arellano and Bover (1995) describe how, if the original equation in levels is added to the system, additional instruments can be brought to bear to increase efficiency. In this equation, variables in levels are instrumented with suitable lags of their own first differences. The assumption needed is that these differences are uncorrelated with the unobserved country effects. Blundell and Bond (1998) show

The Arellano and Bover (1995)/Blundell and Bond (1998)(BB) estimator augments Arellano and Bond (1991) by making an additional assumption, that first differences of instrumenting variables are uncorrelated with the fixed effects. This allows the introduction of more instruments, and can dramatically improve efficiency. It builds a system of two equations—the original equation as well as the transformed one—and is known as "system GMM".

that this assumption in turn depends on a more precise one about initial conditions.[5]

With sequential or weakly exogenous variables $\mathbf{x}$, the implied moments conditions are $E\left(\mathbf{x}_{sj}^{\prime}\Delta\varepsilon_{st}\right) = 0$, for $j = 1, 2, ..., t-1$. These conditions open up a variety of estimation procedures, with $\mathbf{x}_s^{t-1} \equiv (x_{s1}, x_{s2}, ..., x_{st-1})$ and its linear combinations as potential instruments for Δx_{st}, for the equation in first differences. With other forms of endogeneity, the set of potential instruments made up of lags (and leads), varies according to the maintained assumptions. We use the set of available instruments under the maintained assumptions: for τ_{st-j}, we use $(\tau_{s1}, ..., \tau_{st-j-1})$ as instruments. As a practical matter, GMM estimators using many overidentifying restrictions are known to have poor finite sample properties (Wooldridge (2002), pp. 305).

Following Canova (2007), there are at least three issues of practical interest worth discussing when estimating models with homogeneous dynamics and unit specific fixed effects. First, it is well known that OLS estimates of the (common) AR parameters are biased when the model is dynamic, and that the bias is decreasing in T. (The predetermined character of the autoregressive component in dynamic panels motivates the use of instruments to reduce that bias.) Second, we know that GMM is more efficient than IV based on a single instrument, but also that estimates of the weighting matrix converge very slowly. Put differently there is a trade off between bias and efficiency in GMM estimators. The relative size of N and T are crucial in determining the point in the trade off. Using artificial data, with N significantly

[5] For a comprehensive view of dynamic panel data methods, cfr. for example Arellano (2003), and, for a macro perspective,Canova (2007).

lower that T, Canova (2007) concludes the bias induced by the estimation of the optimal weighting matrix is significant and the one-step estimator is always best; that the bias in the two-steps estimator increases, surprisingly, with T and, is larger the larger is the AR coefficient. Second, using two instruments typically produces smaller biases—and up to five instruments the bias is more precisely estimated. As expected, GMM estimators perform better when N is large but, for a fixed N, their performance is weaker. Overall, GMM and OLS biases are similar, when using a one-step estimator.

These source of biases can be addressed combining dynamic panels and instrumental variables estimates for our treatment variables. Using accepted unit root tests—augmented Dickey-Fuller for panels, and Levin et al. (2002)— we reject the unit-root null of our outcome variable in all cases. Since the the sum of AR(p) coefficients, mainly AR(2) and AR(1) , are between 0.6 and 0.85, we cannot neglect the weak instrument problem suggested by Blundell and Bond (1998). Therefore, we present OLS, AB and BB estimation results. As is well know AB estimates a first differenced equation, instrumenting predetermined—first differenced—variables with levels for lagged dependent variables, whereas BB is more flexible allowing also for instruments in first differences.[6] Two-step estimates are omitted from the results shown since, as it is well know, they are not robust, and standard errors tend to be

[6] The approach remove ζ_s by first differencing (3.2), and then instrument for the predetermined lagged dependent variables. When the dependent variable is close to a random-walk, instrumenting first differences with level leads to a weak instrument problem—as first differences are close to a random walk.

severely downward biased.

Data and Results We use the same panel as in Besley and Case (2003) with the additions detailed in Ferrero and Magalhães (2007): 47 states across 39 years, 1960-1998. To begin with, we divide results according to three different samples: All States, include the whole sample; LIV States, include only states with line-item veto with two-thirds veto override requirements, and with no supermajority requirements for tax increases; and BV States, include those states with no line-item veto whatsoever. This partition is justified on several ground, but mainly on the noise that the whole sample carries due to complementary institutions. For example, and as we will discuss in section 3, the adoption of supermajority requirements for a tax increase makes the definition of divided irrelevant for our purposes; in states with no supermajority requirements, a simple majority is needed to control allocation of resources and the notion coincides. Analogously for states with line-item veto with one-half override requirement: the political clout needed to override the veto is the same as that of passing the law, making the line-item veto irrelevant. So, the variables as they are in the whole sample actually fail to identify the budgetary separation effect, and therefore the control divided fails to pick unobserved differences between the budgetary separation effect and pure divided effect. As we will see in the next section, we take these types of variability into account and controlling for divided does work in the desired way. Still, by partitioning the sample here we first eliminate the noise in LIV States, and show that the effect of divided governments differ between LIV States and States featuring only block veto.

All regressions have additional controls as those shown: state and year fixed effects, state population, percentage of people above 65 years (aged), and percentage of kids, whether a state has restrictive rules for tax and expenditure limitations(exp_lim), and supermajority requirements($supmaj$), the party identity of the governor($demgov$); these are the basic controls in the previous paper; not shown but with identical results in the variables of interest are obtained adding federal grants, percentage of black population. Lags of the dependent variable are instrumented with two further lags in levels (AB). Column (7-9) use again the GMM estimation but with instruments in differences in a BB framework. (System GMM, with both equation in levels and differences, are not shown but available upon request.) The number of instruments in both GMM procedures are always from 1 to 5 lags of the predetermined variable—as suggested in Canova (2007)'s simulation results and Arellano and Bond (1991).

According to the results in **Table 1,** when comparing the different methods clear patterns emerge. First, point estimates of the effect of divided government for states with line-item veto are more efficient and slightly higher than divided, when we narrow the sample to LIV States; divided is no longer significant when we restrict the sample to those states with block veto. So, divided captures the budgetary separation effect when using the whole sample, as they overlap for more than 3/4 of the sample. This result is consistent with those found in the static estimation, and are in line with the foregoing discussion.

On impact, the effect of a divided government in a state with line item veto is negative and significant. These results are robust to different sets of controls, and the different estimation methods. The short-run effect ranges from 0.05 to near 0.07 of state taxes over state GDP percentages, as the upper bound in absolute terms. This implies that, for an average state with 6% of taxes over state GDP, taxes increase up to 1.2% on impact when switching status from separation to alignment.

The dynamic specification allows us to compute the expected long-run effect: in steady state, the multiplier $\widehat{m} = \frac{1}{1-\sum \widehat{\rho}_j}$ ranges from 2 to 5, taking the overall long run effect to vary from -0.10 to -0.35—which clearly is a big range of variation. We know that OLS estimates of the autoregressive component are biased, and that GMM ones are less biased when $N > T$. Besides robustness of the estimates, and auto-correlation tests, favor GMM estimates, particularly AB's. Note that AB estimates of the autocorrelation component are the lowest; the inclusion of AR(2) is significant and a sine qua non for the overidentification tests to go through. The lower value of the autocorrelation component $\sum \rho_j$, suggests that the weak instrument problem for the differenced equation is not a problem. However, Sargan and Hansen $J-$tests in the AB specification are only marginally above the 10% p-value. This problem is worked out in the following sections. Finally, autocorrelation tests for the error component are as expected in columns 4-9, with GMM estimation methods instrumenting predetermined variables.

We present only a restricted set of regression outputs. However, the results shown in Table 1 are robust to the exclusion of controls and the addition others—mainly, federal grants and percentage of black population. Proxies for political competition preferences, and voters turnout, are left for a more general discussion in section 4. The results are also robust to changes in the vector of instruments for the autoregressive component, with further lags of instruments going from 1, or 2, to 5, although overidentification and autocorrelation tests are not always valid. They are also robust to a distributive lags specifications, but lagged treatments are never significant in all specifications, once the contemporaneous treatment is included.

As final comments for the section, first, divided government is always positive and not significant in all specifications using the Block Veto sub-sample. This pattern is repeated also in the next exercises—not shown. Second, the point estimates for the budgetary separation effect are robust to the different estimation methods and controls. Third, one step AB procedure yields the most robust estimates, also for the autoregressive component—under different set of instruments and controls. However, overidentification tests are weak, barely exceeding the 10% p-value. This value cannot be improved upon within the current strategy, e.g., maintaining the strong exogeneity assumption for the variable present in all specifications. In section 4, we tackle this issue, and assume that our treatment variable is predetermined (weakly exogenous), or correlated with past shocks, and also endogenous. The result: overidentification tests improve substantially in all specifications.

Before doing so, and addressing potential sources of endogeneity, we revise our measure of separation of powers. In that direction, we need to explore the workings of other fiscal institutions, under the budgetary separation/residual claimant framework.

Table 3.1: Dynamic and Feedback Effects

Explanatory Variables	All States (1)	LIV States (2)	BV States (3)	All States (4)	LIV States (5)	BV States (6)	All States (7)	LIV States (8)	BV States (9)
LIVdivided	-0.00			-0.06			-0.07		
	(0.03)			(0.05)			(0.03)**		
Divided	-0.05	-0.05	0.03	0.01	-0.06	0.08	0.01	-0.06	0.02
	(0.03)*	**(0.02)*** **	(0.04)	(0.05)	**(0.03)** **	(0.05)	(0.02)	**(0.01)*** **	(0.05)
LIV	-0.01			-0.14			-0.05		
	(0.06)			(0.15)			(0.05)		
Lag1_ttax_gdpp	0.84	0.84	0.82	0.62	0.63	0.45	0.68	0.70	0.53
	(0.01)* **	**(0.02)*** **	**(0.05)*** **	**(0.05)*** **	**(0.06)*** **	**(0.09)*** **	**(0.02)*** **	**(0.05)*** **	**(0.10)*** **
Lag2_ttax_gdpp				-0.10	-0.11	-0.11			
				(0.03)* **	**(0.04)*** **	(0.07)			
cycle_trend	0.11	0.10	0.27	0.05	0.06	0.12	0.03	0.05	0.05
	(0.04)* **	(0.05)**	(0.12)**	(0.03)	(0.05)	(0.11)	(0.02)**	**(0.01)*** **	(0.11)
Est. Method	OLS	OLS	OLS	AB	AB	AB	BB	BB	BB
Overrid Test (p)	·	·	·	0.11	0.13	0.12	0.23	0.88	0.03
Res. AR(1) (p)	0.10	0.35	0.58	0.00	0.00	0.00	0.00	0.00	0.00
Res. AR(2) (p)	0.00	0.00	0.41	0.82	0.72	0.48	0.13	0.17	0.26
# Observations	1786	1127	286	1692	1063	266	1739	1095	276

Additional Controls by column: (1-8) state population, aged, and kids, include indgov, demgov, exp_lim, supmaj. Lags of the dependent variable are instrumented with two further lags in levels (AB) and two further lags in differences (BB), for the equation in differences and levels respectively. Column 8 uses again the GMM estimation but with instruments in differences. Standard errors in boldface are also significant with clustering by state. Overidentification tests: Sargan over-identification test is presented for the AB columns, valid under the first specification (homoskedastic error structure); Hansen J-test for the BB. Arellano and Bond (1991) autocorrelation tests for the error component shown under AR-p label. Only observations with line item veto, override requirements of two-thirds, and no supermajority requirements for a tax increase are included in the LIV-States regressions. Number of groups 47. The states of Nebraska, Alaska, and Hawaii are excluded in all regressions. The data set goes from 1960 to 1998. The symbol * is significant at 10%; ** significant at 5%; *** significant at 1%.

