ESSAYS IN POLITICAL ECONOMY AND APPLIED ECONOMETRICS

A Dissertation

Presented to

The Faculty of the Department of Economics

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Pablo J. Garofalo

May, 2014

ESSAYS IN POLITICAL ECONOMY AND APPLIED ECONOMETRICS

An Abstract of a dissertation

Presented to

The Faculty of the Department of Economics

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Pablo J. Garofalo

May, 2014

Abstract

This dissertation is comprised of two essays in political economy and one in micro-econometrics. Each of them proposes an alternative methodology to improve on the estimation of a specific economic phenomenon.

The first essay studies the political allocation of US federal resources to localities taking into consideration that the states are also actively involved in allocating resources to localities. I found that federal funds are biased towards localities within states that are not represented by the same party as the one that represents the federal government. This finding implies that a strategic federal government takes into account that non-aligned states have different spending priorities. These results suggest that past research on the allocation of federal resources to localities has shown biased estimates when the political allocation of resources is not studied in the context of a multi-layered government environment.

The second essay exploits the existence of extended interlude periods (i.e., time between elections and government change date) from Latin American countries to identify a causal effect of a change in the probability of electoral defeat on a change in the budget deficit. Theoretical studies on the strategic use of debt argue that governments issue more debt when facing a higher probability of electoral defeat. Testing this hypothesis has proven challenging since measures of that probability are potentially endogenous. Since my identification strategy is focused on identifying the effects of electoral surprises, I provide a plausible source of exogenous variation. I find that the higher the increase in the probability of electoral defeat (victory), the larger the increase (decrease) in the deficit.

The third essay studies the properties of a maximum likelihood estimator (MLE) of dynamic panel data models with fixed effect when difference GMM methods suffer from weak identification. While previous studies propose moments to solve the weak identification under difference GMM for stationary processes only, this study shows that MLE solves the weak identification issue not only when the process is stationary, but also when it is not.

Acknowledgements

I am grateful to Gergely Ujhelyi for his exceptional advice and encouragement during these years. I thank Bent E. Sørensen and Chinhui Juhn for very helpful comments. This dissertation would not have been possible without their guidance. I also thank Francisco Cantu for serving on my dissertation committee, and Jorge Streb for his helpful comments during these years.

Sine Tepe deserves my deepest gratitude for her never-ending support and belief.

Finally, I want to thank my family for their constant support.

Contents

Chapter 1

Strategic Spending in Federal Government: Theory and Evidence from the US

1.1 Introduction

Incumbent politicians may have several reasons to sway the distribution of resources away from pure welfare maximization. Electoral competition may induce incumbents to allocate more resources to localities with a high proportion of swing voters - voters who are not specifically attached to any of the parties.[1] By contrast, if politicians are risk averse, they might see a safer investment in targeting partisan localities - localities with many voters loyal to the incumbent's party.[2] Apart from their own policy objectives, incumbents may also want to help other politicians from their party: for example, influencing the composition of Congress can help to enact a greater portion of the executive's legislative agenda when a large number of co-partisans reside within Congress.[3]

Studies on the political allocation of resources tend to treat democracies as if they were unitary systems with only one level of government actively involved in the distribution of resources. In reality, however, multiple levels of government each pursue their own political goals. For example, apart from the US federal government, states also allocate resources to localities.[4] In such a system, governments will have an incentive to act strategically. For example, the central government should consider whether to allocate resources to state governments or to localities taking into account

[1] Lindbeck and Weibull (1987, 1993), Case (2001), Strömberg (2008), Arulampalam et al. (2009).

[2] Cox and McCubbins (1986), Ansolabehere and Snyder (2006), Larcinese et al. (2006).

[3] Colleman (1999), Howell et al. (2000).

[4] Studies on the political determinants of state-county transfers include Frederickson and Cho (1974), Ansolabehere et al. (2002), and Ansolabehere and Snyder (2006).

that state governments will also target specific localities. In this paper, I model this type of strategic interaction between different levels of government, and test the model using data on the distribution of federal transfers between states and localities in the US.

Government preferences regarding resource allocation are determined to a large extent by party politics. State and federal governments may be aligned (when they are controlled by politicians from the same party) or non-aligned (when they are controlled by different parties). Because aligned governments have similar preferences regarding resource allocation, they are likely to want to target spending towards the same localities. Non-aligned governments, by contrast, are likely to have different spending priorities. A strategic federal government should take this into account. It should spend more on its preferred localities in non-aligned states, where these localities are likely to be at a disadvantage, than in aligned states, where the same localities are likely to receive state funds as well. Considering federal-to-state rather than federal-to-local transfers, federal transfers to aligned states should be greater than to non-aligned states, since the former would behave as a political partner and the latter as a political competitor of the federal government.

I formalize this idea by setting up a sequential move game with perfect information in which the federal government is the leader and the states are the followers. States can be aligned or non-aligned with the federal government, and each player chooses the intergovernmental transfers made to lower level governments (federal-to-state, federal-to-local, state-to-local). I show that, in equilibrium, the federal government will not transfer funds to localities that are also the target of state spending. Doing so would simply crowd out similar spending by the state. In aligned states, the optimal federal strategy is to target spending towards the state government. By contrast, the federal government does transfer directly to localities in non-aligned states, since these state governments have different spending priorities. The prediction therefore is that we should observe more federal transfer to politically preferred localities within non-aligned states than within aligned ones.

I estimate the predictions of the model using data on the allocation of US federal government transfers between states and counties. I follow a difference-in-difference strategy to test whether the federal government transfers more resources to politically aligned counties within non-aligned states than within aligned ones.[5] Consistent with the model, I find that the federal government increases transfers to politically aligned counties by around 6 percentage points, or roughly $11.50 per capita when the state government changes from being aligned with the federal government to being non-aligned. There is no evidence for such an increase for non-aligned counties. This demonstrates the importance of controlling for the three-way political alignment between local, state, and federal government when studying the determinants of intergovernmental spending. The finding that these interaction terms matter survives a long list of robustness checks - among others, controlling for unobserved heterogeneity at the state-year level, alternative definitions of political alignment, and an IV estimation to control for state transfers to localities.

My study has three broad implications. First, my results suggest that previous findings on the political determinants of federal transfers to localities may contain biased estimates. For example, some previous studies estimate the effect of local-federal political alignment on the allocation of federal transfers without controlling for state-federal alignment (Levitt and Snyder, 1995, 1997; Berry et al., 2010). If local-federal and state-federal alignment are positively correlated, my findings imply that the effect of local-federal alignment will be underestimated since it represents a weighted average of non-aligned states, where I find strong effects, and aligned states,

[5] Section 3 contains a detailed discussion on the construction of these measures of political alignment.

where I find none.

Second, I show that - once the strategic interaction between governments is taken into account - the data shows evidence of political opportunism in the allocation of US federal transfers. The federal government appears to take advantage of the multi-layered system of government in bringing federal dollars to its constituencies. While some previous studies highlight the political incentives present in a federal system (Dixit and Londregan, 1998; Volden, 2005), to my knowledge this is the first paper to test this empirically.

Third, my results have general implications for normative studies of decentralization. Other scholars have studied the efficiency gains from decentralization, either in the sense of aggregate surplus or in the Pareto sense (Oates, 1972; Lockwood, 2002; Besley and Coate, 2003). However, these studies compare public good provision in a pure central system to pure regional or local provision. My results suggest that a federal system with both central and multiple lower governments behaves differently from these extremes. In this type of decentralized system, the federal government might engage in a sort of competition with non-aligned states for mobilizing voters, while cooperating with states that are politically aligned with it.

The rest of the paper is organized as follows. The next section places this paper in the related literature. In section 3, I explain in detail how I define political alignment based on previous studies. In section 4, I present the model. In section 5, I present the data and econometric specification used to test the theoretical predictions. Section 6 contains the main empirical results, and section 7 the robustness checks. Finally, section 8 concludes.

1.2 Related literature

There are three types of studies in the literature on the political allocation of governmental resources: some study the allocation of federal resources to state governments, others the allocation of state resources to localities, and still others the distribution of federal resources to localities. None of the studies in the third group control for the interaction between federal and state governments. In this sense, my study brings together these previous papers by including all three effects.

In the first group of papers, on federal transfers to state governments, Grossman (1994) estimates that federal grants increase when the number of public employees and union membership per capita increase. He also finds that federal grants to states increase when the percentage of seats held by Democrats in the House of Representatives increases. Larcinese et al. (2006) show that federal outlays to states are affected mainly by the President. Contrary to the common belief, the Senate and the House of Representatives have much smaller impact on federal outlays. In particular, the authors find that federal transfers are affected mainly by the alignment between the President and the state governor and by the alignment between the former and the majority of the state delegates in the House. By contrast, the governor's alignment with either the House or the Senate has no effect.

In the second group of papers studying the relationship between states and localities, Ansolabehere and Snyder (2006) examine the effect of party control of the state on the allocation of the state budget. They find that the party that controls the state (which is not necessarily the party of the Governor) skews the distribution of funds towards partisan localities. By contrast, they find weak evidence that swing voters are being targeted.

In the third group of papers, on the allocation of the federal budget to localities, Levitt and Snyder (1995) estimate that, over a period of Democratic control of

Congress, federal programs with higher variability across districts were biased towards districts with more Democrats.[6] Berry et al. (2010) follow Larcinese et al. (2006) but use federal outlays to localities instead of states. They also find that the president has ample opportunities to influence the allocation of high variability funds to localities, both before and after congressional approval of the budget. Specifically, federal spending to counties increases if the county's House Representative is aligned with the President. In contrast, they do not find evidence that congressional committee assignments influence federal spending.

Bringing these results together, if the federal government transfers more funds to aligned states, and states allocate more resources to aligned localities, then some of the federal-to-state transfers might reflect the ultimate objective of targeting localities aligned with the federal government. At the same time, this also implies that the federal government will have more incentive to directly transfer funds to aligned localities within non-aligned states. This is the starting point of my analysis below. My findings will imply that studies such as Levitt and Snyder (1995) and Berry et al. (2010), which do not control for federal-state alignment, are likely to underestimate the effect of political alignment on federal-to-local transfers.

I know of only two (theoretical) studies that consider strategic interaction between different levels of government. Dixit and Londregan (1998) study a model of political platform competition and compare a centralized government with two levels of political competition, central and state. They predict that the central policy implemented is going to be a function of the policy implemented at the state level, since state politicians compete during the second stage of the game. In Volden's (2005) model, state and federal governments may compete in the provision of public goods, leading

[6] High variability programs are assumed to be more discretionary, and hence more likely to reflect political motivations. See more on this in Appendix B.

to over-taxation and over-provision because both seek credit via public spending and they do not want to be blamed for taxing. My contribution relative to these studies is to focus on the role of political alignment in the strategic interaction between governments, and to provide empirical evidence consistent with my model.

1.3 Background: Who controls the budget?

This section discusses the concept of political alignment between governments based on which actor is most likely to have control over the allocation of the budget.

At the federal level, both in the construction of budgets and in their implementation, the President has ample opportunities to affect the geographic distribution of federal outlays since the Budget and Accounting Act of 1921. The President has been responsible for composing a complete budget, which is submitted to Congress in February of each year, and which initiates the actual authorization and appropriations processes. Substantial efforts are made to ensure that the president's budget reflects his or her policy priorities (Berry et al., 2010). The Office of Management and Budget (OMB) is an important vehicle of presidential control. Rather than submitting requests directly to Congress, agencies seeking federal funding must submit detailed reports to the OMB. The OMB clears each of these reports to ensure that it reflects the chief's executive's policy priorities. The end product is a proposed budget that closely adheres to the President's policy agenda. This ability of the President to target funds towards desired areas does not imply that the members of Congress cannot make amendments. However, the threat of a presidential veto gives members of Congress an incentive to keep the budget proposal close to the initial form proposed by the President (McCarty, 2000).[7] The President also has substantial in-

[7] This threat does not apply when a supermajority in Congress would be likely to overturn

fluence over the allocation of federal funds once the budget has been approved. For instance, administrative agencies can be created through executive action; in such a case, they are significantly less isolated from presidential control than are agencies created through legislation (Howell and Lewis, 2002). In addition, the President can reprogram funds within certain budgetary accounts; and with Congress's approval, he can transfer funds between accounts (Berry et al., 2010). In light of these facts, the President's party will be taken in this paper as the party that controls the Federal budget. As discussed in section 2, this is consistent with the empirical findings of Larcinese et al. (2006) and Berry et al. (2010), among others.

Regarding state governments, there are a variety of ways to define party control of the state. One option is to use the governor's party, analogously to the federal level. However, it is important to note that, in contrast with Congress, during my period of study (from 1982 to 2002) there were several instances of a party having a super-majority in both chambers of the state legislature without holding the Governor's seat. In such cases overturning a Governor's veto would have been likely, and this has to be taken account in order to define the state control of the budget accurately. In the main analysis, I will use the measure used by Ansolabehere and Snyder (2006), which accounts for this type of divided government.[8] Based on this measure, the state is under, say, Democratic control if (i) Democrats have a majority in both legislative chambers and the Governor is a Democrat, or (ii) Democrats hold at least two thirds of the seats in both legislative chambers. Republican control is defined analogously.

a presidential veto. In such a case, we might expect the budget to be less representative of the President's priorities. During my period of analysis (1982-2002), there was never a super-majority against the President's party, therefore overturning his veto would have been highly unlikely.

[8] I discuss the robustness of my findings using alternative measures in section 7.

Ansolabehere and Snyder show that, under this definition of party control of the state, state funds are targeted towards localities where the fraction of political supporters is the highest.

1.4 Theoretical framework

I model the political allocation of government expenditures by two levels of government: federal and state. Both governments can spend directly at the local level (by spending funds in specific districts or counties). In addition, the federal government can make intergovernmental transfers to states, giving them discretion in how these funds are ultimately spent.

Consider two states, $i = 1, 2$, with the same number of counties and assume that the party that controls State 1 is aligned with the President and the party that controls State 2 is not.[9] Counties in both states can be politically preferred by the President (represented by the set F_i) and / or politically preferred by the State i (represented by the set S_i). Following the literature discussed in sections 1 and 2 above, a county may be "politically preferred" because it has many loyal voters, or because it is a swing county. The source of political preference will not matter for the theory, but I will consider each of these possibilities separately in the empirical analysis below. Assume that, in state i, m_i counties are politically preferred by both the President and the party in control of the state ($|F_i \cap S_i| = m_i$), n_i counties are politically preferred by the President only ($|F_i \backslash (F_i \cap S_i)| = n_i$), and r_i counties are politically preferred by the party in control of the state only ($|S_i \backslash (F_i \cap S_i)| = r_i$). This

[9] As in section 3, a state is aligned if the party that controls the state budget is the same as the President's party.

is illustrated in Figure 1. As the figure makes clear, it is not unrealistic to assume that the number of counties that are preferred by both the President and the state government is higher in State 1 than in State 2 ($m_1 > m_2$) since the former is aligned with the President. Similarly, the number of counties that are preferred by one level of government only is higher for State 2 ($n_1 < n_2$ and $r_1 < r_2$).[10]

Each county is represented by an elected congressman who may or may not be from the President's party. Let the sets $f_i \subset \{F_i \backslash (F_i \cap S_i)\}$ and $fs_i \subset (F_i \cap S_i)$ denote the counties whose House Representative is aligned with the President, and $\{(F_i \backslash (F_i \cap S_i)) \backslash f_i\}$ and $\{(F_i \cap S_i) \backslash fs_i\}$ the sets of counties whose House Representative is non-aligned. I assume that $\frac{|f_1|}{|F_1 \backslash (F_1 \cap S_1)|} = \frac{|f_2|}{|F_2 \backslash (F_2 \cap S_2)|} = \frac{|fs_1|}{|F_1 \cap S_1|} = \frac{|fs_2|}{|F_2 \cap S_2|} = \alpha$, i.e., there is a constant share of counties aligned with the President within each group for each state.[11]

The President decides in the first stage of the game how much to transfer to each state (T_1^S and T_2^S) and how much to transfer directly to each county j within each state (T_{j1}^C and T_{j2}^C). In the second stage of the game, both state 1 and 2 decide how much to transfer to each county (t_{1j}^C and t_{2j}^C respectively). I will assume that the government's budget is exogenous in order to avoid dealing with another source of political opportunism, that is raising or lowering taxes. The federal government's budget is $\widetilde{B}^F$and states' budgets are $\widetilde{B}^1$and $\widetilde{B}^2$ respectively.

Assuming that all individuals have the same utility function and the same personal income, the representative individual's utility function of locality j in State $i \in (1, 2)$

[10] Intuitively, aligned states have more things in common with the President, hence the preferences over the political allocation of resources are more similar than in the non-aligned state.

[11] One could instead assume that the proportion of aligned localities is higher within the aligned state. The assumption of constant proportion within each state simplifies the algebra without affecting the main result of the model.

is $U^{ij} = H(x^{ij})$, where $H(0) = 0$, $H'(x) > 0$, $H''(x) < 0$, and x^{ij} is the total public spending in the county. Public spending could be financed by either the State i only, State i and the President, or by the President only. Following Oates (1999), I assume that higher level governments are less efficient at spending at the local level than lower level governments that are "closer" to the target of spending.[12] Specifically, I let total public spending be $x^{ij} = \theta T^C_{ij} + t^C_{ij}$, where $\theta \in (0,1)$ represents the relative inefficiency or leakage of President provision compared with the state provision.

The President's payoff is

$$\sum_{i=1}^{2}\Big(\sum_{j\in fs_i}\gamma H\left(\theta T^C_{ij} + t^C_{ij}\right) + \sum_{j\in f_i}\gamma H\left(\theta T^C_{ij}\right) + \sum_{j\in((F_i\cap S_i)\setminus fs_i)} H(\theta T^C_{ij} + t^C_{ij}) + \sum_{j\in((F_i\setminus(F_i\cap S_i))\setminus f_i)} H\left(\theta T^C_{ij}\right)\Big),$$

where $\gamma > 1$ represents a relative preference towards spending in localities that have an aligned House Representative.[13] The President faces the following budget constraint:

$$\widetilde{B}^F = \sum_{i=1}^{2}\Big(\sum_{j\in F_i} T^C_{ij} + T^S_i\Big).$$

[12] Oates (1999) argued that lower level governments should be more efficient in providing local public goods because they are "closer to the people," possessing knowledge of both local preferences and cost conditions that a central agency is unlikely to have. Such local knowledge could also make the political allocation of resources more effective when lower levels of government take the lead.

[13] Presidents may have various reasons to help members of their own party. For example, based on the discussion in Sections 2 and 3, a president can avoid the potential overturn of a future veto, and thereby keep control of the budget, by ensuring that a certain number of co-partisans are elected into office.

State i's payoff is

$$\sum_{j\in(F_i\cap S_i)} H\left(\theta T_{ij}^C + t_{ij}^C\right) + \sum_{j\in(S_i\setminus(F_i\cap S_i))} H\left(t_{ij}^C\right),$$

and it faces the budget constraint

$$\widetilde{B}^i + T_i^S = \sum_{j\in S_i} t_{ij}^C.$$

Note that, because each government only cares about counties that are preferred by it, $t_{ij}^C = 0$ for $j \in (F_i\setminus(F_i\cap S_i))$ and $T_{ij}^C = 0$ for $j \in (S_i\setminus(F_i\cap S_i))$.[14]

Solving the model using Backward Induction yields the following:

Proposition 1. *In a Subgame Perfect Nash Equilibrium, (i) federal transfers to counties that are politically preferred by the President only will be larger when the House Representative is from the President's party ($T_{ij}^C \equiv T_a^C > T_{il}^C \equiv T_{\sim a}^C$ for $i = 1, 2$, $j \in f_i$, $l \in ((F_i\setminus(F_i\cap S_i))\setminus f_i)$); (ii) federal transfers to counties that are preferred by both the President and the state will be equal to zero regardless of the House representative's party ($T_{ij}^C = T_{il}^C = 0$ for $i = 1, 2$, $j \in fs_i$, $l \in (FS_i\setminus fs_i)$).*

Proof. See appendix A. ■

Part (i) of Proposition 1 follows simply from the fact that the President puts higher weight on counties with an aligned representative. Part (ii) is more surprising: it says that the President will not transfer funds to counties that are also politically preferred by the state. To interpret this result, consider the states' reaction function

[14] Similarly to the President, I assume that States only care about their preferred counties. This means that a State's payoff is not affected by federal transfers to its non-preferred counties. This assumption could easily be relaxed: as long as the State attaches a higher weight to preferred counties, allowing non-preferred counties to also have a positive weight would not affect the main implications of the model.

from solving their maximization problem in the second stage of the game:

$$t_{ij}^{C*} = \frac{1}{m_i + r_i}\left[\widetilde{B}^i + T_i^S + \theta \sum_{l\in(F_i\cap S_i)} T_{il}^C\right] - \theta T_{ij}^C, \text{ for } i = 1,2 \text{ and } j \in (F_i \cap S_i) \quad (1)$$

$$t_{ij}^{C*} = \frac{1}{m_i + r_i}\left[\widetilde{B}^i + T_i^S + \theta \sum_{l\in(F_i\cap S_i)} T_{il}^C\right], \text{for } i = 1,2 \text{ and } j \in S_i \backslash (F_i \cap S_i) \quad (2)$$

Consider the President's choice between transferring an extra dollar to county $j \in (F_i \cap S_i)$ or to State i. In the first case, county j would receive a fraction $\theta < 1$ of that dollar. Moreover, given (1), State i would decrease the transfer to that county j by the amount $\triangle t_{ij}^{C*} = (\frac{1}{m_i+r_i} - 1)\theta$ and given (2), it would increase the transfers to all the other counties in the group $S_i - \{j \in (F_i \cap S_i)\}$ by the amount $\triangle t_{ij}^{C*} = (\frac{1}{m_i+r_i})\theta$ to keep the total public spending in each county that belongs to the state i's preferred group S_i equal. Instead, if the President gave the 1 dollar to State i, then the State would increase the transfers to each county in the group S_i by the same amount $\triangle t_{ij}^{C*} = \frac{1}{m_i+r_i}$. Comparing the two strategies, the president can target "indirectly" each of his preferred counties in the group $F_i \cap S_i$ with an extra amount of $(\frac{1}{m_i+r_i})(1-\theta)$ dollars if he transfers one extra dollar to states and not directly to the counties in that group. Then, transferring to his preferred counties in the group $F_i \cap S_i$ is dominated by transferring to the state. This property of the equilibrium comes from the fact that the President is comparatively inefficient at allocating political resources, combined with the fact that he knows that each State i can undo anything he does in the first stage, to meet State i's goals in terms of political allocation.

By contrast, the President does transfer to counties that only he prefers ($j \in F_i \backslash (F_i \cap S_i)$), because State i is not allocating any funds to them. Hence, in equilibrium, the President will allocate resources to his own preferred counties only.

Since the number of counties within each of the three groups ($F_i \cap S_i$, $F_i \backslash (F_i \cap S_i)$ and $S_i \backslash (F_i \cap S_i)$) differs between State 1 and 2, we observe, on average, different federal transfers to the President's preferred counties within the non-aligned state 2 and within the aligned state 1. Formally stated, we have:

Corollary 1. *Average federal transfers to the President's preferred counties are greater in the non-aligned State 2 than in the aligned State 1. The difference between the states is greater in the case of counties that have a House Representative aligned with the President. Formally,* $(\frac{n_2}{(n_2+m_2)} - \frac{n_1}{(n_1+m_1)})T_a^C > (\frac{n_2}{(n_2+m_2)} - \frac{n_1}{(n_1+m_1)})T_{\sim a}^C > 0.$

Corollary 1 is the main result of the theoretical model. On average, we observe greater federal transfers to preferred counties within non-aligned states because (1) there are more counties preferred by the President only, and (2) as stated in Proposition 1, those counties are the ones that the President targets. The difference between states is greater when the county is represented by an aligned House Representative, because these counties have a higher weight in the President's objective function.

The model also has implications regarding federal-to-state transfers. As stated in Proposition 1, transferring federal funds to the President's preferred counties in the group $F_i \cap S_i$ is dominated by the strategy of transferring to the State i. Since the number of counties preferred by both the President and the State is greater for the aligned State 1, that state will receive more federal transfers than the non-aligned State 2. Essentially, the President is more willing to delegate the allocation of funds to State 1 with whom he has more in common.[15] This is formalized in the following corollary:

Corollary 2. *Federal transfers to State 1 are greater than to State 2* ($T_1^S > T_2^S$) *if the endowments of both states are equal* ($\widetilde{B}^1 = \widetilde{B}^2$).

Proof. See appendix A. ■

[15] This result is consistent with the findings of Larcinese et al. (2006) in which federal government transfers more funds to aligned states.