3.3 Budgetary Separation of Powers and Fiscal Institutions

There is a well established literature sustaining that fiscal institutions matter in explaining budgetary outcomes[7].When looked from a fiscal institutions/fiscal outcomes perspective, an interesting feature of the budgetary separation concept is that it approaches the budgetary process in order to explain fiscal outcomes in a very precise way. Even when the process is governed by rules, so far, line-item veto and formal separation of powers, these rules do not work in the vacuum, they interact with organizations, parties, and their relative strength in government.[8] The relative strength is crucial to determine whether a group can raise taxes and (residual) claim the proceeds.

How do fiscal institutions interact with our notion of budgetary separation in shaping expenditure and revenues sizes? We consider two examples of fiscal institutions, discuss their potential effects in light of the budgetary separation concept, and then look at the evidence.

Tax and expenditure limitations Consider, first, the case of a fiscal rule imposing direct constraints on outcomes, as that of *tax and expenditure limitations*. Even

---

[7]See for example Poterba and von Hagen (2000), for a compilation of different papers in the field.

[8]One limitation is that it only accounts for the incentives faced at the formation stage. Even though the incentives and forces governing the implementation and ex-post controls are beyond the scope of this agenda, we include a brief discussion in sub-section 5.1, related to revenue and expenditure composition.

though the budgetary separation approach does not directly encompass limitations on outcomes, it does point to a number of issues that must be taken into account when evaluating such rules. It suggests, for example, the following question: What would happen in a state with an aligned government and a formal rule with tax and expenditure limitations? Either the rule is non-binding, in the sense that the caps imposed by the rule are so high to render it futile,[9] or it generates a conflict between the forces at play, and the formal rule.

Using a soft budget constraint analogy, if groups anticipate that the rule will be enforced and sanctions applied in case of violations, then we could expect the rule to work in keeping the size of government bellow the cap. However, in an aligned government incentives to enforce the rule and punish violations are weak as, by definition, implementation and ex-post control are in hand of the same group.

Empirically, we find that states with formal ceilings on taxes and expenditures are, in short, self-selected as the estimates are positive and significant in most specifications. Formal fiscal rules stipulating tax and expenditure limitations have a positive and significant effect on the average tax rates in all regressions in Table 1 (estimates not shown), and are positive but less robust in Table 2.

Supermajority requirements As a second empirical example of the interaction between fiscal institutions and budgetary separation of powers, consider the case of a supermajority requirement. If a state has a supermajority requirement of two-thirds

[9]This seems to be the case for most states with tax/expenditure limitation rules (see Besley and Case (2003)).

of the House and Senate to approve a tax increase, our model tell us that the we should expect a higher tax rate only if the governor's party has enough votes to pass a tax increase[10]. More precisely, when a state has a supermajority requirement of $\upsilon > \frac{1}{2}$, is not no longer enough for a party (group) to control $\frac{1}{2}$ legislators in order to raise taxes and residual-claim the proceeds; it takes at least υ to do so, making it more difficult to raise taxes along the budgetary process.

Under this interpretation, the effect of a supermajority requirement is not direct but indirect through the broadening on the range of budgetary separation. This is an interesting and distinctive interpretation brought forth by our theoretical framework. The effect of a supermajority requirement for tax increases it to broaden the range of values of governor's strength in the legislature that activate budgetary separation. Conversely, it shrinks the cases in which a government can be considered as aligned from a budgetary perspective.

Based on the foregoing discussion, we modify our original variable to account for the effect of supermajority requirements on the budgetary separation cut-off value. We name the new variable *LIVdivided_SMaj,* which is the same as *LIVdivided* but for states/years in which supermajority requirement is present; in these cases, we fix the new corresponding cut-off level—e.g., 2/3 for a state with a 2/3 requirement—and redefine our divided government variable correspondingly, before interacting it with the line item veto dummy.

[10]See Knight (2000)Knight (2000b) for the analysis of supermajority requirement as treatment itself.

Table 3.2: Supermajority Requirement and Budgetary Separation of Powers

Explanatory Variables	All States (1)	All States (2)	LIV States (3)	All States (4)	LIV States (5)	All States (6)	LIV States (7)
LIVdivided_smaj	-0.18	-0.04		-0.06		-0.07	
	(0.08)**	(0.03)		(0.04)		(0.03)**	
Divided_smaj	0.04	-0.01	-0.05	0.00	-0.06	0.02	-0.06
	(0.07)	(0.03)	**(0.02)*****	(0.05)	**(0.03)****	(0.02)	**(0.02)*****
LIV	-0.04	0.01		-0.09		0.02	
	(0.15)	(0.06)		(0.15)		(0.06)	
supmaj	-0.38	0.08	0.08	-0.19	-0.18	-0.19	-0.23
	(0.06)***	**(0.03)****	(0.03)**	(0.10)*	(0.10)*	(0.03)***	(0.03)
Tax/exp. lim.	0.15	0.04	0.05	0.02	0.02	0.01	0.03
	(0.03)***	**(0.02)****	**(0.02)*****	(0.04)	(0.05)	(0.02)	(0.05)
Lags Dep. Var.	0	1	1	2	2	1	1
Est. Method	OLS	OLS	OLS	AB	AB	BB	BB
Overid Test (p)	.	.	.	0.11	0.04!	0.26	0.75
Res. AR(1) (p)	0.00	0.10	0.27	0.00	0.00	0.00	
Res. AR(2) (p)	0.00	0.00	0.00	0.79	0.98	0.12	0.15
# Observations	1833	1786	1598	1692	1457	1739	1275

Controls not shown in the table by column: (1-7) state population, aged, kids, demgov, indgov, restrict, cycle_trend. The results are robust to the inclusion of federal grants, percentage of blacks in the state population, and the previous notion of divided government. Lags of the dependent variable are instrumented with two further lags in levels (AB), and in differences (BB), for the equation in differences and levels respectively. Standard errors in boldface are also significant with clustering by state (Huber-White). Overidentification tests: Sargan over-identification test is presented for the AB columns, valid under the first specification (homoskedastic error structure); Hansen J-test for the BB. Arellano and Bond (1991) autocorrelation tests for the error component shown under AR-p label. Only observations with line item veto, override requirements of two-thirds, and no supermajority requirements for a tax increase are included in the LIV-States regressions. Number of groups 47. The states of Nebraska, Alaska, and Hawaii are excluded in all regressions. The data set goes from 1960 to 1998. The symbol * is significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2 shows the results with the new variables. Columns 1 to 3 show results with fixed effects, the first column being a static specification; AB and BB follow in columns 4-5 and 6-7, respectively. Overall, estimates are very similar to those of *LIVdivided* in Table 1 but with some differences. First, it picks the differences between divided and budgetary separation, as it is robustly significant including the whole sample. Estimates vary even less, in all specifications, and are more precisely estimated. SupMaj has a slightly lower less robust effect on taxes, but it is still significant in some specifications and sizeable. Obviously, we must control for endogeneity problems in SupMaj as well, since our variable of interest is now by construction correlated with SupMaj, which may itself be endogenous. We do so in the next section.[11] However, it is noticeable the improvement in overidentification tests in all GMM regressions, compared with those in Table 1.

As opposed to the results shown in Table 1, Ferrero and Magalhães (2007), the modified budgetary separation treatment has an important (close to the RD, static, point estimate) effect in the OLS regressions in levels (column 1, Table 2). Note also that *divided* is no longer significant as in the dynamic specifications, and always positive. In the following alternative specifications, this result does not change, nor does it for the BV sub-sample; we can safely omit it from the table to save space—all

[11]It is important to note that we ran multiple instrumental variables specifications in most cases, including supermajority requirement in a static set up. To allow for a straightforward relation with the related literature, we followed a similar strategy as in Knight (2000) to choose additional instruments for SupMaj. The variable was instrumented with two dummy variables: one for whether a state permits citizens's initiatives (Initiative), and the other for states that permit initiatives for constitutional amendments (Conammend). In contrast to Knight's result, we found that SupMaj had no direct effect on the average tax level, once its effect on budgetary separation was taken into account.

results are available upon request of course.

Again, we present only a restricted set of regressions outputs. The results shown in Table 2 are robust to the exclusion of controls and the addition others—mainly, federal grants and percentage of black population. The results are also robust to changes in the vector of instruments for the autoregressive component, with further lags of instruments going from 1, or 2, to 5, although overidentification and autocorrelation tests are not always valid. They are also robust to a distributive lags specifications, but lagged treatments are never significant in all specifications, once the contemporaneous treatment is included. The inclusion of the second lag in AB columns is a sine qua non condition to get AR tests on the error components right.

Another distinctive feature is that supermajority requirements still have an independent negative effect. This result is not expected, in the sense that we expect these requirements to work in the way described above, along the budget process, and not to have an effect *per se*. However, it is less robust and less precisely estimated, particularly in GMM regressions. More interestingly, *supmaj* is no longer significant when potential sources of endogeneity are taken into account. This suggests that the effect of supermajority requirements for tax increases works through its effect on the cut-off value for budgetary separation.

The autoregressive components on the average tax rates repeat the same patterns as those present in Table 1. OLS estimates are high, with the long run multiplier around 5, while GMMs estimates are near .5, and multiplier around 2. AB estimates of the autocorrelation component are the lowest; the inclusion of AR(2) is again significant and a sine qua non for the overidentification tests to go through. The lower value of the autocorrelation component $\sum \rho_j$, suggests that the weak instrument problem for the differenced equation is not a problem. AB is the most robust to changes in the set of instruments and controls, both in terms of the treatment estimates and of those of the autoregressive components (which in turn pin down the long run treatment effect). Sargan and Hansen $J-$tests in the AB specification improve substantially.

In this subsection, we have shown clear evidence that a corrected measure for budgetary separation of powers has a significant and sizeable effect on state average tax level. It points in the same direction that our previous results, while enhancing the validity of our model. It suggests that institutions do not work in the vacuum but their workings depend on the complex of institutions governing interactions and actual forces at play.

Of course, a few caveats are in order. The intensity of the treatment may not be the same at $\upsilon = \frac{1}{2}$ and at $\upsilon = 2/3$. Plus, in the full sample it remains the noise generated by states with $\frac{1}{2}$ override requirements. The restricting the sample still provides with a rude way to isolate better the budgetary separation effect. It is important bear in mind then that under the restricted sample LIV States, divided, LIVdivided, and LIVdivided_Smaj, coincide.

3.4 Voters' Preferences, Political Competition, and Weak Exogeneity

Before moving to a discussion on variable treatment effects, we stop to revise our previous results and their identifying assumptions. So far, we have assumed that, conditional on the set of controls, the different measures of budgetary separation are strictly exogenous—orthogonal to contemporaneous and past realizations of the error component in (3.2). It is commonly argued in the literature that preferences for public expenditures, correlated with preferences for political competition, is a key source of endogeneity in this kind of regressions. This is also one main concern, for the restricted sample results: the possibility of omitted variables that are correlated both with the tax level and divided would imply a more severe bias as our treatment and divided overlap. In order to make claims of a causal effect between a divided government and a lower tax rate in states with line item veto we move on to two strategies. First we present two variable that may proxy for the omitted variable political preferences. We then include voters' turnout in State elections, and finally, we treat our treatment variables as predetermined and endogenous, and instrument them with their own lags.

A starting point to address the potential endogeneity of our treatment variable due to voter preferences for taxes and political competition is to find proxies for those preferences. We propose two proxies that are appropriate for our study provided by Ansolabehere and Snyder (2002). They have collected election results for a number of directly elected state offices other than governorship; they include Attorney General, Lieutenant Governor, Secretary of the State, etc. We include two proxies: a dummy for unopposed elections in lower office elections ($Unoposed$), and another non-partisan measure of political competition in lower level elections: $\mathbf{p}_{st} = -abs\left(DemLowOff_{st} - 0.5\right)$ with $DemLowOff$ as the share of votes for the democrat candidate in lower level elections. Both variables are non partisan measures of political competition in the state lower level elections. An additional important control for voters' direct influence on the average tax level is the rate of voters turnout in each state election.

Netting out the effect of preferences for political competition, and turnout, from the treatment variable, that may also affect directly preferences for the average tax level, we expect to find a higher effect in absolute terms, however, as they are positively correlated with $LIVdivided$.[12]

[12] We expect our estimates in Tables 1 and 2 to be downward biased in absolute terms. Let the variables with tildes denote the residuals from its linear projection on a vector of observables $\mathbf{x} \in \mathbf{R}^K$, which includes state-time fixed effects as well as lagged dependent variables. We denote variables with tilde, those once the effect of x's has been partialed out, formally, for any variable y, $\widetilde{\mathbf{y}} = \mathbf{M_X y} = \left(\mathbf{I} - \mathbf{P_X}\right)\mathbf{y} = \left(\mathbf{I} - \mathbf{X}\left(\mathbf{X'X}\right)\mathbf{X}\right)\mathbf{y}$.

Thus, we can re-express the structural equation in (3.2) as

$$\widetilde{\tau}_{st} = \delta\widetilde{Div}_{st} + \widetilde{u}_{st},$$

which implies that the plim $\hat{\delta}_{OLS} = \frac{\mathrm{Cov}\left(\widetilde{Div},\widetilde{\tau}\right)}{\mathrm{Var}\left(\widetilde{Div}\right)} = \delta + \frac{\mathrm{Cov}\left(\widetilde{Div},\widetilde{u}\right)}{\mathrm{Var}\left(\widetilde{Div}\right)}$, where the variance and covariance terms refer to population moments. If pclow and turnout have a negative effect on taxes, but positively correlated with our treatment variables, as expected and as it is the case, it implies that $\frac{\mathrm{Cov}\left(\widetilde{Div},\widetilde{u}\right)}{\mathrm{Var}\left(\widetilde{Div}\right)} > 0$, reducing the point estimate for the treatment in absolute terms.

Another concern is that even when doing our best controlling for observables, our treatment may be not strictly exogenous, but correlated with past or simultaneous unobservables present in the error term. To take on this issue, we take advantage of the flexibility and possibilities offered by the AB and BB, Generalized Method of Moments, instrumenting procedures.

Results are again as expected and robust. In **Table 3**, we present results only when treatment variables are taken as predetermined or endogenous—only due to space considerations, as results are not affected in all specifications. All specifications include divided or divided_smaj, as standard controls. In columns 1-2, with the whole sample, and in columns 3-4 with the restricted one, we treat either LIVdivided or LIVdivided_Smaj, as predetermined. The sample in column 4 and 6, differ from that of 3 and 5, as States with supermajority requirements are included. In column 5 and 6, also voters turnout is treated as predetermined.

As expected, point estimates are either slightly higher or unaffected by the introduction of proxies for political competition and turnout. The autoregressive components, not shown, reduced their variability substantially, summing up to .5, with a multiplier oscillating between 2 and 2.5, in the GMM specifications. Overidentification and autocorrelation tests perform well in all specifications. The same holds with other specifications of the vector of instruments, from 2 to 5 lags also for the predetermined treatments. When the *LIVdivided* and *divided* are treated as strictly exogenous, conditional on the additional controls for voters preferences, the only important difference with the results shown above is that the overidentification tests are weaker although the null is always rejected.

Note that when the dynamics is accounted for, proxies for voters preferences and voters turnout are not significant. This is not the case in a static specification, especially for voters' turnout rates. Another point to make is that supmaj is not significant in most regressions.

Table 3.3: Budgetary Separation of Powers, Political Competition and Preferences

Explanatory Variables	All States (1)	All States (2)	LIV States (3)	LIV States† (4)	LIV States (5)	LIV States† (6)
LIVdivided_smaj		-0.12		-0.06		-0.07
		(0.06)**		(0.03)**		**(0.03)****
LIVdivided	-0.10		-0.05		-0.06	
	(0.05)**		(0.03)*		(0.03)**	
supmaj	-0.17	0.17		-0.14		-0.12
	(0.09)**	**(0.09)****		(0.09)		(0.09)
pclow	-0.23	0.33	0.21	0.26	-0.33	-0.19
	(0.21)	(0.21)	(0.29)	(0.27)	(0.28)	(0.26)
unopposed	0.01	0.01	0.02	-0.02	0.03	0.02
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
turnout	-0.17	-0.14	-0.21	-0.17	-0.67	-0.56
	(0.20)	(0.20)	(0.29)	(0.24)	**(0.33)****	(0.25)**
Lags Dep. Var.	2	2	2	2	2	2
Est. Method	AB	AB	AB	AB	BB	BB
Overid Test (p)	0.89	0.86	0.67	0.21	0.97	0.80
Res. AR(1) (p)	0.00	0.00	0.00	0.00	0.00	0.00
Res. AR(2) (p)	0.96	0.96	0.34	0.46	0.82	0.62
# Observations	1565	1565	1014	1194	1014	1194

Additional Controls by column: (1-8) state population, aged, and kids; (4)-(8) include indgov, demgov, exp_lim, supmaj, grants, pbl. Lags of the dependent variable are instrumenting with two further lags in levels (AB). Column (7) treats LIVdivided as predetermined, and is instrumented with 1, 2, and 3, lags with robust results. Column 8 uses again the GMM estimation but with instruments in differences. Overidentification tests: from Column 1-7, Sargan over-identification test is presented, valid under the first specification (homoskedastic error structure); column 8 presents Hansen J-test. Second brackets are robust to heteroskedasticity and intra-state arbitrary autocorrelation structure. Control always include divided, LIV, and dividedsmaj, when using the full sample.The sample in column 4 and 6 (†), differ from that of 3 and 5, as States with supermajority requirements are included.

Comparing results, for a given methodology, estimates under the modified treatment effect are again slightly higher in absolute terms, and more precisely estimated. It is important to stress that columns with the restricted sample are not comparable, however. With the modified treatment, in the restricted sample column, there are states with and without supermajority requirements; it is only restricted to those states with line-item veto and two-thirds override requirements, whereas for the standard treatment on top of that there are no states with supermajority requirements.

Finally, we can also interpret the relative strength of the modified treatment as an additional support of the notion of budgetary separation. This is so since we depart a bit more from factors such as political competition and divided governments.

Note that there are no statistical reasons, to exclude the possibility of using pclow and unopposed as instruments for LIVdivided with the restricted sample—as they are basically political competition proxies not correlated with average taxes directly in all specifications. In first stage regressions, they do perform very well, and so do second stage, always in statistical terms—not shown. The identifying assumption would be that preferences for political competition only affect the tax level through those that can actually modify it, meaning through changes in parties' strength in the legislature, or the party identity of the governor. Results in a dynamic set up are robust, and better for the LIV States subsample. This is a natural result since the instruments are mostly proxies for political competition. It is also the case that estimates are more precisely estimated for the subsample excluding the 90's;

again this is justifiable as some states have introduced direct mechanisms, as popular referenda, to approve tax increases during the 90's. Additionally, once you control for the modified treatment, using the same instruments for supermajority requirements as those used by Knight (2000), we find that supermajority requirements have no direct effect, only through the extension of the cutoff value for budgetary separation. This interesting result is in line with those in column (4) and (6) of Table 3, where supermajority requirement is no longer significant when states with no line-item veto are excluded from the sample.

Overall, voters preferences over political competition, and voters turnout, have no effect on the average tax levels directly. This is also expected within this framework, as the budget preparation, approval, and implementation, is delegated by voters to governments, guided by clear incentives and constraints, beyond those of individual voters.

3.5 Extensions: Variable Effects, Budget Composition, and Parties

We extend the foregoing framework to allow for variable treatment effects of budgetary separation of powers. We first take on the analysis of budget composition—instead of size. In a broader interpretation of the model developed in Ferrero and Magalhães (2007), the governor's incentive to line-item veto spending programs should be higher the easier they are targeted to groups outside her constituency. On the other hand, general spending programs should not be as sensible to budgetary separation as specific transfers.

Next, we explore the variable effects of differential parties' characteristics: party preferences for overall spending and taxation, and variable party alignment. Another way of looking at the modelling assumptions in Ferrero and Magalhães (2007), is that parties are perfectly aligned, and that the constituent groups for legislators and for the governor perfectly overlap when they share party affiliation. There is no within party conflict. Although locally valid, this is a stringent assumption. More generally, as target constituencies diverge between legislators and the executive of the same party affiliation, within party conflict in terms of desired budgetary composition rises. Thus, within party conflict provides with an active role for line-item veto even when governments are, party-wise, aligned. Finally, it is well known that the two traditional parties in the American States, have well differentiated preferences for general public expenditures; we expect that shifts to budgetary separation should imply a differential effect depending on the political identity of the governor. Overall, we find preliminary evidence supporting our priors.