1.5 Data and econometric specification

1.5.1 Data

The Census of Governments provides reliable and comparable data on the distribution of Federal expenditures. It collects data on Government spending at five year intervals throughout the U.S. I use the years 1982, 1987, 1992, 1997 and 2002, providing county level data for around 3100 counties. The dependent variable for my analysis is the sum of federal transfers to all local governments inside the county, as a percentage of county personal income (from the Census Bureau). Importantly, the data allows me to identify whether federal funds go directly to any local governments inside the county (federal to county transfers), or indirectly through the state (federal to state transfers).[16]

To what extent are federal to county transfers discretionary, as opposed to strictly formula based? In Appendix B, I study this question in detail, using techniques from the literature to measure the extent of discretion. In particular, Levitt and Snyder (1995, 1997) and Berry et al. (2010) argue that variability in spending provides evidence of discretion, and I show that the variable I use displays more variance than even the highly discretionary programs from CFFR. In the Appendix, I also propose an alternative, more stringent test for measuring the variability of federal programs

[16] Some previous studies have used data from the Consolidated Federal Fund Report (CFFR hereafter). This data details federal transfers by programs and recipients every year, but one cannot identify whether those funds go directly to a locality through federal agencies, or indirectly through state agencies. This distinction is crucial for my study. Another advantage of using data aggregated across programs is that federal programs from an integrated and complex federal budget are often linked, so using aggregate data controls for this correlation, avoiding the simultaneous equation bias that might arise if specific programs were studied instead.

and show that the variable I use appears highly discretionary based on this test as well.

Other data used here include controls that are standard in the public finance literature (see Appendix C for detailed sources). I use county level income per capita, black population, population under 18, population over 65, total population, and presidential elections statistics, all from the Census Bureau. The information about Congressional districts was collected from the Atlas of Congressional Districts, taking into account the changing district boundaries. I also use voting data about Governors, state legislatures, and US House Representatives from multiple sources described in Appendix C.

1.5.2 Econometric specifications

Based on Corollary 1, I estimate the difference in federal transfers to counties in aligned vs. non-aligned states depending on whether the county is represented by a House Representative from the President's party ("aligned counties"). I present two econometric models. In the first one, I do not try to identify which counties are "preferred" by the President, i.e., these could be either partisan or swing counties. In the second one, I explicitly study which of these two groups drives the results.

The first econometric specification is as follows:

$$T^C_{jit} = \alpha + \beta_{FS} FS_{it} + \beta_{FC} FC_{jit} + \beta_{FS\times FC} FS_{it} \times FC_{jit} + \mathbf{X'_{jit}b} + \beta_{pos} pos_{jit} + \beta_{close} close_{jit} + \mathbf{D_t} + \mathbf{u_j} + e_{jit} \quad (3)$$

Here, T^C_{jit} is federal transfer to county j in State i during year t, $\mathbf{X'_{jit}}$ are various time varying controls (natural log of real income per capita, percentage of blacks, percentage of people under 18, percentage of people over 65 and natural log of population) and FC and FS are political alignment dummy variables. Namely, FS is an indicator that represents federal-state political alignment for the current and the previous two years.[17] Based on the discussion in Section 3, this variable takes a value of 1 if the party that controls the state budget is the same as the President's party. Similarly, FC is an indicator that represents federal-county alignment for the current and the previous two years. It takes a value of 1 if the congressional district in which county j lies has a US House Representative from the same party as the President.[18] The variables *pos* and *close* are indicators of the last presidential election vote margin. The former takes a value of 1 if the vote margin was higher than 0.10, and the latter takes a value of 1 if the margin was between -0.10 and 0.10.[19] These variables are included because of the potential correlation between alignment categories and previous electoral vote margins, in which case excluding them could lead to an omitted variable bias. The specification also includes fixed effects: time fixed effects ($\mathbf{D_t}$) are used to control for country-wide effects, such as the political and economic environment at the federal level, and county fixed effects ($\mathbf{u_j}$) control for time-invariant unobserved heterogeneity at the county level, such as the number of local government

[17] My results below are virtually unchanged if I use the previous two years (ignoring the current year).

[18] If the county is divided into many congressional districts, as it happens with highly populated counties, I categorize the county as being aligned with the President if at least 70% of its House Representatives are from the President's party. In section 7, I show that the results are robust if I exclude these cases from the analysis.

[19] Margin is a continuous variable taking values between [-1,1]. For example, if the president is a Democrat and 55% of the electorate in county j voted for Democrats and 45% for Republicans, the margin will be 0.10. However, if the President were Republican, the margin would have been -0.10.

units within each county, or urban vs. rural areas where the President might have different political incentives.

Based on the prediction of the model in Section 4, we expect the difference-in-difference estimator $\beta_{FS\times FC}$ in (3) to be negative. As stated in Corollary 1, $\beta_{FS\times FC} = (\frac{n_2}{(n_2+m_2)} - \frac{n_1}{(n_1+m_1)})(T^C_{\sim a} - T^C_a) < 0$. This means that the change in federal transfers when the State becomes non-aligned with the President (changing the party that controls the state budget) has to be greater, on average, for aligned counties than for non-aligned ones.

By not conditioning the difference-in-difference estimate $\beta_{FS\times FC}$ on "preferred" counties, equation (3) is likely to provide an underestimate of the true effect. This is the average effect between the President's preferred and non-preferred counties. Based on the theory, the effect should only be present among the preferred counties.

In the second econometric model, I investigate which counties, partisan or swing, are more likely to be preferred. For example, if preferred counties are the partisan counties, we expect the difference-in-difference to be stronger for this group. I incorporate in equation (3) the effect of partisanship on the change of federal to county transfers due to changes in alignments by fully interacting the alignment variables with the presidential vote margin categories: negative partisan (margin below -10%), swing (margin between -10% and 10%) and positive partisan (margin above 10%). I run the following regression:

$$\begin{aligned} T^C_{jit} = {} & \alpha + \beta_{FS} FS_{it} + \beta_{FC} FC_{jit} + \beta_{FS\times FC} FS_{it} \times FC_{jit} + \beta_{pos} pos_{jit} + \beta_{close} close_{jit} \\ & + \beta_{FS\times pos} FS_{it} \times pos_{jit} + \beta_{FS\times close} FS_{it} \times close_{jit} + \beta_{FC\times pos} FC_{jit} \times pos_{jit} \\ & + \beta_{FC\times close} FC_{jit} \times close_{jit} + \beta_{FS\times FC\times pos} FS_{it} \times FC_{jit} \times pos_{jit} \\ & + \beta_{FS\times FC\times close} FS_{it} \times FC_{jit} \times close_{jit} + \mathbf{X}'_{\mathbf{jit}}\mathbf{b} + \mathbf{D_t} + \mathbf{u_j} + e_{jit}, \end{aligned} \tag{4}$$

where pos_{jit} stands for positive partisan, $close_{jit}$ for swing, and the excluded category is negative partisan counties.

1.6 Main results

In this section I present the main empirical findings of the paper. In Table 1, I regress federal to county transfers on federal-state and federal-county alignment, and the time varying covariates listed under Equation (3). In Column (1) and (2), we see that federal transfers to counties are not significantly affected by the alignment between the President and the party that controls the state (FS) or by the alignment between the President and the House Representative of a county (FC). However, the coefficients in both regressions have the correct sign. Namely, in Column (1) federal transfers to counties are 2.5 percentage points smaller inside aligned states, suggesting that the President has more interests in targeting counties within non-aligned states compared with aligned ones. Transfers are 1 percentage point greater when the county is aligned with the President, as seen in Column (2). The latter result is in line with the findings in Berry et al. (2010), where an aligned Representative with the President receives more federal funds for his district.

The estimation of equation (3), presented in Column (3), controls for the differential effect between aligned and non-aligned states on federal transfers to aligned and non-aligned counties. Consistent with the model, I find that the President targets spending towards counties represented by an aligned Representative more within non-aligned states. The coefficient estimate $\hat{\beta}_{FS}$ is almost zero, which means that the transfer to a non-aligned county does not change if the State changes from non-aligned to aligned with the President. Instead, when this difference is conditional on aligned counties, the transfer decreases by 5.6 percentage points as shown by the linear combination $\hat{\beta}_{FS} + \hat{\beta}_{FS\times FC}$ in panel B. This finding is explained by Proposition 1: There is no incentive to spend in aligned counties within aligned states, since that would simply crowd out similar spending by the State. There is, however, an incentive to spend in aligned counties within non-aligned states.

Based on Corollary 1, the precision of the estimation can be increased by conditioning on counties preferred by the President. For that purpose, Table 2 presents the results from estimating equation (4). The table shows the estimators conditional on positive partisan (in panel A: margin $>$ 0.10), swing (in panel B: margin$\leq |0.10|$) and negative partisan counties (in panel C: margin $<$ -0.10). For each case, I report the estimates for the change in federal transfers to an aligned county when the state changes from non-aligned to aligned with the President (first row in all panels of Table 2), the change in federal transfers to a non-aligned county when the state changes from non-aligned to aligned (second row in all panels of Table 2), and the difference between these two changes (reported in the third row in all panels of Table 2).

In the first row of Panel A, federal transfers to an aligned partisan county are 5.9 percentage points smaller when the state changes from non-aligned to aligned with the President. The second row of the same panel shows the same difference when the county is not aligned with the President. This estimate is close to zero and insignificant, indicating that the President does not change the allocation of resources to non-aligned counties when the state's political alignment changes. As the third row of the panel shows, the two estimates are significantly different from each other, indicating that there is a difference in the President's behavior regarding aligned and non-aligned partisan counties.

As panel B shows, I do not find similar differences for swing counties (although the coefficients have the expected sign). This suggests that, in this context, the President has a preference towards targeting aligned partisan counties rather than swing counties.

In Panel C, the estimate $\hat{\beta}_{FS}$ indicates a significant effect for aligned negative partisan counties. However, as the last row shows, there is no evidence of a difference between aligned and non-aligned counties in this case. Given the small number of aligned negative partisan counties, results for this category should be interpreted with care.[20]

The results in this section are in line with the theoretical model. The President has an incentive to allocate funds strategically to aligned counties only within non-aligned states. This effect appears to be stronger within partisan counties, suggesting that these might be the counties viewed as politically valuable by the President in this setting.

1.7 Robustness checks

In this section, I explore the robustness of the above results by estimating specifications (3) and (4) on different sub samples, by changing how the dependent or independent variables are measured, and by controlling for various sources of unobserved heterogeneity.

[20] For aligned states (i.e., $FS = 1$), we have 190 negative partisan counties, compared with 360 swing and 990 positive partisan counties. For aligned counties (i.e., $FC = 1$), we have 250 negative partisan counties, compared with 760 swing and 2400 positive partisan counties. For aligned counties inside aligned states (i.e., $FS \times FC = 1$) we have 31 negative partisan counties, compared with 100 swing and 570 positive partisan counties.

1.7.1 Redefining the dependent variable: Federal transfers in per capita terms

The dependent variable throughout this study is federal-county transfers as a percentage of county personal income. If income can also fluctuate due to political cycles, the dependent variable might have an unclear interpretation because every time the federal government changes transfers due to political alignment, both the numerator and the denominator will be moving in the same direction. As a robustness check, I use real federal transfers (prices of 2000) in per capita terms as the dependent variable.[21]

The results can be seen in Column (2) of Tables 3 and 4. In Column (2) of Table 3, we see that the main results do not change, although the difference in difference became non-significant.[22] Column (2) of Table 4 separates partisan and swing counties, and shows that the results are qualitatively the same as above. I find a significant and negative difference-in-difference for partisan counties but not for swing counties.

1.7.2 Alternative party control of the state definition

The party in control of the state can be defined in alternative ways (see Section 3). Above, I have used the measure proposed by Ansolabehere and Snyder (2006). In this section, I change that definition slightly to show that the main results are not sensitive to changes in the way of defining party control of the state.

[21] The drawback of this variable compared to income is that people can move due to public good provision as in the Tiebout sorting model.

[22] This could be explained by Tiebout sorting. It could also be due to an attenuation bias because a linear extrapolation was used to get population at 5 year intervals from the decennial census.

A governor's veto power is an important element of control over the state budget. However, a veto can in some cases be overturned by two thirds of the legislators. The greater the share of co-partisan legislators, the smaller the probability of overturning a Governor's veto, and the more likely that the governor's preferences will determine the budget. To capture this, I use the following definition of party control: if the Governor's party has a simple majority in one of the legislative chambers and holds at least one third of the seats in the other chamber, then the state is controlled by the Governor's party. Intuitively, a veto overturn is unlikely in this case since the legislature needs more than two thirds in both chambers for overturning a Governor's veto. I use this new definition to construct the federal-state alignment variable, and re-estimate equations (3) and (4).

The results can be seen in Column (3) of Table 3 and 4. As we can see in Table 3, the results change little, with the difference in difference increasing somewhat in absolute value for this new definition of party control of the state. Column (3) of Table 4 shows a similar pattern: the finding of a differential effect for partisan counties is reinforced compared to the measure used earlier.

1.7.3 Addressing unobserved heterogeneity at state level

The results could be subject to an omitted variable bias if federal transfers to counties were correlated with unobserved state-time varying covariates. One example of this is federal to state transfers. Since these are potentially endogenous, controlling for them would not be appropriate. Using state-time fixed effects will address this and other potential state-time level heterogeneity.