3.5.1 Variable Treatment Effects 1: Expenditure and Revenue Composition

Suppose that instead of having—easy identifiable—group specific transfers, as in Ferrero and Magalhães (2007), we have that programs vary in the extent that they can be targeted to a specific group. More formally, we still have a general size of the budget $g = \sum f^i$. However, group utility takes a smooth form on public expenditure of the following form: $H\left(f^j + \sum_{i \neq j} \alpha^i f^i\right)$, where $\alpha^i \in [0, 1]$. The different programs may have a different α parameter over that interval; the case presented in the baseline model is that of $\alpha^i = 0$, for all i, that is, group i only benefits from specific programs targeted to the group. This would be one of the extreme cases, while the other is that of $\alpha^i = 1$ for all i, that is only pure public goods enter the budget.

The group not controlling agenda setting powers in the legislature, but the line item veto prerogative, will be more permissive with broad based programs. In a nutshell, we do not expect all the items of the budget to be affected homogeneously, when a government switches to budgetary separation. More general expenditure programs, with higher α, should be less affected by line-item veto, whereas specific transfers should me more responsive to the treatment.

In order to test these predictions, data requirements are more stringent. Not only do we need detailed information on the budget broken down by programs, but also to be able to classify them according to hard it is to target them to a specific constituency. A more pragmatic approach is to take single items of the budget, and check how responsive they are to changes in government composition in the presence of line item veto. This approach could be useful also to retrieve constituencies/groups that are favored, for example, by one party or governor—see next subsection.

Based the data available, we generate a variable that captures, with limitations, expenditure composition. First, we construct a variable summing up transfers to specific activities, such as, agriculture, games, fish, mining, and the like; then we construct another variable simply dividing these total specific transfers by total state expenditures (specif_totalexp). Additionally, we construct in the same way a general expenditure variable, including spending programs such as in parks, recreation, forestry, and the like. And we use them as the outcome variable in the very same specifications used above. We always include federal grants, however, as there might be federal earmarked transfers to specific programs. **Table 4** shows the results only for the restrictive sample. The predicted behavior is supported empirically: the effect of budgetary separation is negative and significant for specific transfers, and neutral for general transfers.

Even though results in Table 4 are robust, we take the evidence only as an illustration of an important potential application. Of course, the variables constructed can be subject to criticism, as they are not comprehensive enough. The excuse is that of data limitations. Another caveat is that programs may have implicit different partisan preferences.

Table 3.4: Expenditure Composition

Explanatory Variables	specif_totalexp (1)	specif_totalexp (2)	specif_totalexp (3)	specif_totalexp (4)	gral_totalexp (5)	gral_totalexp (6)
LIVdivided	-0.05	-0.07	-0.05	-0.07	-0.03	-0.03
	(0.02)**	(0.03)***	(0.02)***	(0.02)***	(0.05)	(0.05)
Fed. Grants		0.09		0.01		-0.05
		(0.02)***		(0.02)		(0.06)
Lags	0	1	1	1	1	1
Sample	LIV States	LIV States$^{(+)}$	LIV States$^{(-)}$	LIV States$^{(+-)}$	LIV States$^{(-)}$	LIV States$^{(+-)}$
Est. Method	OLS	OLS	AB	OLS	AB	AB
Overid Test (p)	.	.	0.18	0.21	0.01!	0.01
Res. AR(1) (p)	0.00	0.00	0.00	0.00	0.00	0.00
Res. AR(2) (p)	0.00	0.00	0.61	0.38	0.59	0.74
# Observations	1047	990	992	881	992	881

Additional Controls by column: (1-8) state population, aged, and kids; (4)-(8) include indgov, demgov, exp_lim, supmaj, grants, pbl. Lags of the dependent variable are instrumenting with two further lags in levels (AB). Column (7) treats LIVdivided as predetermined, and is instrumented with 1, 2, and 3, lags with robust results. Column 8 uses again the GMM estimation but with instruments in differences. Overidentification tests: from Column 1-7, Sargan over-identification test is presented, valid under the first specification (homoskedastic error structure); column 8 presents Hansen J-test. Second brackets are robust to heteroskedasticity and intra-state arbitrary autocorrelation structure. Control always include divided, LIV, and dividedsmaj, when using the full sample.The sample in column 4 and 6 (†), differ from that of 3 and 5, as States with supermajority requirements are included. Sample (+) implies that there are missing values for a variable (grant). (-)observations lost with instrumenting predetermined or endogenous variables with lags. Specif_totalexp is constructed summing over the programs to specific activities, available, and multiplied by one hundred, this variable in real terms in then divided by total expenditures.

3.5.2 Variable Treatment Effects 2: Partisan Preferences, and Party Alignment

It is a well know feature of the American politics that republicans tend to prefer lower overall taxation. One explanation could be that lower taxation, may have a net positive redistributive effects for its constituent groups; or simply that partisan preferences differ. Formally, this can be captured adding a preference parameter $\alpha^p H(g)$, with $p = R, D$, and $\alpha^D > \alpha^R$.

In the context of budgetary separation this can be easily tackled interacting the treatment variable with the party identity of the governor. Letting ι^p, be an indicator variable taking value one when the governor belongs to party p, and zero otherwise, $\iota^p \cdot LIVdivided$ would pick the effect of a divided government when the governor, with party affiliation p has line-item veto.

Another extension refers to variable party alignment. In the previous work, the local interpretation of the model and of the core empirical strategy, allowed us to treat parties as unified groups with perfectly aligned interests and constituencies. One can argue that parties are not perfectly aligned groups. It may be the case that some legislators, of the same party identity as that of the governor, fiercely represent the interests of groups in his district of origin, which in turn may not be of interest for governor—for example, low relative voters mobility in Dixit and Londregan (1996) wording. But as long as the overlapping of constituency groups between the governor and legislators of the same party is higher than under that of two with different party affiliation—much weaker and plausible assumption—, the effect of budgetary separation will still hold as the governor will (line-item) veto more programs when the legislature is dominated by the opposition. However the effect would vary. It should be higher the higher the party alignment or, equivalently, the lower the conflict —constituency wise— within party. Within party conflict would provide line-item veto with an active role even when the Legislative and the Executive are aligned.

The empirical problem is how to proxy for party alignment. In a different context—credibility of policy platforms for governor elections—, Grossman and Helpman (2005) argue that the higher the share of independents in the voting population, the higher the within party conflict as target constituencies diverge.[13] By analogy, in our case, the higher the share of independent voters, the lower party alignment, and the strength of the budgetary separation of powers. As legislators effort to target programs to their own constituency, there is room for active line-item veto even when the Legislature is controlled by the same party as the party identity of the governor. (We sketch a formal discussion in the appendix, adapting Dixit and Londregan (1996), Persson et al. (2000), and Grossman and Helpman (2005) to motivate this idea.) The empirical implication is clear: a higher share of independent voters weakens party strength, and thus switches on and off budgetary separation will have

[13] In their case, this conflict is anticipated by voters, and the governor candidate when he sets his platform. Still, optimal platforms diverge both from the candidate's bliss policy (determined by her constituencies' preferences), but also from the ex-post implemented budget if she wins the election. Credibility is an issue, since the lower party alignment, the less credible platforms are.

a lower effect on the equilibrium tax level.

We use CBS/New York Times national polls on ideology party identification to proxy for independent voters. Unfortunately, the surveys cover a limited time span, 1976-2003, but they are available for all states.[14] Part of the survey focuses on ideological issues, and classify respondents intro one of three groups: liberals, moderates, or conservatives. We use the share of moderates, less ideological, voters as a proxy for voters responsiveness (mobility) to policy.

The estimating equation is:

$$\tau = \mathbf{x}\beta + \delta_1 treat + \delta_2 treat \cdot moderates + \delta_3 + u \tag{3.3}$$

and we expect $\delta_2 > \delta_1$. So, under budgetary separation we have that on average taxes are $\delta_1 + \delta_2 \cdot indip$, $\delta_2 > 0$, which can be thought of as a inverse proxy for party alignment, or a proxy for within party budgetary conflict.

The interpretation is the following: consider first the extremes, if all voters are independent, mobile and policy motivated, then each legislator is going to try to shift resources to his district as his chances for re-election are not secured—he must be more attached to his local constituents rather than adhere to the party line; on the other extreme, suppose all voters are partisan, groups are then perfectly

[14] CBS News/ New York Times interviews voters as they leave the polls. It asks questions on partisanship, and a number of questions on attitudes toward ideological issues, such as abortion, death penalty, religion, and the like. Based on it, they classify voters according to their partisan affiliation in democrats, independents, and republicans, and according to the ideology in liberal, moderate, and conservative. The data is downloadable from http://mypage.iu.edu/˜wright1/. — Gerald C. Wright webpage, Department of Political Science, Indiana University

aligned and well differentiated according to their party affiliation, with no intra party conflict. The interpretation of the pure LIVdivided effect now changes: it is the upper bound of the budgetary separation effect when groups controlling the Legislature are perfectly aligned those of the Executive. This happens when voters are fully aligned to a party line, and there are no independent voters in the population. The overall effect of budgetary separation coincides with that of LIVdivided, and equals $\delta_1 < 0$. On the other extreme, the effect of budgetary separation is $\delta_1 + \delta_2$, which should be neutral as groups do not intersect, even within a party. As for δ_3 we have no

theoretic prior. (As we omitted this component in the foregoing estimations, we expect to find a $\widehat{\delta}_1 < \widehat{\delta}_I$, with the overall effect for the average share of independent voters, $\delta_1 + \delta_2 \cdot \overline{indip}$, around the same value.)

We show a few results both on the effects of variable party preferences, and party alignment, on budgetary separation in **Table 5**. As can be seen in the last row, the sample size drops substantially due to the survey coverage. We keep the number of states in 47, but the time span goes from 1976-1998. We use the same specifications as in the previous estimations. In columns 1-4 we find that party preferences seem to matter. The budgetary separation effect is much higher under a republican governorship. The point estimates at least doubles the one under a Democrat's. However, point estimates vary substantially, and the democrat budgetary separation effect is not significant. Still, the pattern systematic in that partisan preferences seem to matter, and that the average effect under R, is higher than the overall average, which in turn is higher than the average effect under D.

In columns 5-7, we present the results for (3.3). Again results are as expected. A higher point estimate for the average effect, which is robust around -.4, and a positive effect for the within party conflict parameter δ_2. The average share of moderate voters is 40%, which yields 0.30-0.32 as a within party conflict effect under budgetary separation. The overall effect of budgetary separation for a state with average share of moderates is thus,

$$\widehat{\delta_1} + \widehat{\delta_2} \cdot 0.4 \in (-.10, -.08),$$

which is very close to are previous estimates, in this dynamic specification. This is again on impact. The long run multiplier, with predetermined treatments as those shown, are again slightly above two.

The variables used can be subject to criticism. The validity of our proxy, for example, is constrained to the distribution across districts of moderate voters, if we interpret the within party conflict in terms of geography. As in the previous section, and even though results in Table 5 are robust, we take the evidence as another example of important potential applications of the framework developed in Ferrero and Magalhães (2007). The above approaches may be interacted, for example, by using both partisan preferences and programs to identify overlapping and disjunct constituent groups between parties.

Table 3.5: Partisan Voters, Party Alignment and Heterogenous Effects

Explanatory Variables	ttax_gdpp (1)	ttax_gdpp (2)	ttax_gdpp (3)	ttax_gdpp (4)	ttax_gdpp (5)	ttax_gdpp (6)	ttax_gdpp (7)
budsep_D	-0.04	-0.02	-0.05	-0.02			
	(0.06)	(0.06)	(0.06)	(0.07)			
budsep_R	-0.08	-0.11	-010	-0.13			
	(0.05)*	(0.05)**	(0.05)**	(0.05)***			
LIVdivided					-0.40	-0.39	-0.40
					(0.20)**	(0.20)*	(0.22)*
budsep_moderates					0.82	0.78	0.79
					(0.43)**	(0.43)*	(0.46)*
moderates%					-0.01	0.01	0.01
					(0.01)	(0.01)	(0.01)
Lags	2	2	2	2	2	2	2
Sample	LIV States	LIV States$^{(+)}$	LIV States$^{(++)}$	LIV States$^{(++)(+)}$			
Est. Method	AB/pred	AB/pred	AB/pred	AB/pred	AB/pred	AB/pred	AB/pred
Overid Test (p)	0.66	0.86	0.67	0.85	0.20	0.16	0.20
Res. AR(1) (p)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Res. AR(2) (p)	0.74	0.85	0.66	0.76	0.53	0.87	0.54
# Observations	532	532	532	532	532	532	532

Additional Controls by column: (1-8) state population, aged, and kids; (4)-(8) include indgov, demgov, exp_lim, supmaj. Lags of the dependent variable are instrumented with two further lags in levels (AB). LIVdivided as predetermined, and is instrumented with 1, 2, and 3, lags with robust results. Overidentification tests: from Column 1-7, Sargan over-identification test is presented, valid under the first specification (homoskedastic error structure). Sample (+) allows for simultaneous correlation (endogeneity), (-) observations lost with instrumenting predetermined or endogenous variables with lags. (++) includes additional controls as robustness checks: percentage of black, partisan, turnout.

3.6 Summary and Discussion

We introduced a dynamic approach to studying the effects of budgetary separation, and we took on an extensive look at the effects of budgetary separation of power as interacted with features of fiscal institutions and politics. The introduction of dynamic framework not only allow us to control for potential feedbacks to voting decision, but also to explore the time series properties of the variables of interest—serving as robustness checks of our previous results. We presented the alternative estimation methods in this set up, conditional on the time span available. The result that budgetary separation of powers matters, negatively affecting the average tax level, goes through.

The interaction of fiscal and political institutions, together with relevant organizations and their characteristics, seems to be promising. In this case, we do find that the notion of budgetary separation of powers has important, additional implications, when interacted and in interpreting politics and other fiscal policy institutions. In terms of fiscal institutions, the use of rules that set constraints on outcomes, may generate a tension along the budgetary process, between the budget preparation and implementation stage, and that of enforcement and ex-post control. Empirically, we find that states with formal ceilings on taxes and expenditures are, in short, self-selected as the estimates are positive and significant in most specifications. Instead, fiscal rules affecting the costliness of tax raises, seem to be more effective. This is the case of supermajority requirements for tax increases, as it broadens the extent of budgetary separation.

Another important finding is that voters preferences over political competition, and voters turnout, have no effect on the average tax levels directly. This is also expected within this framework, as the budget preparation, approval, and implementation, is delegated by voters to governments, guided by clear incentives and constraints.

We also extended the framework to the analysis of budget composition—instead of size. As expected, we find that general spending programs are not be as sensible to budgetary separation as specific transfers. The interaction of this finding together with that of party heterogeneity, may be applied to a wide range of issues. For example, to trace back target groups, lobbies, and program that are easy to target to a specific constituency. The variable party alignment and its effect on budgetary separation relates to a new research agenda in political economics. The precise mechanism through which variable party alignment works in shaping budgetary outcomes, is again clearly addressed by the budgetary separation framework. Evidence, overall, point to the expected directions in all cases.

Chapter 4

Spatial Analysis and Endogenous Information Acquisition: a Structural Model of Voting and Participation in the US Presidential Elections

4.1 Introduction

Spatial models assume individuals may be placed on a metric together with politicians. Individuals should vote for the candidate closest to them. If both candidates are at the same distance(indifference) or if the preferred candidate is too far away(alienation), individuals may prefer to abstain.

Poole and Rosenthal (1984) brought this model to the data using the American National Election Studies(ANES) from various presidential elections. It was successful in depicting movements in the polarization of the electorate and in explaining individual preferences for the closest candidate. It was not able to account for abstentions. We estimate a similar spatial model in which abstentions only arise due to uncertainty about the location of candidates. This modification improves the capacity of the model to predict abstentions.

Voters and candidates were represented in a metric retrieved by least-squares unfolding techniques. This procedure was based on the Thermometer Felling questions asked in the ANES surveys about the presidential candidates and around ten other well known political figures. Individuals are asked to grade each political figure(stimuli) with a grade from 0 to 100, where 50 means neutral. The intuition of the algorithm they use is that 'for each candidate-respondent pair, one can define prediction error as the difference between the distance and the normalized quantity $(100 - T)/50$, where T is the original thermometer rating. The coordinates are then chosen to minimize the sum of squares of these error'.

The utility of an individual i of announcing candidate j is:

$$U_{ij} = \alpha_j + \beta * distance_{ij}^2 + \epsilon_{ij}, \tag{4.1}$$

where the distance between i and j is based on the retrieved metric. They estimate the parameters α_j and β with a logit model and produce the predicted probability of each possible action: either vote for the republican, democrat, or independent candidate(when there is one), or abstain.

One way to evaluate the model is to assign as the predicted choice the one with the highest estimated probability. The model consistently underpredicts abstention. The best result if for the 1972 election where the model predicts 50% of the total abstention rate but less than half of it is made of correct individual predictions. The spatial model seemed to fail in explaining abstentions.

The literature turned to other variables that correlate with turnout: individual characteristics such as age, education, income, partisanship, occupation, religion; election specific conditions such closeness of elections, campaign spending; or random aspects such as local weather condition. Matsusaka and Palda (1999) use three dozen of such explanatory variables in a logit model to estimate the individual predicted probability of voting. They conclude that even though most demographic and contextual variables are significant in the regression, random guessing would be more accurate in predicting whether someone goes out to vote or abstains.

Theoretically, different models have been forwarded to explain participation in elections[1]. In pivotal-voter models the motivation for voting is to affect the electoral outcome, that is, going out to vote is optimal if individuals have a non-negligible probability of being pivotal. The closest of an empirical application of this type of models in mass-elections is to check whether the closeness of the race may induce

[1]For a more detailed overview of turnout models in the political economy literature see Merlo (2005).

a higher turnout. Such a variable, however, does not improve the estimates of Matsusaka and Palda (1999).

Ethical-voter models are based on the notion of rule-utilitarianism. Individuals follow the voting rule that, if followed by everyone else in their group, would maximize the group's aggregate utility. The rule implies that within a group some individuals should vote and others should abstain. The model generates significant turnout in equilibrium independently of the probability of being pivotal. Coate and Conlin (2004) use the model to estimate group behavior with data on Texas liquor referenda.

The theory of turnout of most interest to this paper is the uncertain-voter model. As in the spatial model, individuals and candidates are assumed to have positions on a metric. Individuals prefer the closest candidate. If they know the position of candidates in the space they have no reason to abstain, they just choose the closest candidate. Uniformed individuals, who do not know the exact position of the candidates, may make a mistake by voting on a candidate that turns out not to be the closest. The cost of making that mistake may outweigh the benefits of voting.

Degan and Merlo (2006) estimate the model with individual level data from the National Election Survey of 2000. They take as given the metric and spatial positions of the candidates as those estimated by Poole and Rosenthal's NOMINATE scores[2]. The place of each individual on that metric is retrieved from their revealed political choices(whether they participated in the elections and who they voted for)

[2]For a detailed description on how the politicians positions are estimated see Poole (2005); for a more intuitive introduction and political analysis with the estimated political space see Poole and Rosenthal (1997).

and demographic characteristics: age, income, education, religion, race, gender, and partisanship. From the same statistics they are also able to estimate the voters probability of being informed and whether they have high or low civic duty.

The model is able to replicate observed patterns of abstentions and goes further, reproducing split ticketing, and selective abstention (voting in the presidential election but not on the congressional election). Their estimation redeems the spatial model and its ability to account for abstentions.

In this paper I also use the combination of the spatial model with the uncertain-voter model to predict individual choices in presidential elections. Instead os relying on a revealed preferences argument and on demographic statistics I go back to Poole and Rosenthal (1984) and estimate the metric in which candidates and individuals are placed only from the Thermometer Feeling questions on political figures in the National Election Survey. To identify the probability of each individual of being informed I look at questions in the NES regarding exposure to information about the campaign and I show that political preferences and variables on information are sufficient statistics to estimate and predict voting and participation behavior in America Presidential Elections.

4.2 The Political Space

The position of each individual in the one-dimensional space is observed by the econometrician, together with the position of the two candidates. It is retrieved by a multidimensional unfolding algorithm, ALSCAL[3]. This type of procedure is commonly used in marketing to place costumers and products(stimuli) on a space, where products can be compared by how 'close' they are to each other. The stimuli in our case as in Poole and Rosenthal (1984) are the Feeling Thermometer questions asked in the ANES about the presidential candidates and other major political figures.

The algorithm may estimate either one or bi-dimensional space in which the candidates and individuals are placed . I apply the algorithm to the 2000 election and use the one-dimensional space. The estimate results in the candidates being located in the extreme of the space. Implying individuals see themselves as moderates compared to the candidates. This results are similar to those in Poole and Rosenthal (1984). The estimated space goes from -2.42 to 2.30. The average individuals is placed at 0.03 and the median at 0.02. Kerry is at the extreme left of the space, -2.40; and Bush at the extreme right 2.26.

4.3 Model

4.3.1 Announcement

Citizens face an election for president between a democrat and a republican candidate. The democrat candidate is drawn from a population of possible candidates with the distribution function $F(y^D|y^D < Y^m)$, where y^m is the individuals in the median

[3]For a detailed explanation of how the algorithm works see Takane and W. (1977). The version I use is for FORTRAN90 and is available at Forrest Young's ALSCAL webpage

location of the metric. The republican candidate is drawn from $F(y^R|y^m < Y^R)$. Thus we are capturing that democrats are expected to be on the left side of the political spectrum and republicans on the right.

Individuals announce for which candidate they intend to vote or whether they intend to abstain. This is observed as a pre-election question in the ANES, asked three months before the election itself. We modelled this choice as in the original spatial model in Poole and Rosenthal (1984), as in 4.1. The only difference is that at this stage all individuals are uninformed, their choice of candidate is determined by the expected position of each candidate, j. Out of a sample of 961 respondents 86 announce they will abstain but also announce a preferred candidate. Only 11 announce they will abstain and and do not report a preferred candidate.