In Column (4) of Table 3 and 4, we can see the estimation of equations (3) and (4), respectively, once these fixed effects are included. In Column (4) of Table 3, as before, aligned counties receive significantly higher transfers only in non-aligned states. In aligned states, the effect of federal-county alignment is negative and not statistically significant.[23] In Column (4) of Table 4, the estimates change little, and the difference in difference for partisan counties is still significant. Note that there is a considerable loss in degrees of freedom in these regressions due to the inclusion of around 250 new fixed effects. Based on these results, state level heterogeneity does not appear to affect the main findings of the paper.

1.7.4 Elected council-executive counties

There are three basic forms of county government: Commission, Administrator and Council-Executive. The last one differs from the others in that the executive is independently elected by county voters instead of being appointed by a council or commission board. The county board remains the legislative body, but in this case the executive can veto ordinances enacted by the commission. The county executive has as much power as a mayor-council in a strong municipality or city. For counties with such a strong executives the President might care about the party of the executive more than about the party of the House Representative.

I am not aware of any dataset that would contain the party affiliation of the county executives or the date this form of governments was first introduced in each county. However, the National Association of Counties (NACO) identifies which counties currently have this form of government. In Column (5) of Table 3, I drop these 400 counties and re-estimate the model. The estimator $\hat{\beta}_{FS\times FC}$ is still negative but not significant. Nevertheless, the linear combination $\hat{\beta}_{FS} + \hat{\beta}_{FS\times FC}$ is significant and negative, and $\hat{\beta}_{FS}$ is close to zero, just like in Column (1). When we control for

[23] Although not significant, the sign of $\hat{\beta}_{FS\times FC}$ remains unchanged.

partisan and swing counties as shown in Column (5) of Table 4, the results are very similar to the ones shown in Column (1). These results reinforce the main findings of the paper. They also suggest that either the organizational form of the counties and the party affiliation of their executives are not correlated with the party affiliation of the House Representative, or that, even council-executive counties, the President cares more about the party of the House Representative.

1.7.5 States with only one congressional district

If a state has only one congressional district, this increases the correlation between the federal-state and the federal-county alignment measures. If we assume the extreme case in which all the states have only one congressional district as large as themselves, then neither the model of equation (3) nor equation (4) would be identified. Even though the situation is away from this extreme case, there are states with one congressional district that increases the correlation between those two measures of alignment. This could reduce the significance of the individual parameter estimates, while still resulting in significant linear combinations like in panel B of Table 1 and Table 2. In order to see whether the results are affected by the correlation between FS and FC, I drop from the sample the states with only one congressional district (Alaska, Delaware, North Dakota, Vermont, Wyoming and South Dakota). These results are in Column (6) of Table 3 and 4, and the estimates are very similar to the ones obtained earlier. Hence, we can conclude that states with only one congressional district are not driving the results found above.

1.7.6 Multi congressional districts counties

The most populous counties are divided into many congressional districts. Since the unit of observation is the county, federal-county alignment could be measured in different ways in these cases. For the estimates above, I defined a multi-district county as being aligned if at least 70% of the House Representatives were aligned with the President. To check whether these counties are biasing the results I drop them from the sample. The result of re-estimating the specifications in this manner are in Column (7) of Tables 3 and 4. In Table 3, the difference-in-difference estimate is significant at 10% as in the main estimation of Column (1) with an increase in absolute value. In Table 4, the results change very little compared with estimation in Column (1) of the same table. The definition of alignment for counties with multiple congressional districts does not drive the findings above.

1.7.7 Controlling for state transfers to localities

Since state-county transfers are endogenous based on the model, an instrument is required in order to include them in the regression. Here I instrument state transfers to county j with the average transfer inside the congressional district where county j lies, but without county j. Formally, for a congressional district l, I estimate

$$t^C_{jlit} = a + \phi \left(\frac{1}{R^l - 1} \sum_{k \neq j}^{R^l} t^C_{klit} \right) + \mathbf{W}'_{\mathbf{jlit}}\mathbf{c} + \mathbf{D_t} + \mathbf{u_j} + \varepsilon_{jlit} \quad (5)$$

$$T^C_{jlit} = \alpha + \eta t^C_{jlit} + \mathbf{W}'_{\mathbf{jlit}}\mathbf{d} + \mathbf{D_t} + \mathbf{u_j} + e_{jlit} \quad (6)$$

where $\mathbf{W_{jlit}} = (FS_{it}, FC_{lit}, FS_{it} \times FC_{lit}, \mathbf{X_{jlit}}, pos_{jlit}, close_{jlit})'$, $\mathbf{d} = (\beta_{FS}, \beta_{FC}, \beta_{FS \times FC}, \mathbf{b}, \beta_{pos}, \beta_{close})'$, and t^C_{jlit} represents the state transfer to county j, which lies within congressional district l, inside state i, during year t. Equation (5) represents the first stage of a just identified system of equations composed by (5) and (6), where the excluded instrument for state to county j transfers is $\left(\frac{1}{R^l-1}\sum_{k\neq j}^{R^l} t^C_{klit}\right)$ because it is less likely that e_{jlit} is correlated with $\left(\frac{1}{R^l-1}\sum_{k\neq j}^{R^l} t^C_{klit}\right)$ than with t^C_{jlit}. Equation (4) can be instrumented in a similar manner.

The results are in Column (8) of Table 3 and 4. Since the instrument cannot be constructed if the county is divided in multiple congressional districts, I exclude these counties from the regression.[24] As we can see, the instrument is fairly strong. In the bottom panel of Table 3 and Table 4, the F-statistic of the first stage is higher than 55 in both cases, the adjusted R^2 of the first stage regression is around 0.36, and the coefficient of the instrument is significant at 1%.

The estimated coefficient on state-county transfers is close to zero, while comparing Column (7) and (8) in both tables shows little change in the coefficient estimates $\hat{\beta}_{FS}$, $\hat{\beta}_{FC}$ and $\hat{\beta}_{FS\times FC}$. This reinforces the validity of the OLS estimates presented above.

1.8 Conclusion

To this point, scholars have been studying the political allocation of federal resources without considering the involvement of state governments. Because state governments allocate resources based on some of the same considerations, a strategic federal government should take this into account. Controlling for this fact using party alignment between these two layers of governments, I have found that the President skews the distribution of funds towards counties whose House Representatives are from the

[24] The IV estimates from Column (8) can be compared with the OLS estimates of Column (7) because the sub-samples are the same.

President's party, but only within non-aligned states. Specifically, federal transfers to such counties decrease by around 6 percentage points when the party that controls the state becomes aligned with the President. Consistent with my model, no effect has been found for counties whose House Representatives are not from the President's party. This demonstrates the importance of controlling for the three-way political alignment between county, state, and federal government when studying the determinants of intergovernmental spending. The finding that these interaction terms matter survives a long list of robustness checks, as shown in Section 7 above.

This paper has important implications for normative studies of decentralization. My results suggest that in a highly decentralized federal system such as the US, the federal government might engage in a sort of competition with non-aligned states for mobilizing voters, while cooperating with states that are politically aligned with it. Understanding the welfare impact of the strategic interaction between different layers of governments is outside the scope of this paper, but my findings do imply that taking this interaction into account is important for welfare analysis.

The standard view of decentralization is that it removes political power from the center. The findings in this paper indicate the presence of an offsetting effect. After decentralization, a strategic central government may be able to rely on some local governments to further his political goals, and could concentrate more direct spending on those areas where his power has declined. The ultimate impact on the central government's de facto power may be ambiguous.

Appendix A. Proof of propositions of section 4

I solve the model using Backward Induction. In the second stage, each State $i = 1, 2$ maximizes the following Lagrangian:

$$L_i = \sum_{j \in (F_i \cap S_i)} H\left(\theta T_{ij}^C + t_{ij}^C\right) + \sum_{j \in (S_i \backslash (F_i \cap S_i))} H\left(t_{ij}^C\right) + \mu_i(\tilde{B}^i + T_i^S - \sum_{j \in S_i} t_{ij}^C), \text{ for } i = 1, 2$$

The first order conditions are:

$$H'(\theta T_{ij}^C + t_{ij}^C) = \mu_i, \text{ for all } j \in (F_i \cap S_i) \tag{1}$$

$$H'(t_{ij}^C) = \mu_i, \text{ for all } j \in (S_i \backslash (F_i \cap S_i)) \tag{2}$$

$$\tilde{B}^i + T_i^S = \sum_{j \in S_i} t_{ij}^C \tag{3}$$

working with (1), (2) and (3) yields state i's reaction functions:

$$t_{ij}^{C*} = \frac{1}{m_i + r_i}\left[\widetilde{B}^i + T_i^S + \theta \sum_{l \in (F_i \cap S_i)} T_{il}^C\right] - \theta T_{ij}^C, \text{ for all } i = 1, 2 \text{ and } j \in (F_i \cap S_i) \tag{4}$$

$$t_{ij}^{C*} = \frac{1}{m_i + r_i}\left[\widetilde{B}^i + T_i^S + \theta \sum_{l \in (F_i \cap S_i)} T_{il}^C\right], \text{ for all } i = 1, 2 \text{ and } j \in S_i \backslash (F_i \cap S_i) \tag{5}$$

Given this, the Lagrangian for the President's maximization problem in the first stage is given by

$$L^P = \sum_{i=1}^2 \left(\sum_{j \in fs_i} \gamma H\left(\theta T_{ij}^C + t_{ij}^{C*}\right) + \sum_{j \in f_i} \gamma H\left(\theta T_{ij}^C\right) + \sum_{j \in ((F_i \cap S_i) \backslash fs_i)} H\left(\theta T_{ij}^C + t_{ij}^{C*}\right) + \right.$$

$$\left. \sum_{j \in \{((F_i \backslash (F_i \cap S_i)) \backslash f_i\}} H\left(\theta T_{ij}^C\right) \right) + \sum_{i=1}^2 \left(\sum_{j \in F_i} v_{ij} T_{ij}^C + v_i T_i^S \right) +$$

$$\lambda \left(\widetilde{B}^F - \sum_{i=1}^2 \left(\sum_{j \in F_i} T_{ij}^C + T_i^S \right) \right),$$

The first order conditions for maximization are,

$$L^P{}_{T_{ij}^C} = 0: \; H'(\theta T_{ij}^C + t_{ij}^{C*}) \left(\frac{\theta}{m_i + r_i} - \theta \right) + \sum_{l \in ((F_i \cap S_i) \backslash fs_i), l \neq j} H'(\theta T_{il}^C + t_{il}^{C*}) \frac{\theta}{m_i + r_i} +$$

$$\sum_{l \in fs_i} H'(\theta T_{il}^C + t_{il}^{C*}) \frac{\gamma \theta}{m_i + r_i} - \lambda + v_{ij} = 0, \text{ for all } i = 1, 2 \text{ and } j \in ((F_i \cap S_i) \backslash fs_i) \tag{6}$$

$$L^P{}_{T_{ij}^C} = 0: \; H'(\theta T_{ij}^C + t_{ij}^{C*}) \left(\frac{\gamma \theta}{m_i + r_i} - \gamma \theta \right) + \sum_{l \in fs_i, l \neq j} H'(\theta T_{il}^C + t_{il}^{C*}) \frac{\gamma \theta}{m_i + r_i} +$$

$$\sum_{l \in ((F_i \cap S_i) \backslash fs_i)} H'(\theta T_{il}^C + t_{il}^{C*}) \frac{\theta}{m_i + r_i} - \lambda + v_{ij} = 0, \text{ for all } i = 1, 2 \text{ and } j \in fs_i \tag{7}$$

$$L^P{}_{T_{ij}^C} = 0 : \theta H'(\theta T_{ij}^C) - \lambda + v_{ij} = 0, \text{ for all } i = 1,2 \text{ and } j \in ((F_i \backslash (F_i \cap S_i)) \backslash f_i \quad (8)$$

$$L^P{}_{T_{ij}^C} = 0 : \theta\gamma H'(\theta T_{ij}^C) - \lambda + v_{ij} = 0, \text{ for all } i = 1,2 \text{ and } j \in f_i \quad (9)$$

$$L^P{}_{T_i^s} = 0 : \sum_{l \in ((F_i \cap S_i) \backslash fs_i)} H'(\theta T_{il}^C + t_{il}^{C*}) \frac{1}{m_i + r_i} + \sum_{l \in fs_i} \gamma H'(\theta T_{il}^C + t_{il}^{C*}) \frac{1}{m_i + r_i} - \lambda + v_i = 0,$$

for $i = 1, 2$ (10)

$$L^P{}_\lambda = 0 : \widetilde{B}^F - \textstyle\sum_{i=1}^2 (\sum_{j \in F_i} T_{ij}^C + T_i^S) = 0, \text{ for } i = 1,2 \quad (11)$$

Lemma 1: $T_{ij}^C = 0$ for all $j \in (F_i \cap S_i)$, $T_{ij}^C > 0$ for all $j \in (F_i \backslash (F_i \cap S_i))$, $T_i^S > 0$ for all $i = 1, 2$ is an equilibrium.