4.3.2 Information Acquisition

Individuals may acquire information at a cost. The American National Election Studies asks whether respondents read about campaigns on magazines, newspapers, or hear about the campaign on TV. I allow a different cost for each of these three possibilities, f, λ_f. Acquiring information increases the probability of being informed about the exact location of the candidates. Candidates that have announced abstention and no preferences are assumed in the model not acquire information.

Reading about the campaign in magazines (294 individuals) or newspapers(522) is one way to be informed. Our full sample is 961 individuals. Getting information about the campaign in both indicates more information at a greater cost. Otherwise, for most people TV is the way to they get information about campaigns. Those that report having seen no information about the campaign are 91 and we assume they are uninformed with probability one. From these 91, 35 announced they would abstain.

The acquisition of information does not guarantee being informed. Two sets of questions asked on the ANES help us assess with what probability is an individual informed. The first set asks the respondent to place Kerry and Bush in a liberal conservative scale from 1 to 7. I create a indicator, $uninflibcon$, that takes value one if the individual places Bush as morel liberal than Kerry, 258 make such a mistake. ANES also asks the respondent to identify out of four possibilities the position of Bush and Kerry on abortion (Under no circumstances, in case of danger to mother's health, the need must be clearly established, the woman's choice). I create another indicator variable $uninfab$ to respondents who gave Bush a more pro-choice position than Kerry, 229 made such a mistake.

The probability of being informed is estimated for each type of information that may be acquired, s, magazines and newspapers(M), magazines or newspapers(N), and TV. I also allow for it to differ between those that announced a preference for republicans or for democrats (those that do not announce a preference (11 observations) are assumed not to acquire information):

$$\pi_{sj} = \frac{exp(\alpha_s + \beta_{1S} uninflibcon + \beta_{2S} uninfab)}{1 + exp(\alpha_s + \beta_{1S} uninflibcon + \beta_{2S} uninfab)}.$$

4.3.3 Voting

Informed voters observe the exact position of each candidate and vote for the closest one, the cost of voting for the informed is zero. Uniformed voters must calculate the cost of making a mistake, that is, of voting for a candidate that may turnout to be the farthest.

The cost of voting for the uniformed (CVU) is given by:

$$CVU_D = \int [u_i(y^D) - u_i(y^R)] dF_i^e(y^D, y^R)$$

where the integration is over all the possible combinations of candidate positions in which the individual would had been better off voting for the other candidate:

$$\{(y^D, y^R) \in Y \times Y : u_i(y^D) < u_i(y^R)\}.$$

In this last stage, individuals who had announced a certain candidate and stay uninformed will compare the choice between voting for the announced candidate and abstaining. If she votes for the announced candidates the pay-off is:

$$\delta_t - CVU_j,$$

where δ_j is the civic duty reward that is allowed to differ for those that announced Bush, Kerry, and abstention (b, k, a). If the prefers to abstain, the pay-off at this stage is zero.

Individuals who are informed learn the candidates position and their choice is between the two candidates. The cost of voting for the closest candidate is zero and the cost of voting for the farthest candidate may be normalized to one. Let's call it CVI_j The individual also receives δ_j.

The final pay-off of a individual i, announcing j, acquiring information f is given by:

$$U_{ijf} = \alpha_j + \beta * distance^2_{ij} - \lambda_f + \delta_j * I[1\ if\ voted] -$$
$$CVU_j * I[1\ if\ uninformed] - CVI_j * I[1\ if\ informed] + \epsilon_{ijf},$$

where ϵ_{ijf} is defined as measurement error in placing the individual in the one-dimensional political space.

4.3.4 Estimation

Following Amemiya (1985), I assume

$$F(\epsilon_1, \epsilon_2, ..., \epsilon_{44}) = exp\Big(- \sum_j b_j \Big\{ \sum_{f \in J} a_f \big(\sum_{v \in F_f} exp(-\rho_f^{-1} \epsilon_v) \big)^{\frac{\rho_f}{\sigma_j}} \Big\}^{\sigma_j} \Big),$$

where $v = (K, B, A)$, whether the individual voted for Kerry, Bush, or abstained. This assumption allow us to write the conditional probabilities of the different events, announcement, information acquisition, and voting as in a nested logit model. As an example let's take and individuals who announced and voted Bush, and got information from TV, we can write the probability of each of these events as:

$$P(B|b, TV) =$$

$$\pi_{TV}(x) \frac{exp\Big((\alpha_B + \beta * distance_{iB}^2 - \lambda_{TV} + \delta_B - CVI_B)/\rho_{TV}\Big)}{exp\Big((... - CVI_B)/\rho_{TV}\Big) + exp\Big((\alpha_B + \beta * distance_{iB}^2 - \lambda_{TV} + \delta_B - CVI_K)/\rho_{TV}\Big)} +$$

$$(1 - \pi_{TV}(x)) \frac{exp\Big((\alpha_B + \beta * distance_{iB}^2 - \lambda_{TV} + \delta_B - CVU_B)/\rho TV\Big)}{exp\Big((... - CVU_B)/\rho_{TV}\Big) + exp\Big((\alpha_B + \beta * distance_{iB}^2 - \lambda_{TV})/\rho_{TV}\Big)},$$

$$P(TV|b) = \frac{a_{TVb}\Big(\sum_{v \in F_{TV}} exp(U_v/\rho_{TVb}))\Big)^{\frac{\rho_{TVb}}{\sigma_b}}}{\sum_{f \in F} a_{fb}\Big(\sum_{v \in F_f} exp(U_v/\rho_{fb}))\Big)^{\frac{\rho_{fb}}{\sigma_b}}},$$

$$P(B) = \frac{b_b\Big(\sum_{f \in F} a_{fb}\Big(\sum_{v \in F_f} exp(U_v/\rho_{fb}))\Big)^{\frac{\rho_{fb}}{\sigma_b}}\Big)^{\sigma_b}}{\sum_v b_v\Big(\sum_{f \in F} a_{fv}\Big(\sum_{v \in F_f} exp(U_v/\rho_{fv}))\Big)^{\frac{\rho_{fv}}{\sigma_v}}\Big)^{\sigma_v}}.$$

4.4 Results

In table 4.1 we can see preliminary estimates of the nested logit model described above. All parameters relating to the correlation in the error term according to different groups have been set to one. The estimation procedure is a multinomial logit. The estimates have the expected sign. An interesting result is the estimated probability of being informed. As we can see even individuals that acquire information may have a probability of being uninformed higher than one-half (when $\alpha_{fj} + \beta_{1fj} + \beta_{2fj}$ is less than zero); this happens when individuals make mistakes in the 'tests' about political knowledge of the space.

Table 4.1: Multinomial logit parameter estimates

Parameters		**Estimates**	
1^{st} stage announcement:	Bush	Common	Kerry
		Announcement Stage	
α	3.64		1.02
β-distance		-0.33	
		Information Acquisition	
Cost Info MagAndNews-λ_M		14.58	
Cost Info MagOrNews-λ_N		17.79	
Cost Info TV-λ_{TV}		1.02	
		Probability of being informed-π	
α_M	1.30		1.46
β_{1M}	0		-0.25
β_{2M}	0		-0.54
α_N	0.70		0.86
β_{1N}	0		-0.08
β_{2N}	0		-0.87
α_{TV}	0.68		0.49
β_{1TV}	-0.53		0
β_{2TV}	-0.38		-1.42
		Voting Stage	
Civic Duty-δ	5.35		8.08
Civic Duty-δ_a		0	

Estimates were obtained maximizing the log-likelihood function with all the correlation parameters ρs, σs, as, and bs set to one. This implies the model estimated was a multinomial logit.

In table 4.4 I show a contingency table comparing the predicted choices by the model and the actual choices. The prediction is made assigning probability one to the event with the highest estimate probability given individual characteristics and the parameter estimates. I show the results for the conditional choice, taken as given the announcement made by the individuals and their information. In the diagonal we have the correct prediction of the model and the off-diagonal are the mistakes. I focus on this test for the model because I am interested in explaining individual level behavior that explains turnout and not aggregate turnout.

Table 4.4 indicates that the model is an improvement on the previous work by Poole and Rosenthal (1984) in explaining abstentions. For a comparison I reproduce here one of their tables, the one for the 1972 elections and the best result in terms of turnout explanation, see table 4.4. The predictive power of the spatial model aggregated with the uncertain-voter model increases the capacity to explain turnout. And yet, this is only a partial estimation of the proposed model. I still have room to allow for the error term to be correlated and estimated a nested logit. The model can go farther without the use of demographics characteristics.

As it stands, the model does not fair well in explaining announcement of abstention in the pre-election survey. Further work on the estimation and the use of a panel with responses for the 2004 and 2000 presidential elections may help the model fulfill its potential.

Table 4.2: Voting Choice conditional on announcement and information

Contingency Table		**Predicted Vote**		
		Kerry	Bush	Abstained
	Kerry	328	5	42
Actual	Bush	8	346	38
	Abstained	44	47	103

The probability of each action for each individual is calculated with the estimated parameters, the spatial position of the individuals, and the individual's answer to the 'test' questions. The option with the highest predicted probability is assigned as the predicted choice. Values in the diagonal are correct individual predictions. Values outsides are mistakes.

Table 4.3: Poole and Rosenthal (1984) example – the 1972 Elections

Contingency Table		Predicted Vote		
		McGovern	Nixon	Abstained
	McGovern	411	67	74
Actual	Nixon	30	884	81
	Abstained	149	290	81

The probability of each action for each individual is calculated with the estimated parameters αs and β, and the spatial position of individuals. The option with the highest predicted probability is assigned as the predicted choice. Values in the diagonal are correct individual predictions. Values outsides are mistakes.

Bibliography

Abrams, B. A. and Dougan, W. R. (1986). The effects of constitutional restraints on government spending. *Public Choice*, 49(2):101–16.

Alm, J. and Evers, M. (1991). The item veto and state government expenditure. *Public Choice*, 68(1-3):1–15.

Alt, J. E. and Lowry, R. C. (1994). Divided government, fiscal institutions, and budget deficits: Evidence from the states. *American Political Sience Review*, 88(4):811–828.

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press.

Ansolabehere, S. and Snyder, J. M. J. (2002). The incumbency advantage in u.s. elections: An analysis of state and federal offices, 1942-2000. *Election Law Journal 1(3), 315-338*, 1(3):315–338.

Arellano, M. (2003). *Panel Data Econometrics*. Advanced Texts in Econometrics. Oxford University Press, first edition.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economics Studies*, 58(2):277–297.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, pages 249–275.

Besley, T. and Case, A. (2003). Political institutions and policy choices: Evidence from the united states. *Journal of Economic Literature*, 41.

Blundell, R. and Bond, S. (1998). Initial conditions an moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87:115–143.

Bohn, H. and Inman, R. (1996). Balanced budget rules and public deficits: Evidence from the u.s. National Bureau of Economic Investigation Working Paper, 5333.