Rewriting conditions (4) to (11) by imposing the restrictions in Lemma 1 shows that the first order conditions hold, we therefore have an equilibrium.

Using Lemma 1 to rearrange conditions (8) and (9) yields the following,

$T_{ij}^C \equiv T_{\sim a}^C$ for all $i = 1, 2$ and $j \in ((F_i \backslash (F_i \cap S_i)) \backslash f_i$; $T_{ij}^C \equiv T_a^C$ for all $i = 1, 2$ and $j \in f_i$; and $T_a^C > T_{\sim a}^C$ (12)

Lemma 1 combined with condition (12) verifies Proposition 1 and Corollary 1.

Lemma 2: $T_{il}^C > 0$ for any $l \in (F_i \cap S_i)$, or $T_{il}^C = 0$ for any $l \in F_i \backslash (F_i \cap S_i)$ cannot be an equilibrium.

One can easily verify that rewriting conditions (4) to (11) based on the restrictions imposed in Lemma 2 will lead to a contradiction. Thus, Lemma 2 shows the uniqueness of the equilibrium stated in Proposition 1.

Rewrite (8), (9) and (10) based on Lemma 1 and condition (12) to get

$$\frac{m_2 + r_2}{m_1 + r_1} (\widetilde{B}^1 + T_1^S) \frac{H'^{-1}(\frac{(1-\alpha)m_1}{m_1 + r_1} + \gamma \frac{\alpha m_1}{m_1 + r_1})}{H'^{-1}(\frac{(1-\alpha)m_2}{m_2 + r_2} + \gamma \frac{\alpha m_2}{m_2 + r_2})} - \widetilde{B}^2 = T_2^S.$$

It is easy to see that $T_1^S > T_2^S$ for $\widetilde{B}^1 = \widetilde{B}^2$, since $m_1 > m_2$ and $r_1 < r_2$. This proves Corollary 2.

Appendix B. The discretionary nature of federal transfers to counties

Berry et al. (2010), among others, used data from CFFR. To separate broad-based entitlement programs from federal programs that represent discretionary spending, Levitt and Snyder (1995, 1997) and Berry et al. (2010) calculate coefficients of variation for each program and they separate them into two categories: low and high variability programs (using as threshold a coefficient of variation of 3/4), because they assume that high variability represents discretionary spending. Unfortunately, I cannot follow the same methodology since the data from the Census does not allow me to identify each source of spending individually. However, I can compare the data from Census of Governments with high-variability programs from CFFR to show that the former is highly discretionary as well.

In Table B1 column (4) we can see that the coefficient of variation associated with Federal to county transfers is 1.45, by far higher than the threshold 3/4 proposed by Levitt and Snyder (1995). The composition of federal to county transfers is detailed in Column (1).[25] There, we can see that almost half of it, on average, is composed by Housing and community development, a highly discretionary set of programs based on the coefficient of variation.[26] Education is the second highest component of federal

[25] The data I am using from the Census of Government does not allow me to identify each source of spending individually at county level. However, each source can be observed aggregated at state level. That is to say the sum of all the federal to county transfers inside the state divided by program, which is what I am using to calculate the shares in column 1.

[26] The magnitudes of federal to county transfers cannot be compared with federal funds from CFFR because the former only accounts for direct transfers to localities, while the second one contains both direct and indirect transfers.

transfers to counties at 19%, also fairly discretionary. Health and Highways are the third and fourth, with 4% and 3% respectively, and these are unlikely to exert much influence overall.

A high coefficient of variation may not be due to discretion, but instead to large demographic or economic changes in a county during a period of time. If this were the case, the coefficient of variation would mistakenly indicate that the program is highly discretionary when it is not. In order to address this potential issue, I will compare the variance of the residual that comes from a regression of each program against all the observable demographic and economic characteristics with the variance of the program itself. To compute the former, I estimate

$$y_{jit} = \alpha + \mathbf{X}'_{\mathbf{jit}}\beta + \mathbf{D}'_{\mathbf{i}}\mathbf{D}_{\mathbf{t}} + \mathbf{u}_{\mathbf{j}} + e_{jit}, \tag{B1}$$

where y_{jit} is a given federal outlay in county j within State i in year t as a percentage of personal income; $\mathbf{X}'_{\mathbf{jit}}$ is a matrix of demographic and economic county level-time varying controls (natural log of real income per capita, % of blacks, % under 18 years old, % over 65 years old and natural log of population); $\mathbf{D}'_{\mathbf{i}}\mathbf{D}_{\mathbf{t}}$ captures state by state level heterogeneity per year and $\mathbf{u}_{\mathbf{j}}$ is a county fixed effect that captures unobserved fixed heterogeneity; and e_{jit} is the residual.

If the ratio $v\hat{a}r(\hat{e}_{jit})/v\hat{a}r(y_{jit})$ is close to one, it means that the model did not absorb much variation of y_{jit}, in which case demographic and economic changes did not explain the variability, hence the program could be considered as highly discretionary. The opposite is concluded if that ratio is close to zero.

The results can be seen in Column (6) of Table B1, federal to county transfers are not less discretionary than the variables used in previous studies, detailed in Column (8). Even more, it is at least as discretionary as all of them except for highway programs.

Appendix C. Data sources

All the data comes from the *Census Bureau - USA Counties*, unless indicated.

http://www.census.gov/support/USACdataDownloads.html#INC

Intergovernmental transfers from Federal government to Counties. *U.S. Census Bureau - USA Counties, Census of Governments* (1982, 1987, 1992, 1997, 2002).

Intergovernmental transfers from State government to Counties. *U.S. Census Bureau - USA Counties, Census of Governments* (1982, 1987, 1992, 1997, 2002).

Regional Consumer Price Index (CPI) for all urban consumers, not seasonally adjusted. Yearly value obtained by averaging across months. *U.S Department of Labor: Bureau of Labor Statistics.*

Personal Income. *Bureau of Economic Analysis - USA Counties.*

Percentage of Blacks. *Race Data, U.S. Census Bureau - USA Counties.*

Percentage of People Under 18. *Age, U.S. Census Bureau - USA Counties.*

Percentage of People Over 65. *Age, U.S. Census Bureau - USA Counties.*

Population. *U.S. Census Bureau - USA Counties.*

Presidential election Outcomes, Democrat and Republican vote share. *CQ Press - USA Counties.*

Matched Counties with Congressional district and Redistricting. *Congressional District Atlas: 95th to 109th Congress.*

President, Governors, and United States House Representatives' Parties. *Library of Congress Web Archive; OurCampaigns.com.*

State legislative seats held by each party. *Burnham, W Dean, "Partisan Division of American State Governments, 1834-1985". ICPSR Study No. 00013; Council of State Governments, Book of the States.*

Elected council-executive counties. *National Association of Counties (NACO).*

Table 1: Federal-County transfers conditional on State and County alignment. Estimation of equation (3)

Panel A: Estimation Results		(1)	(2)	(3)
	Estimators			
	$\hat{\beta}_{FS}$	-0.025		-0.007
		[0.017]		[0.015]
	$\hat{\beta}_{FC}$		0.010	0.021**
			[0.011]	[0.009]
	$\hat{\beta}_{FS\times FC}$			-0.049*
				[0.029]
	Observations	15,067	15,054	15,054
	R2 within	0.180	0.179	0.181
	Number of counties	3,071	3,071	3,071
Panel B: Linear combination of estimators				
(1)	$\hat{\beta}_{FS} + \hat{\beta}_{FS\times FC}$			-0.056**
				[0.027]
(2)	$\hat{\beta}_{FC} + \hat{\beta}_{FS\times FC}$			-0.027
				[0.029]

Notes: This table shows how federal to county transfers increase within non-aligned states compared to aligned ones. In panel A column 1 and 2 I estimate the effect of state and county alignment on federal to county transfers. Panel A column 3 shows the result of estimating equation (3). In Panel B column 3, row (1) shows the difference in Federal transfers to an aligned county between aligned and non-aligned states. The row (2) shows the difference within an aligned state between an aligned and a non-aligned county. All regressions include county and year fixed effects, as well as the natural log of income per capita, natural log of population, % of blacks, % of inhabitants younger than 19 and % of inhabitants older than 65. The highest 2% values of the dependent variable were considered outliers and dropped from the sample. Robust standard errors clustered at the state level are reported in parenthesis. *, ** and *** denote 10, 5 and 1% level of significance, respectively.

Table 2: Federal-County transfers conditional on State and County alignment as well as partisan or swing counties. Linear combination of estimators from the estimation of equation (4)

(1)		**(2)**
Panel A: Positive partisan (margin $>$ 0.1)		
(1)	$\hat{\beta}_{FS} + \hat{\beta}_{FS\times FC} + \hat{\beta}_{FS\times pos} + \hat{\beta}_{FS\times FC\times pos}$	-0.059*
		[0.032]
(2)	$\hat{\beta}_{FS} + \hat{\beta}_{FS\times pos}$	0.018
		[0.018]
(3)	$\hat{\beta}_{FS\times FC} + \hat{\beta}_{FS\times FC\times pos}$	-0.077**
		[0.037]
Panel B: Swing (margin $\leq \lvert 0.1 \rvert$)		
(1)	$\hat{\beta}_{FS} + \hat{\beta}_{FS\times FC} + \hat{\beta}_{FS\times close} + \hat{\beta}_{FS\times FC\times close}$	-0.044
		[0.035]
(2)	$\hat{\beta}_{FS} + \hat{\beta}_{FS\times close}$	-0.021
		[0.023]
(3)	$\hat{\beta}_{FS\times FC} + \hat{\beta}_{FS\times FC\times close}$	-0.024
		[0.044]
Panel C: [Omitted category] Negative partisan (margin $<$ -0.1)		
(1)	$\hat{\beta}_{FS} + \hat{\beta}_{FS\times FC}$	-0.016
		[0.076]
(2)	$\hat{\beta}_{FS}$	-0.046**
		[0.02]
(3)	$\hat{\beta}_{FS\times FC}$	0.031
		[0.082]

Notes: This table shows how federal to county transfers increase within non aligned states, compared to aligned ones, for three different categories of the last presidential electoral vote share. Each cell represents a linear combination of estimators obtained from estimating equation (3). The number of observations is 15,054, the number of counties is 3,071. R2=0.179. The highest 2% values of the dependent variable were considered outliers and dropped from the sample. Robust standard errors for the linear combinations clustered at the state level are reported in parenthesis. *, ** and *** denote 10, 5 and 1% level of significance, respectively.

Table 3: Robustness checks. Different subsamples and specifications. Estimation of equation (3)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Method:				OLS				IV
estimators								
$\hat{\eta}$								0.011
								[0.008]
$\hat{\beta}_{FS}$	-0.007	-1.715	0.008	-	-0.013	-0.020	-0.000	-0.000
	[0.015]	[3.316]	[0.011]	-	[0.017]	[0.015]	[0.016]	[0.012]
$\hat{\beta}_{FC}$	0.021**	3.471**	0.020**	0.013*	0.018*	0.023***	0.020**	0.014*
	[0.009]	[1.654]	[0.009]	[0.007]	[0.009]	[0.009]	[0.009]	[0.007]
$\hat{\beta}_{FS\times FC}$	-0.049*	-9.659	-0.054*	-0.034	-0.046	-0.036	-0.059*	-0.055***
	[0.029]	[6.003]	[0.028]	[0.024]	[0.031]	[0.029]	[0.030]	[0.017]
Observations	15,054	15,066	15,054	15,054	13,218	14,260	13,292	13,133
R-squared	0.181	0.110	0.180	0.217	0.167	0.181	0.173	0.167
Number of Counties	3,071	3,077	3,071	3,071	2,699	2,976	2,927	2,892
$\hat{\beta}_{FS}+\hat{\beta}_{FS\times FC}$	-0.056**	-11.37**	-0.046	-	-0.059**	-0.056*	-0.059**	-0.056***
	[0.027]	[5.668]	[0.028]	-	[0.028]	[0.029]	[0.027]	[0.013]
$\hat{\beta}_{FC}+\hat{\beta}_{FS\times FC}$	-0.027	-6.188	-0.034	-0.022	-0.028	-0.014	-0.039	-0.04
	[0.028]	[6.017]	[0.027]	[0.023]	[0.032]	[0.029]	[0.029]	[0.016]
First Stage R2								0.360
First Stage F test								70.12
First Stage excluded instrument coefficient								0.520***
Standard error								[0.015]

Note: This table shows the same result as in Table 2 in Column 1. In Column 2, I use federal transfers in per capita terms as the dependent variable. In Column 3 I use an alternative measure of party control of the State. In Column 4 I control for State level Heterogeneity by using State*year dummy variables. In Column 5 I eliminate from the sample those counties in which voters in a county elect a council-executive. In Column 6 I eliminate states with one congressional district. In Column 7 I eliminate counties divided in many congressional districts. In Column 8 I perform an IV estimation where State-County transfers are instrumented with the average of State-County transfers inside the district but outside the country. All regressions include county and year fixed effects, as well as the natural log of income per capita, natural log of population, % of blacks, % of inhabitants younger than 19 and % of inhabitants older than 65. The highest 2% values of the dependent variable were considered outliers and dropped from the sample. Robust standard errors clustered at the state level are reported in parenthesis. *, ** and *** denote 10, 5 and 1% level of significance, respectively.