Canova, F. (2007). *Methods for Applied Macroeconomics Research*. Princeton University Press.

Coate, S. and Conlin, M. (2004). A group rule-utilitarian approach to voter turnout: Theory and evidence. *American Economic Review*, 94:1476–1504.

de Figueiredo Jr., R. J. P. (2002). Budget institutions and political insulation: Why states adopt the item veto. Forthcoming Journal of Public Economics.

Degan, A. and Merlo, A. (2006). A structural model of turnout and voting in multiple elections. Working Paper at http://www.ssc.upenn.edu/ merloa/.

Dixit, A. and Londregan, J. (1996). The determinants of success of special interests in redistributive politics. *Journal of Politics*, 58(4):1132–1155.

Ferrero, L. and Magalhães, L. M. d. (2007). Budgetary separation of powers in the american states: A regression discontinuity analysis. Mimeo.

Grossman, G. M. and Helpman, E. (2005). Party discipline and pork barrel politics. Working Paper.

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

Holtz-Eakin, D. (1988). The line item veto and public sector budgets: Evidence for the states. *Journal of Public Economics*, 36:269–292.

NCSL (2005).

Ichimura, H. and Todd, P. E. (2006). Implementing nonparametric and semiparametric estimators. Prepared for the Handbook of Econometrics Volume 5.

Klor, E. F. (2005). A positive model of overlapping income taxation in a federation of states. *Journal of Public Economics*, 90:703–723.

Knight, B. G. (2000a). Supermajority voting requirements for tax increases: evidence from the states. *Journal of Public Economics*, 76:41–67.

Knight, B. G. (2000b). Supermajority voting requirements for tax increases: evidence from the states. *Journal of Public Economics*, 76(1):41–67.

Lee, D. S. (2005). Randomized experiments from non-random selection in u.s. house elections. Version June 2005.

Levin, A., Lin, C.-F., and Chu, C.-S. J. (2002). Unit root tests in panel data: Asymptotic and finite sample propertites. *Journal of Econometrics*, 108:1–24.

Lijphart, A. (1999). *Patterns of Democracy: Government Forms and Perfomance in Thirty Six Countries*. New Haven, Yale University Press.

Matsusaka, J. G. and Palda, F. (1999). Voter turnout: How much can we explain? *Public Choice*, 98:431–466.

Merlo, A. (2005). Wither political economy? theories, facts and issues. Prepared for delivery in the symposium on Political Economy at the 2005 World Congress of the Econometric Society, London, August 19-24, 2005.

Pagan, A. and Ullah, A. (1999). *Nonparametric Econometrics*. Cambridge University Press, first edition.

Persson, T., Roland, G., and Tabellini, G. (2000). Comparative politics and public finance. *The Journal of Political Economy*, 108(6):1121–1161.

Persson, T. and Tabellini, G. (2003). *The Economic Effects of Constitutions*. Munich Lectures in Economics. MIT Press.

Poole, K. T. (2005). *Spatial Models of Paliamentary Voting*. Cambridge University Press.

Poole, K. T. and Rosenthal, H. (1984). The polarization of american politics. *The Journal of Politics*, 46(4):1061–1079.

Poole, K. T. and Rosenthal, H. (1997). *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press.

Poterba, J. (1994). State responses to fiscal crisis: The effects of budgetary institutions and politics. *The Journal of Political Economy*, 102(4):799–821.

Poterba, J. M. and von Hagen, J. (2000). *Fiscal Institutions and Fiscal Perfomance*. MIT Press.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica*, 56(4):931–954.

Shugart, M. S. and Carey, J. M. (1992). *Presidents and Assemblies: Constitutional Design and Electoral Dynamics*. Cambridge University Press.

Takane, Y. and W., Y. F. (1977). Nonmetric individual differences multidimensional scaling; an alternating least squares method with optimal scaling features. *Psychometrika*, 42(1):7–67.

Urofsky, M. I. and Finkelman, P. (2002). *March of History: a Constitutional History of the United States*, volume 2. Oxford University Press, 2 edition.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Appendix A

A.1 Definition of Equilibrium with Line Item Veto

The equilibrium in the legislative game with line item veto is a vector of policies $q_t^V(\omega_t)$ and a vector of reservation utilities ω_t^V, such that, in any period t, when all players take as given the equilibrium outcomes of periods $t+k, k \geq 1$:

1. for any given ω_t, at the veto stage, the line item veto legislator E prefers $q_t^V(\omega_t)$ to the status quo and to the policy vector proposed by the legislator $p_t^V(\omega_t)$ in which $f_t^{LV}(\omega_t)$, $f_t^{EV}(\omega_t)$, $r_t^{EV}(\omega_t)$, and $r_t^{LV}(\omega_t)$ is greater or equal than in $q_t^V(\omega_t)$;

2. for any give ω_t, the legislator L prefers $q_t^V(p_t^V(\omega_t))$ to any other policy satisfying the conditions above;

3. the reservation utilities ω_t^{iV} are optimal for the voters of each type i, when one takes into account that policies in the current period are set according to

$q_t^V(\omega_t)$ and takes as given the reservation utilities of the type $j \neq i$, the identity of the agenda setter, and of the executive.

A.2 Kernel and Bandwidth Choice

To estimate each nonparametric regression we have used the local linear procedure as described in Pagan and Ullah (1999) p.93.. The method consists in minimizing for m:

$$\sum_{i=1}^{n} \left\{y_i - m - (x_i - x)\right\}^2 K\left(\frac{x_i - x}{h}\right),$$

where $K(.)$ is the kernel function and h the bandwidth.

The local linear regression method, as argued in Hahn et al. (2001), fairs relatively better than other methods at the boundaries and, therefore, is the most appropriate for regression discontinuity analysis. Let's calls $s = \frac{x_i - x}{h}$, our choice of Kernel is:

$$K = \frac{15}{16}(s^2 - 1),\ for\ s \leq 1\ and\ 0\ otherwise.$$

Monte Carlo studies have shown that the choice of Kernel does not affect the estimation by much. Similar results would be found using the normal density as the kernel function. What matters is the choice of bandwidth h. Many methods have been devised to find the optimal bandwidth. The rule-of-thumb bandwidth is given by $h^* = 1.06 * \sigma * n^{-\frac{1}{5}}$, where n is the number of observations and σ the standard deviation (in our case of *gov_strength*). Monte Carlo exercises have shown, however, that the rule-of-thumb bandwidth over-smooths the estimation. In our case

$h^* = 0.057$. All estimation results are made with an $h = 0.05$. The point estimate does not change by much if we vary the bandwidth from 0.45 to 0.6 as we show in Table 5. Any lower and we run into the problem of trying to run a local linear regression with less than 2 data points for some point we are estimating, which is not identified.

A.3 RDD – Polynomial Strategy

In this section we show another strategy to identify the shape of the non-linear part of the model and the discontinuity. It illustrates, at least in our case, the dominance of the semiparametric method. The results are highly dependent on the choice of polynomials.

One Polynomial Specification

Our first strategy is to include in the panel estimate along with the controls and the state and year dummies, various polynomial degrees of *gov_strength*. To identify the discontinuity we still include the dummy *divided*. If we are able to correctly estimate the shape of the function, the dummy gives us the causal effect of *divided* on the tax level at the discontinuity.

The results can be seen in **Table 4**, where we try different specifications in different columns, from a 3-degree to a 6-degree polynomial. We attempted with higher degree polynomials, but *divided* remained statistically insignificant.

Table A.1: Regression Discontinuity - One Polynomial

	(1) ttax_gdpp	(2) ttax_gdpp	(3) ttax_gdpp	(4) ttax_gdpp
divided	-0.14	-0.14	0.02	0.04
	(0.07)**	(0.07)**	(0.08)	(0.08)
	(0.12)	(0.13)	(0.12)	(0.12)
gov_strength	4.53	5.41	13.68	29.85
	(1.21)***	(2.24)**	(3.50)***	(5.27)***
	(3.54)	(6.32)	(8.61)	(12.26)**
gov_strength2	-11.59	-15.54	-75.56	-241.80
	(2.60)***	(7.82)**	(18.95)***	(40.39)***
	(7.03)	(20.94)	(42.76)*	(95.83)**
gov_strength3	8.59	14.73	176.74	833.26
	(1.63)***	(10.66)	(45.21)***	(140.80)***
	(4.12)**	(27.72)	(95.13)*	(337.79)**
gov_strength4		-3.07	-185.01	-1,398.84
		(4.97)	(47.89)***	(242.83)***
		(12.71)	(95.81)*	(585.84)**
gov_strength5			72.33	1,128.77
			(18.44)***	(202.46)***
			(35.59)*	(488.68)**
gov_strength6				-349.01
				(65.07)***
				(156.39)**
Sample LIV 2/3 1159 observations Baseline controls				

Huber-White robust and clustered-robust standard errors in parentheses. Number of clustered groups 47. The states of Nebraska, Alaska, and Hawaii are excluded in all regressions. The symbol $*$ is significant at 10%; $**$ significant at 5%; $***$ significant at 1%. Baseline controls include: state population, state population squared, percentage of aged, and of kids; state and year dummies; *cycle_trend*, *demgov*, *indgov*, *restrict*.

Graphically the result for the 3-degree specification can be seen in **Figure** A.1. Since in this case *divided* is significant we add the jump at 0.5. In **Figure** A.2 we show the graphical result for the 6-degree polynomial specification, since *divided* is not significant we do not add the jump at 0.5.

Results depend on the polynomial degree specification. The identification of the jump depends on correctly identifying the shape of the function. Later we use a semiparametric estimate of its shape that does not rely on the choice of polynomial degrees. We can use it as a reference of our semiparametric estimates and compare whether the polynomials can reproduce its shape. Of course, as we add more polynomials, asymptotically the results should be the same.

One function for each side of the discontinuity

Another parametric strategy is to estimate two sets of polynomials from both sides of the discontinuity. That's what we do in **Table5**. The variable g_sLeft is defined as $gov\_strength \times divided$ and g_sRight as $gov\_strength \times (1 - divided)$. In column 1 we show a 3-degree specification on both sides of the discontinuity; in column 2, 4 on the left and 5 on the right; and in column 6 on the left and 7 on the right. In the last column $g\_sRight4$ is dropped due to multicollinearity. We show the graphical results for these different specifications in **Figures** A.3, A.4, and A.5.