Table 4: Robustness checks. Different subsamples and specifications. Estimation of equation (4)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Method:				OLS				IV
$\hat{\eta}$								-0.0006
								[0.006]
$\hat{\beta}_{FS\times FC} + \hat{\beta}_{FS\times FC\times pos}$	-0.078**	-14.438*	-0.091**	-0.053*	-0.074*	-0.065*	-0.091**	-0.085***
	[0.037]	[7.708]	[0.038]	[0.031]	[0.041]	[0.038]	[0.038]	[0.023]
$\hat{\beta}_{FS\times FC} + \hat{\beta}_{FS\times FC\times close}$	-0.024	-6.723	-0.031	-0.021	-0.016	-0.017	-0.032	-0.029
	[0.044]	[7.474]	[0.043]	[0.040]	[0.056]	[0.044]	[0.046]	[0.034]
$\hat{\beta}_{FS\times FC}$	0.031	3.29	-0.027	0.043	0.006	0.034	0.039	0.042
	[0.082]	[13.852]	[0.030]	[0.081]	[0.086]	[0.083]	[0.089]	[0.062]
Observations	15,054	15,066	15,054	15,054	13,218	14,260	13,292	13,133
R-squared	0.179	0.109	0.179	0.216	0.165	0.181	0.172	0.169
Number of Counties	3,071	3,077	3,071	3,071	2,699	2,976	2,927	2,892
First Stage R2								0.36
First Stage F test								56.05
First Stage excluded instrument coefficient								0.572***
Standard error								[0.021]

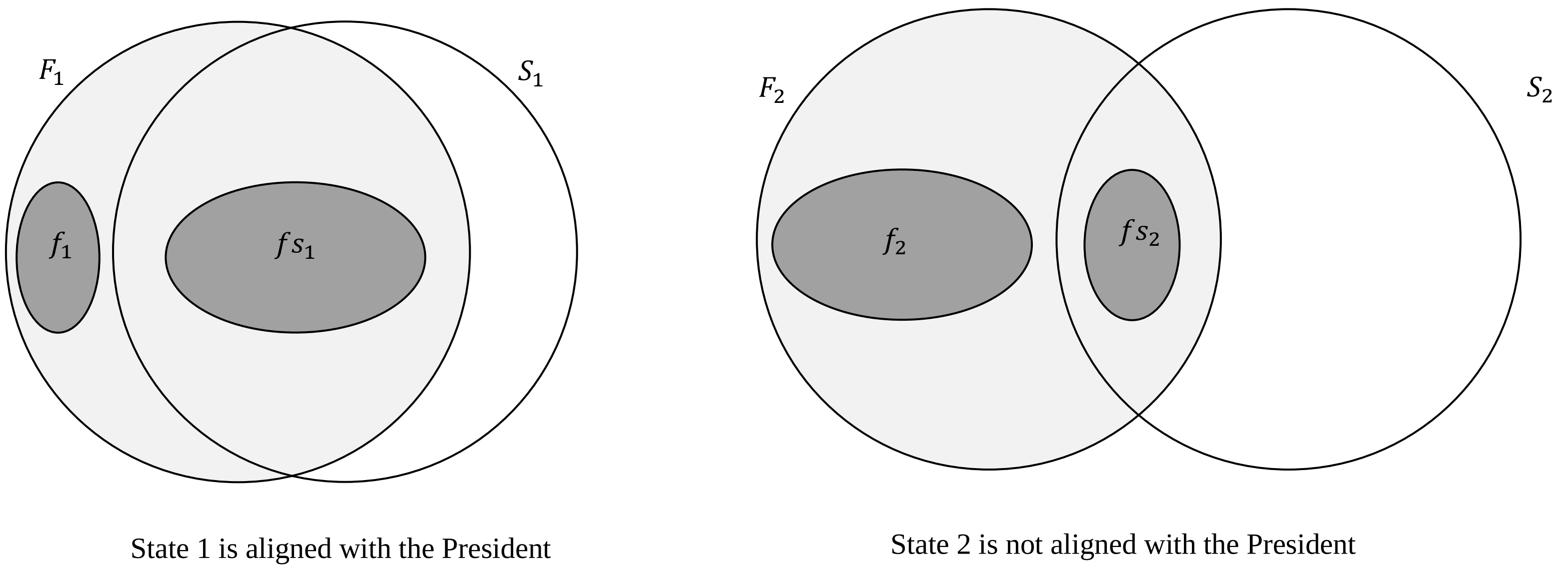
Note: This table shows the same result as in Table 3 in Column 1, but only for the linear combinations. In Column 2, I use federal transfers in per capita terms as the dependent variable. In Column 3 I use an alternative measure of party control of the State. In Column 4 I control for State level Heterogeneity by using State*year dummy variables. In Column 5 I eliminate from the sample those counties in which voters in a county elect a council-executive. In Column 6 I eliminate states with one congressional district. In Column 7 I eliminate counties divided in many congressional districts. In Column 8 I perform an IV estimation where State-County transfers are instrumented with the average of State-County transfers inside the district but outside the country. All regressions include county and year fixed effects, as well as the natural log of income per capita, natural log of population, % of blacks, % of inhabitants younger than 19 and % of inhabitants older than 65. The highest 2% values of the dependent variable were considered outliers and dropped from the sample. Robust standard errors clustered at the state level are reported in parenthesis. *, ** and *** denote 10, 5 and 1% level of significance, respectively.

Table B1. The discretionary nature of federal transfers to counties. Comparison with other transfers and programs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
variable y_{jit}		mean(y_{jit})	$\hat{var}(y_{jit})$	CV	$\hat{var}(e_{jit})$	$\frac{\hat{var}(e_{jit})}{\hat{var}(y_{jit})}$	Source	Used in previous studies by:
Federal-County transfers as % PI		0.50	0.52	1.45	0.24	0.46	Census of Governments	-
State-County transfers as % PI		4.53	5.71	0.53	0.73	0.13	Census of Governments	Frederickson and Cho (1974) Ansolabehere et al. (2002) Ansolabehere and Snyder (2006)
Federal-State transfers as % PI		3.33	1.43	0.36	0.13	0.09	Census of Governments	Grossman (1994) Ujhelyi (2013)
Federal funds on Health as % PI	4%	5.61	38.64	1.11	5.56	0.14	CFFR	Albouy (2013) Berry et al. (2010)
Federal funds on Education as % PI	19%	0.43	0.40	1.47	0.15	0.38	CFFR	
Federal funds on Highway as % PI (Dept. of transportation)	3%	0.52	1.24	2.13	1.01	0.82	CFFR	
Federal funds on Housing and Community development as % PI	49%	0.83	1.09	1.26	0.52	0.47	CFFR	

Note: Column 1 shows the composition of federal-county transfers as % of personal income (PI) by program. Column 2, 3 and 4 show simple means, variances and coefficient of variation, respectively. Column 5 presents the estimated variance of the residual that comes from regressing equation B1 using clustered errors at State level (for Federal-State transfer as % GSP robust standard errors were used, instead). Column 6 shows the ratio between $\hat{var}(e_{jit})$ and $\hat{var}(y_{jit})$ as a measure of variability of the federal program (close to 1 is high variability, close to 0 is considered low variability). Column 7 shows the sources where the federal funds come from. And Column 8 presents authors who used the mentioned variables in previous studies. For calculating columns 2, 3,4, 5 and 6 the highest 2% values of the dependent variable were dropped from the sample because of being considered outliers.

Figure 1: Graphic representation of counties preferred by the President and states.



F_i: set of counties preferred by the President.

S_i: set of counties preferred by the State *i*.

$(F_i \cap S_i)$: set of counties preferred by both the President and State *i*.

f_i: set of counties preferred only by the President, with an aligned Representative.

fs_i: set of counties preferred by both the President and State *i*, with an aligned Representative.

Chapter 2

Horizon Effect in Government Deficit: Evidence from Presidential Interludes in Latin America

2.1 Introduction

Positive studies on budget deficits and debt accumulation have argued that a government anticipating a possible defeat in the next election can use debt strategically in order to influence the policy of its successor. Such opportunistic behavior leads to an over-issuance of government debt relative to what is optimal (Persson and Svensson, 1989; Alesina and Tabellini, 1990; Tabellini and Alesina, 1990). In particular, Alesina and Tabellini (1990) predict that governments, regardless of their party ideology, will issue more debt when facing a higher probability of electoral defeat.

Estimating the predictions of strategic debt models has proven challenging. Since incumbents may use the budget to help their reelection prospects (Rogoff and Sibert 1988; Rogoff, 1990; Shi and Svensson, 2006), proxies of the probability of reelection may be affected by deficits, leading to a reverse causality problem.

This paper exploits the existence of extended interlude periods (i.e., time between election and government change date) from Latin American presidential democracies to identify a causal effect of a change in the probability of electoral defeat on a change in the budget deficit. Namely, since uncertainty about the incumbent's successor is revealed during the interludes, by definition, the probability of electoral defeat becomes, respectively, one or zero if the incumbent was voted out or reelected during the elections. Then, using ex-post electoral outcomes as a proxy of the probability of electoral defeat during the pre-electoral period allows me to construct an exogenous change in that probability between the pre-electoral and the interlude periods to estimate its impact on the change in debt accumulation.

This identification strategy has an added benefit. Estimating the effect of changes in the probability of being defeated between pre-electoral and interlude periods allows me to control for unobserved incumbent characteristics, such as the ability or desire to use the government budget for political gain. This is possible because the incumbent remains in office until government change date, so I observe the same incumbent choosing monthly spending before and after elections *even in the case of an electoral defeat*. Previous studies have tested the same hypothesis using data in levels instead of testing it by using changes (Pettersson-Lidbom, 2001; Lambertini, 2003), in which case such unobserved incumbent characteristic may be an additional concern.

I first formalize the effect of a change in the probability of being defeated on debt accumulation by extending Alesina and Tabellini's (1990) model to include interlude periods. Like in their model, I assume that the individuals are identical except for the preferences over the composition of public expenditure using a separable utility framework. And unlike them, I assume that the incumbents use lump sum taxes fixed to unity. In equilibrium, I find that the higher the increase (decrease) in the probability of being defeated, the higher the increase (decrease) in the deficit between interlude and pre-electoral periods.

Second, I test the theoretical prediction of the model using panel data on federal government deficit from presidential Latin American democracies reported by International Financial Statistics (IFS). I regress monthly deficits on indicators for pre-electoral period interacted with the probability of being voted out, interlude, and interlude if the incumbent and his successor are not from the same party.[1] Consistently with the model, the estimation results show that the deficit increases considerably when the incumbent's party is voted out and this was perceived as unlikely before the elections. By contrast, the deficit *decreases* considerably when the incumbent's party is re-elected and this was perceived as unlikely before the elections. For example, the budget deficit as a percentage of GDP –in absolute value– changes by around 2 percentage points (the mean is 1.6) when the magnitude of the electoral

[1] I consider partisanship as a proxy of similarity in policy choices.

surprise is 0.35 (mean).[2] The deficit does not change significantly when the electoral outcome conforms to expectations. This shows that only large changes in the probability of being defeated generate large changes in the budget deficit; i.e., when a victory or a defeat is perceived as a surprise.

My paper improves on the identification strategy of earlier studies of strategic debt. Pettersson-Lidbom (2001) uses ex post vote shares from Swiss municipalities in order to construct the probability of electoral defeat. To address the reverse causality problem arising if politicians use government budgets to enhance their reelection prospects, he uses municipality fixed effect as instruments because municipality dummies appear to be valid instruments. The unequal dispersion of government change across Swiss municipalities suggests that municipality fixed effects can be used as predictors of the probability of defeat. Thus the fixed effects measure the average frequency of party change and can be interpreted as capturing the latent instability of voters' preferences in a particular municipality. Under this empirical strategy, however, it is not possible to control for fixed unobserved heterogeneity at the municipality level in the second stage estimation, which may lead to biased estimates. Lambertini (2003) uses ex-ante opinion polls (i.e., the fraction of interviewed individuals that would vote in favor of the incumbent if elections were held at the time the poll is taken) as a proxy of the probability of electoral defeat for the United States and OECD countries. Since opinion polls are usually conducted one or two months before elections take place, they may suffer from the same endogeneity problem as the ex post vote shares. Indeed, the literature suggests that incumbents start to enhance

[2] The magnitude of an electoral surprise is calculated as follows: when the incumbent perceives an electoral defeat with probability 0.65 and he indeed lost elections (probability of electoral defeat goes to 1). Or when the incumbent perceives an electoral defeat with probability 0.35 and he finally wins (probability of electoral defeat goes to 0).

their reelection chances up to one year before elections (Brender and Drazen, 2005; Shi and Svensson, 2006; Streb et.al, 2012). In contrast to these papers, my identification strategy is focused on identifying the effects of electoral surprises, which provides a plausible source of exogenous variation.

Substantively, my results suggest that the existence of interludes may generate significant welfare losses. I find significant changes in fiscal deficit –and consequently in public expenditure– due to the electoral surprises. Under this fact, citizens suffer from a decrease in the inter-temporal utility levels because the "principle of consumption smoothing" is violated by the incumbents.

The rest of the paper is organized as follows. In the next section I present the model. In section 3 I describe the data and the econometric specification. In section 4 I present the empirical results, and section 5 the robustness checks. Finally, section 6 concludes.

2.2 Theoretical framework

I consider a simplified version of Alesina and Tabellini's (1990) model. In this economy only lump sum taxes are available (assumed fixed to unity) and citizens have separable utility functions. The current incumbent can carry deficit from one period to the next and all the cumulative debt has to be cancelled in full at the end of the game. I consider three periods in this game: in the first period, the incumbent faces uncertainty about who will be his successor since elections take place at the beginning of the next period. In period two, the incumbent learns who will be his successor but he remains in office until the beginning of period three (this represents the interlude). And in the third period the successor takes office.

Citizen i's utility function in each period is as follows (the incumbent is one of the citizens):

$$v^i(c_t,\, l_t,\, g_t^A,\, g_t^B) = h(c_t) + v(l_t) + \alpha^i u(g_t^A) + (1-\alpha^i)u(g_t^B)$$

where c_t and l_t are private consumption and leisure respectively, and $h^{'}(x) > 0$, $h^{''}(x) < 0$, $v^{'}(x) > 0$, $v^{''}(x) < 0$. Also $u^{'}(x) > 0$, $u^{''}(x) < 0$, $u^{'}(1) = 1$ and g^A and g^B represents spending on public goods A and B respectively. $\alpha^i \in (0,1)$ represents the relative importance consumer i attaches to public good A.

Since the incumbent charges citizens a fixed amount of lump sum taxes equal to one each period, the per-period budget constraint is

$$Total\,time_t = c_t + l_t + 1.$$

Under this scenario, consumption and leisure per period will be the same for all citizens, thus we can focus on public consumption only.

In this economy there are only two parties, D and R. Without loss of generality, assume that $\alpha^D = 1$ and $\alpha^R = 0$, and that party D is the incumbent in period 1. Under this setting, the inter-temporal utility function of party D is

$$V(g_1^A,\, g_2^A,\, g_3^A,\, \delta) = u(g_1^A) + \delta E_1[u(g_2^A)] + \delta^2 E_1[u(g_3^A)].$$

The expectation reflects the uncertainty during period 1 about who will be elected at the beginning of period 2 for taking office at the beginning of period 3. The probability of being reelected is assumed to be exogenous and equal to $(1-p^R)$. Under these assumptions, the incumbent's optimization problem is the following:

$$\max_{\{g_1^A,\, g_2^A,\, \widetilde{g}_2^A,\, g_3^A,\, \widetilde{g}_3^A\}} V(g_1^A,\, g_2^A,\, \widetilde{g}_2^A,\, g_3^A,\, \widetilde{g}_3^A,\, p^R,\, \delta)$$

$$st: \ g_1^A = 1 + d_1; \ g_2^A = 1 + d_2; \ \widetilde{g}_2^A = 1 + \widetilde{d}_2; \ g_3^A = 1 - \tfrac{d_1}{q^2} - \tfrac{d_2}{q}; \ \widetilde{g}_3^A = 1 - \tfrac{d_1}{q^2} - \tfrac{\widetilde{d}_2}{q};$$
$$-1 \leq d_1 \leq q + q^2; -1 \leq d_2 \leq q - \tfrac{d_1}{q}; -1 \leq \widetilde{d}_2 \leq q - \tfrac{d_1}{q}.$$

Where d_1 is the deficit generated during period 1 carried forward to period 3, and d_2 ($\widetilde{d}_2$) is the deficit generated during period 2 after the incumbent learns that his party was re-elected (not re-elected), also carried forward to period 3. δ is the discount factor and $q = \frac{1}{1+r}$ the inverse of the gross interest rate. Note that the deficit during period 1 cannot be greater than the present value of all the future government revenues (i.e., $d_1 \leq q + q^2$), and surplus during period 1 cannot be greater than the current income during period 1 (i.e., $d_1 \geq -1$). The constraints on d_2 and $\widetilde{d}_2$ are derived similarly, taking into consideration the deficit or surplus during period 1. Solving the problem above yields

Proposition 1. *The fiscal deficit in period 2 is larger if the incumbent's party is not reelected ($\widetilde{d}_2^* > d_2^*$).*

Proof. The model is solved using backward induction. When the information about who the successor will be arises, the current incumbent updates during period 2 his inter-temporal consumption path. In particular, if the incumbent's party is not reelected the problem becomes

$$\max_{\widetilde{d}_2} u\left(1 + \widetilde{d}_2\right) \qquad st: \ -1 \leq \widetilde{d}_2 \leq q - \frac{d_1}{q}.$$

This yields the upper corner solution,

$$\widetilde{d}_2^* = q - \frac{d_1}{q}. \tag{2.1}$$

Instead, if the incumbent's party is reelected, the consumption plan is updated following the optimization:

$$\max_{d_2} u\left(1+d_2\right)+\delta u\left(1-\frac{d_2}{q}-\frac{d_1}{q^2}\right) \qquad st:\ -1 \leq d_2 \leq q-\frac{d_1}{q},$$

which yields the following interior solution,

$$d_2^* = \left(1-\frac{d_1}{q^2}-\left[u^{'-1}(\frac{\delta}{q})\right]^{-1}\right)\left(\left[u^{'-1}(\frac{\delta}{q})\right]^{-1}+\frac{1}{q}\right)^{-1}. \tag{2.2}$$

Since the boundaries are the same for both problems, and the solution is interior when the incumbent's party is re-appointed, but it is the upper bound when he is not, then it follows that $\widetilde{d}_2^* > d_2^*$. ■

If the incumbent learns during period 2 that his successor will have different preferences over policy choices, he knows that in period 3 any remaining resources will be used to finance the provision of public goods that he does not favor (represented by public good B). Given this, as an optimal strategy, the incumbent over spends during period 2, providing the public goods he considers important and leaving no resources to the successor. Instead, if he or his party is reelected, he knows the successor will implement the policies he favors. Then, he will want to smooth public consumption by moderating public good provision during period 2, and leave resources for the last period.

In addition, when the information about the successor's identity arrives after the election, motivates incumbents to change suddenly the budget deficit before the government change date. Under this fact, we expect the following:

Proposition 2. *If the incumbent discounts the future at the same rate as the market does (i.e., $\delta = q$), then in equilibrium (i) the deficit generated during period 1 is at least as large as the deficit generated during period 2 if the incumbent's party is reappointed; but (ii) the deficit of period 1 is at most as large as the deficit generated during period 2 if the incumbent's party is not reappointed ($d_2^* \leq d_1^* \leq \widetilde{d}_2^*$). In addition, (iii) d_1^* approaches d_2^* [$\widetilde{d}_2^*$] as the magnitude of a surprise win [defeat] decreases.*

Proof. If $\delta = q$, then d_2^* showed in equation (2.2) becomes,

$$d_2^* = -\frac{d_1}{q\,(q+1)}. \tag{2.3}$$

The optimization problem during the first period is

$$\underset{d_1}{max}\, u\,(1+d_1) + \\ \delta\left[\left(1-p^R\right)u\,(1+d_2^*) + \left(p^R\right)u\left(1+\widetilde{d_2^*}\right) + \delta^2\left(1-p^R\right)u\left(1-\frac{d_2^*}{q}-\frac{d_1}{q^2}\right)\right]$$

$$st:\ -1 \leq d_1 \leq q^2 + q.$$

The first order condition for an interior solution is

$$u^{'}\,(1+d_1) + \delta\left[\left(1-p^R\right)u^{'}\,(1+d_2^*)\,\frac{\mathrm{d}\,(d_2^*)}{\mathrm{d}\,(d_1)} + \left(p^R\right)u^{'}\left(1+\widetilde{d_2^*}\right)\frac{\mathrm{d}\left(\widetilde{d_2^*}\right)}{\mathrm{d}\,(d_1)}\right] + \\ \delta^2\left(1-p^R\right)u^{'}\left(1-\frac{d_2^*}{q}-\frac{d_1}{q^2}\right)\left(-\frac{1}{q^2}-\frac{1}{q}\frac{\mathrm{d}\,(d_2^*)}{\mathrm{d}\,(d_1)}\right) = 0,$$

and after some algebra we get

$$u^{'}\,(1+d_1) = \left(1-p^R\right)u^{'}\,(1+d_2^*) + \left(p^R\right)u^{'}\left(1+\widetilde{d_2^*}\right) \tag{2.4}$$

from which directly follows that $d_2^* \leq d_1 \leq \widetilde{d_2^*}$. ■

Parts *(i)* and *(ii)* of proposition 2 can be interpreted in the following way: since the incumbent does not know whether his party will be re-appointed, he increases the deficit moderately. Instead, when he fully learns in period 2 who will come next, he updates the public consumption plan either by increasing the deficit even more (if a politician with different policy preferences was elected), or by decreasing it (if a politician with similar policy preferences was elected).

Understanding part *(iii)* of proposition 2 requires the inspection of function (2.4) combined with (2.3) and (2.1) for all $p^R \in [0, 1]$. First, for the extreme values $p^R = 0$ and $p^R = 1$ we know that $d_1^* = d_2^* = 0$ and $d_1^* = \widetilde{d_2^*} = q^2(1 + q)$, respectively. In both extreme scenarios, the change in the deficit is zero indicating that knowing exactly what will happen during elections (i.e., electoral victories or defeats are not perceived as surprises) does not alter the public consumption plan, neither the deficit plan. Instead, this plan changes when the result of the upcoming elections is unknown (i.e., the incumbent does not know whether he will lose or not). I show how the change in p^R affects the change in the deficit using two plots: in the first one, I construct the change in the deficit when the probability of being defeated decreases from $p^R \in (0, 1)$ to 0 during the interlude (i.e., the incumbent was then re-elected), shown in panel A of Figure 2; in the second one, I construct the change in the deficit when that probability increases from $p^R \in (0, 1)$ to 1 (i.e., the incumbent was voted out), shown in panel B of Figure 2.

As this figure makes clear, changes in the deficit occur as a result of "surprises" in the electoral outcome, and the higher the magnitude of the electoral surprise the higher the change in the deficit. For example, panel A shows that when the incumbent's party is re-elected but this was perceived as unlikely before the elections, the deficit decreases considerably. Instead, the deficit does not decrease when the incumbent's party is re-elected and this was perceived as the likely outcome. The case when the incumbent's party is voted-out, shown in panel B, can be interpreted similarly.

Note that, in the context of this model, changes in the deficit due to electoral surprises generate significant welfare loses since the principle of public consumption smoothing is violated. Extended interlude periods result in large fluctuations in public spending, especially when the electoral results are unexpected.

Below, I estimate the effect of the magnitude of the electoral surprise on the change in the budget deficit. I expect to find that the higher the magnitude of the electoral surprise the higher the change in the deficit, as stated in proposition 2. Specifically, when the magnitude represents a surprise win (i.e., the incumbent's party unexpectedly won reelection) I expect the deficit to decrease. Instead, when the magnitude represents a surprise defeat (i.e., the incumbent's party unexpectedly lost the election) I expect the deficit to increase.