The point estimation itself comes from calculating the difference between the value of the left side function at 0.5 and the right side function at 0.5 as well. To

Figure A.1: Three degree polynomial

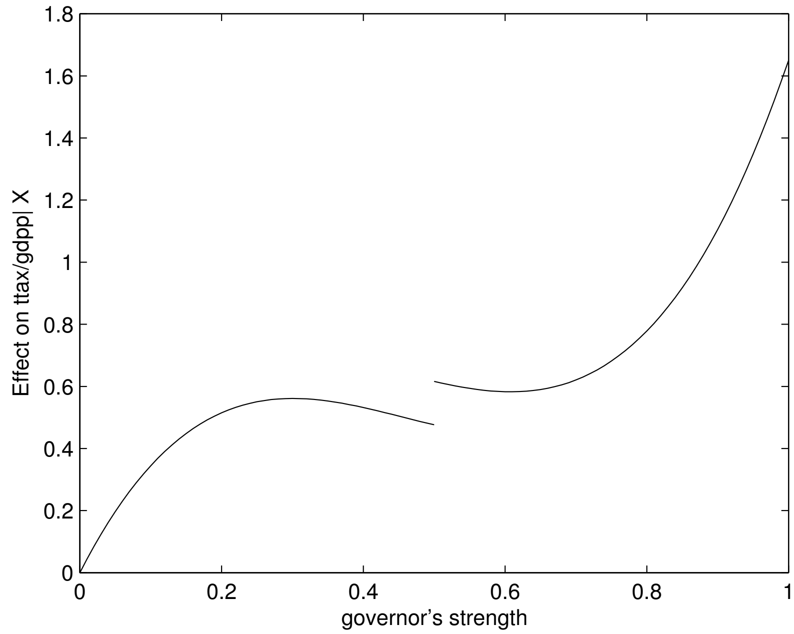


Figure A.2: Six degree polynomial

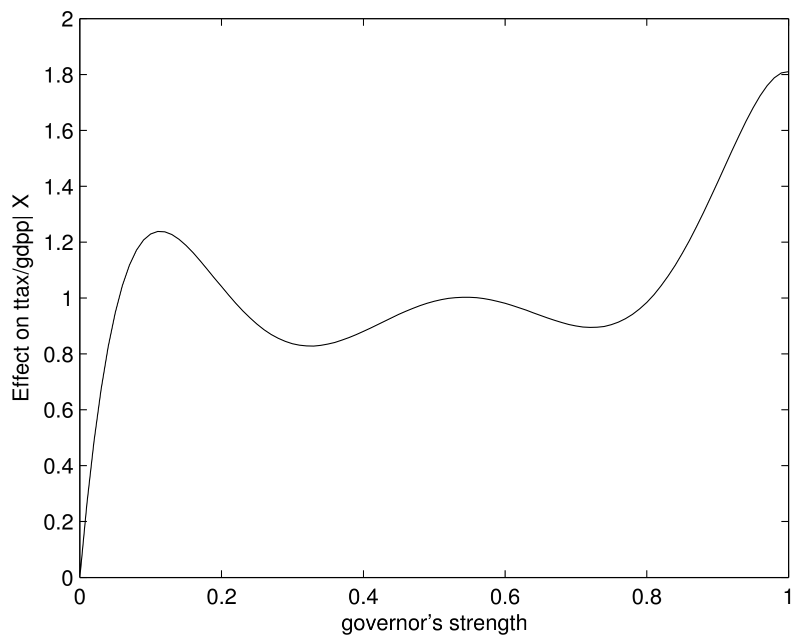


Table A.2: Regression Discontinuity - Two Polynomials

	(1) ttax_gdpp	(2) ttax_gdpp	(3) ttax_gdpp
g_sLeft	10.18 (2.80)*** (7.29)	34.67 (5.37)*** (12.19)***	47.16 (13.19)*** (22.43)***
g_sLeft2	-39.80 (10.62)*** (26.37)	-264.53 (38.05)*** (88.20)***	-552.85 (210.28)*** (315.26)***
g_sLeft3	45.62 (12.29)*** (29.41)	724.62 (104.63)*** (242.84)***	3,103.63 (1,474.25)*** (2,113.44)
g_sLeft4		-657.94 (97.13)*** (223.35)***	-9,628.49 (5,092.39)* (7,179.04)
g_sLeft5			15,772.93 (8,497.24)* (11,881.66)
g_sLeft6			-10,497.35 (5,469.69)* (7,615.32)
g_sRight	6.32 (1.38)*** (3.29)*	-49.43 (29.43)* (66.09)	-375.45 (80.92)*** (130.69)***
g_sRight2	-14.72 (3.03)*** (6.31)**	301.94 (162.95)* (372.95)	2,116.87 (452.46)*** (738.03)***
g_sRight3	10.34 (1.92)*** (3.80)**	-642.87 (333.18)* (773.25)	-3,648.41 (781.02)*** (1,287.40)***
g_sRight4		585.81 (297.76)** (697.78)	- - -
g_sRight5		-193.13 (98.18)** (231.42)	6,610.08 (1,435.74)*** (2,416.10)***
g_sRight6			6,968.28 (1,529.91)*** (2,600.68)***
g_sRight7			2,267.72 (503.73)*** (864.82)**
Sample LIV 2/3	1159 observations	Baseline controls	

Huber-White robust and clustered-robust standard errors in parentheses. Baseline controls include: state population, state population squared, percentage of aged, and of kids; state and year dummies; *cycle_trend*, *demgov*, *indgov*, *restrict*.

Figure A.3: Two polynomials 3-left 3-right

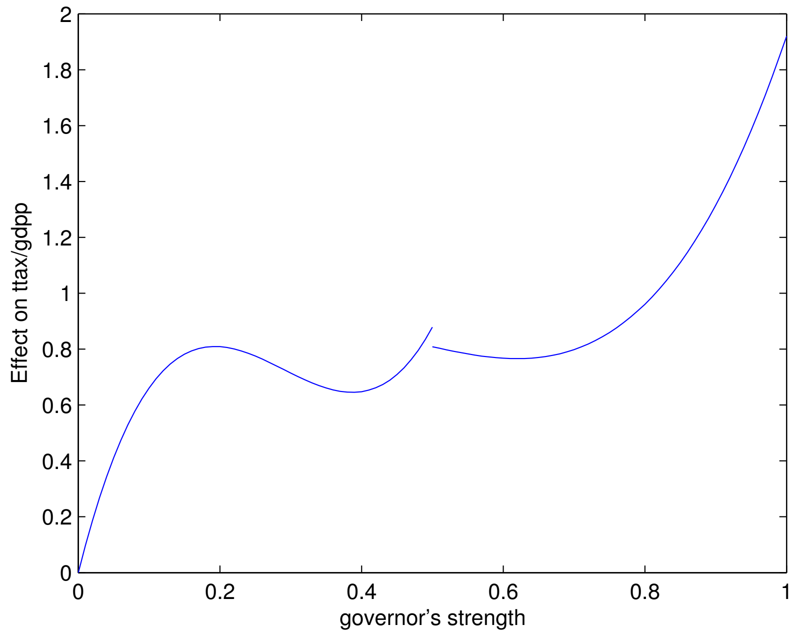


Figure A.4: Two polynomials 4-left 5-right

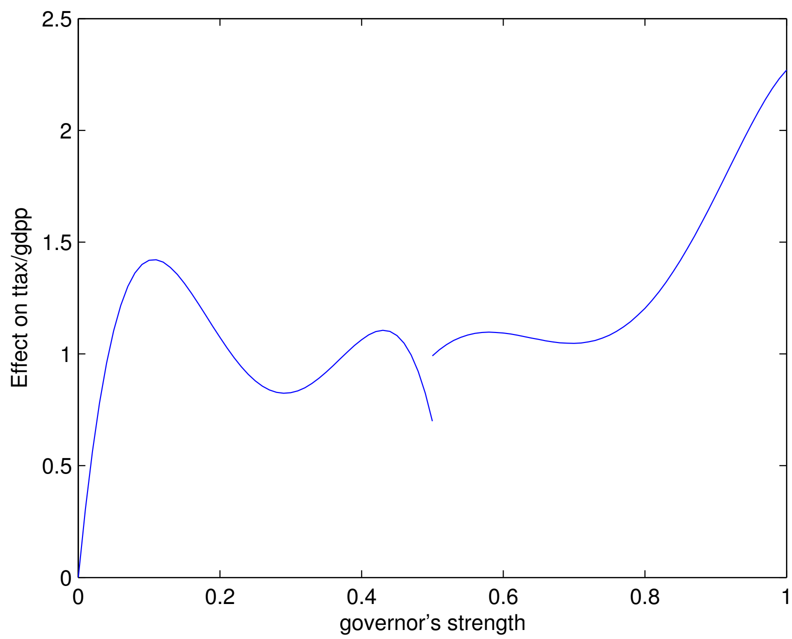
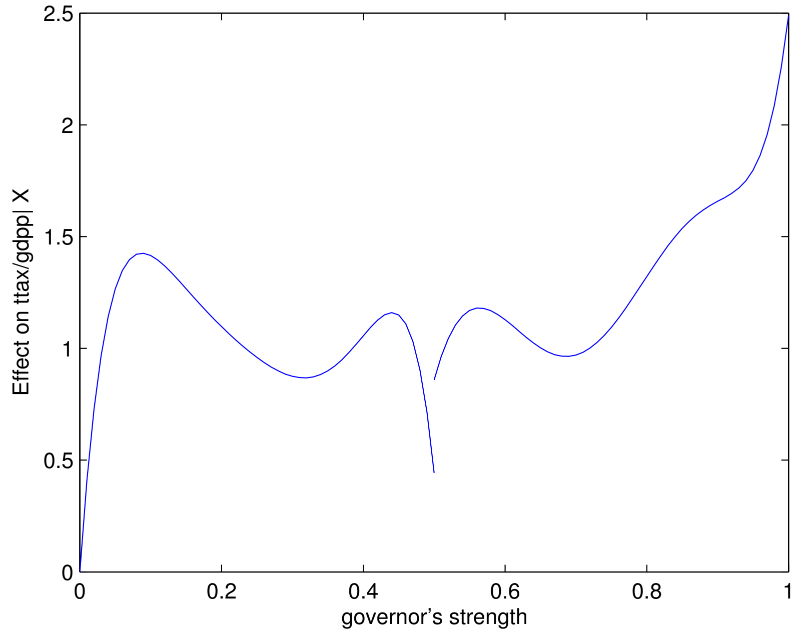


Figure A.5: Two polynomials 6-left 7-right



establish a standard error around our estimate of the discontinuity we bootstrap the polynomial estimation 50 times and at each time save the difference between the point estimate of the left and right side function at 0.50. As can be seen from the figures, results vary according to the specifications. For the specification in column 1 the point estimate is 0.07 and not significant. For the specification in column 2, the point estimate is -0.33 and it is significant at the 1% level. The last point estimate for column 3 is -0.42 and significant at the 1% level as well.

Table A.3: Bootstrapped results

Polynomial Specification	**Estimated Jump**	**(Std. Err.)**
3-left & 3-right	0.067	(0.080)
4-left & 5-right	-0.327***	(0.106)
6-left & 7-right	-0.422***	(0.149)

Bootstrapped standard errors were retrieved resampling 50 times with replacement.

If we compare the graphical results in Figures A.3 to A.5 we see that they do a better job of capturing the shape of the function estimated semiparametrically in Figure 2.1. However, the results are still dependent on the polynomial degree assumption we make.