2.3 Data and Empirical strategy

2.3.1 Data and variable definitions

To test for horizon effects, I use monthly data on Presidential democracies from Latin America during the periods 1980:1-2005:12. The countries in the sample are Argentina, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, Honduras, Mexico, Nicaragua, Panama, Peru, Uruguay and Venezuela. I only consider democratic periods based on Polity IV Project. To define the relevant election date, I use presidential elections following the classification in the Database of Political Institutions (DPI). The electoral calendar and the vote share outcome per party were taken from the Center on Democratic Performance at Binghamton University, SUNY.[3] Presidential elections take place every 4 to 6 years. All these countries have at least one

[3] The probability of being defeated was constructed taking into consideration the vote share of the two parties that got the highest share (it usually represents more than 70% of the total electorate). For example, if the incumbent's party gets the highest vote share, say 40%, and the second party gets 35%, the incumbents vote share is calculated as 0.40/(0.40 + 0.35). Instead, if the incumbent got the third place or below, the corresponding probability is considered as zero.

month of interlude per election, which is necessary in order to have enough within group variation.[4] Table 1 shows the summary statistics of monthly fiscal deficit as a percentage of GDP from the central government and other variables used in the analysis, available at International Monetary Fund's International Financial Statistics. To construct monthly GDP figures in nominal terms, I follow Fernandez (1981) distribution procedure, available in MATLAB, using monthly import series. This allows me to compute the ratio of the budget deficit to GDP on monthly basis. I follow a similar procedure to distribute real GDP.

2.3.2 Methods and econometric specification

Based on Propositions 1 and 2, I estimate the effect of the change in the probability of being defeated between the pre-electoral period and the interlude on the change in the budget deficit between these two periods using the following econometric specification:

$$d_{it} = \sum_{n=1}^{12} \beta_n d_{i,t-n} + \xi \triangle real\, gdp_{it} + \beta_{E\times P} E_{it} \times P_{it} + \beta_{E\times P^2} E_{it} \times P_{it}^2 + \beta_I I_{it} + \beta_{I\times \widetilde{P}} I_{it} \times \widetilde{P}_{it} + \beta_G G_{it} + e_{it} \quad (3.1)$$

[4] If there is a second round election and the incumbent's party runs during that round, the interlude is constructed taking into consideration only the months that lies between the second round and the government change date. Instead, if the incumbent's party does not run during the second round, the interlude is defined as the months that lie between the first round and the government change month. Since the vote share does not chance considerably between the first and second round when the incumbent's party runs both, I only consider the first round vote share for the construction of the probability of electoral defeat.

Here d_{it} is fiscal deficit as percentage of GDP for country i during month t. The first terms on the right hand side are twelve distributed lags.[5] Monthly real GDP growth rate is used as a control. E_{it} is a dummy that takes value 1 for the last nine months before an election at month t takes place. $1-P_{it}$ is the ex-post electoral vote share of the incumbent's party, P_{it} squared is also included to capture the fact that politicians may be risk averse, as assumed in the model of section II. I_{it} takes value one if the month t lies within an interlude period. $\widetilde{P}_{it}$ is a dummy variable that takes value 0 when the incumbent learns that his party was re-appointed, 1 if he learns that his party was voted out. G_{it} is a dummy that takes value 1 for the nine months after constitutional government change in order to test whether the political budget cycle is reversed. Year-country and monthly-year fixed effects are used to capture annual country trend and country seasonal effect, respectively.

Since GDP was disaggregated based on Fernández's (1981) method using monthly imports in USD as an indicator, the disaggregation of GDP can potentially attenuate the estimates since the lagged dependent variables used as regressors include an error term from the disaggregation method itself. I address this using an instrumental variable strategy described in Appendix A.

As stated in the introduction, even though P_{it} is potentially endogenous, the change in that probability from the pre-electoral to interlude periods is given (i.e., $\widetilde{P}_{it} - P_{it}$) allowing me to determine a causal effect on the change of fiscal deficit as a percentage of GDP. Following equation (3.1), the effect of the change in the probability of being voted out on the change in the fiscal deficit is as follows,

$$\triangle d = \beta_{I\times\widetilde{P}}\widetilde{P} + \beta_I - (\beta_{E\times P}P + \beta_{E\times P^2}P^2). \tag{3.2}$$

Based on proposition 1, we expect $\beta_{I\times\widetilde{P}} > 0$, i.e., the fiscal deficit during interludes is larger if the incumbent's party is not reelected. Similarly, based on proposition 2

[5] The selection of the amount of distributed lags is explained in Appendix B.

we expect:

1. $\left(\hat{\beta}_{I\times\widetilde{P}}+\hat{\beta}_I\right)-\left(\hat{\beta}_{E\times P}P_1+\hat{\beta}_{E\times P^2}P_1^2\right)>\left(\hat{\beta}_{I\times\widetilde{P}}+\hat{\beta}_I\right)-\left(\hat{\beta}_{E\times P}P_0+\hat{\beta}_{E\times P^2}P_0^2\right)>0$, for $P_0>P_1>0$, i.e., the higher the magnitude of the surprise defeat, the higher the increase in the deficit; and

2. $\left(\hat{\beta}_{I\times\widetilde{P}}+\hat{\beta}_I\right)-\left(\hat{\beta}_{E\times P}P_0+\hat{\beta}_{E\times P^2}P_0^2\right)\approx 0$, for $P_0\cong 1$, i.e., that no surprises in the electoral defeat does not generate changes in the deficit. Similarly,

3. $0>\left(\hat{\beta}_I\right)-\left(\hat{\beta}_{E\times P}P_1+\hat{\beta}_{E\times P^2}P_1^2\right)>\left(\hat{\beta}_I\right)-\left(\hat{\beta}_{E\times P}P_0+\hat{\beta}_{E\times P^2}P_0^2\right)$, for $P_0>P_1>0$,

4. $\left(\hat{\beta}_I\right)-\left(\hat{\beta}_{E\times P}P_1+\hat{\beta}_{E\times P^2}P_1^2\right)\approx 0$, for $P_1\cong 0$, which can be interpreted similarly to (i) and (ii) for a surprise win.

2.4 Estimation results

2.4.1 Basic model

In this section I show the main empirical finding of the paper. In Column 1 and 2 from Table 2 I regress, as shown in Equation (3.1), fiscal deficit as a percentage of GDP on a dummy for the last nine months before elections interacted with the probability of being defeated ($E\times P$), its squared ($E\times P^2$), a dummy for interludes (I), a dummy for interludes interacted with a dummy that takes value 1 if the incumbent's party was not appointed for next term ($I\times\widetilde{P}$), a dummy for the nine months after government change date (G), and the time varying covariates listed in Equation (3.1) not reported on Table 2. Column 2 differs from Column 1 by the fact that the IV method corrects for the attenuation bias generated by the potential measurement error of the lagged dependent variables, as explained in appendix A.

The results in Table 2 support Proposition 1 (i.e., $\hat{\beta}_{I\times\widetilde{P}}$ is positive and significant). Namely, during the interludes the deficit increases by 4.3-4.9 percentage points when incumbents learn that their successor will be from the opposite party.

To test Proposition 2, I calculate the linear combinations stated in previous section using the estimators of Table 2-Column 2. The first linear combination tests the hypothesis that the larger the magnitude of a surprise electoral win, the larger the decrease in the deficit; while the second one test the hypothesis that the larger the magnitude of a surprise electoral defeat, the larger the increase in the deficit.

The results are shown in Panels A and B of Figure 3. These correspond to the theoretical predictions in Figure 2. In panel A, when the magnitude of a surprise electoral win is small (represented with a small P), the decrease in the deficit is small and non-significant, as shown by the 95% confidence interval (dotted lines). But when the magnitude of a surprise electoral win is large (above 0.5), the decrease in the deficit is large and significant. The intuition behind these results follows from Proposition 2 part (iii). If the incumbent believes that his party's chances of getting reelected are high, it would not be a surprise if he learned during the interlude that his party was indeed re-appointed. As a consequence, the deficit before elections is almost as low as the deficit during interludes ensuring that his successor from the same party has sufficient resources to implement his favored policies. Instead, if the incumbent believes that there is a low chance of being re-elected he increases the budget deficit before elections. Then, learning during the interlude that his party was re-elected is unexpected. As a result, the deficit has to decrease during the interlude in order to leave resources to the successor.

The intuition for a surprise defeat is similar. If the incumbent believes that his party's chances of getting reelected are low, and he is indeed voted out, then the deficit before elections is almost as large as the deficit during interludes. This reflects the incumbent's incentive to strategically restrict the successor's policy choices. Instead, if the incumbent believes that there is a high probability of reelection, learning during the interlude that his party was voted out is perceived as a surprise. Then, in order to restrict the policy choices of his successor, the incumbent has to create a large increase in the deficit during the interlude compared to the pre-electoral period.

The results of Table 2 also show an interesting finding in the context of the literature of political budget cycles (PBCs). Empirical tests of PBCs usually find that fiscal deficit increases before elections and decreases afterwards, offsetting the effect of the initial fiscal expansion (Akhmedov and Zhuravskaya, 2004; Streb et al., 2009; Streb et al., 2012).[6] However, I find in this paper that the fiscal contraction after interlude periods does not fully offset the deficit generated before (the estimator $\hat{\beta}_G$ is small and non-significant). Then, after every election we observe a permanent impact on public debt.

2.4.2 Reelected incumbent vs. successor from the same party

In this section I extend the analysis to study the behavior of reelected incumbents vs. co-partisans during the interlude period. In the model, I assumed that preferences of decision makers from the same party were perfectly aligned. In reality, an incumbent who knows that a co-partisan will replace him next period may exert less effort to control the deficit than if he is reelected himself. In order to test this hypothesis I

[6] Technically, the permanent impact on public debts specially occurs when the probability of being defeated is high before elections. I conduct the linear combination $9\hat{\beta}_G - \left(\hat{\beta}_{I\times\widetilde{P}} + \hat{\beta}_I + 9\left(\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P\right)\right)$ and I find a strong negative coefficient for $P = 0.5$. This implies that the deficit generated during the 9 months before an election takes place plus the deficit generated during the interlude (for the case of one month of interlude) is not offset during the nine months after government change date.

add in Equation (3.1) a dummy for interlude periods when the president is reelected himself ($I \times reelected$). As shown in Table 3, the estimator $\hat{\beta}_{I \times reelected}$ is small and non-significant indicating that there is no difference in the behavior compared with the case when a co-partisan takes office. In this case, the assumption made in the theoretical section II that incumbents care about parties as much as they care about themselves is upheld in the data.[7]

2.5 Robustness checks

In this section, I explore the robustness of the above results by estimating specification (3.1) on different subsamples, and by changing the length of the electoral dummies.

2.5.1 The timing of the pre-electoral fiscal manipulation

In the results above, I use a dummy that takes value 1 for the last nine months before an election takes place interacted with the probability of being defeated at the election date. Since the incumbent's expectation of that probability may be updated as the election gets closer, the fiscal deficit can change each month before elections due to changes in that probability, as argued in the strategic use of debt literature. In that case, I would have biased estimates when using a 9 months dummy variable

[7] These results have to be interpreted with care because the amount of episodes of reelected governments is nine, as shown in Table 1. This happens because only 6 out of the 13 countries used in this study allow one immediate reelection. Those countries are Argentina, Brazil, Colombia, Dominican Republic, Peru and Venezuela. In appendix D I show the dates those countries changed the Constitution in order to either allow or forbid an immediate reelection.

to control for pre-electoral periods. To address this potential issue, I use a dummy variable for the last two months before elections take place, since incumbents would have during those months an accurate idea of the right probability of being defeated at the electoral date, and consequently, would generate a constant deficit. Choosing two months was not an arbitrary decision. I first regress fiscal deficit as percentage of GDP on monthly dummies before elections and covariates,

$$d_{it} = \sum_{n=1}^{12} \beta_n d_{i,t-n} + \xi \triangle real\, gdp_{it} + \beta_{E_9} E_{i9} + ... + \beta_{E_1} E_{i1} + \beta_{E_0} E_{i0} + e_{it}. \tag{5.1}$$

Where E_{ik} is a dummy variable that takes value one if the month is k months away from the electoral month. Second, I perform F-tests for the equality of those dummies' estimators starting from the first two closest electoral month (i.e., $\hat{\beta}_{i0} = \hat{\beta}_{i1}$). Then, if that F-test is not rejected, I test $\hat{\beta}_{i0} = \hat{\beta}_{i1} = \hat{\beta}_{i2}$ and continue this procedure as long as the F-tests are not rejected. I found that the F-test $\hat{\beta}_{i0} = \hat{\beta}_{i1}$ is not rejected at 95% confidence interval. However, the test $\hat{\beta}_{i0} = \hat{\beta}_{i1} = \hat{\beta}_{i2}$ is rejected. This may imply that the deficit associated with the probability of being defeated may be similar only the first two months before elections. Given this, I conclude that incumbents may have an accurate idea of the electoral outcome two months before an election takes place, but perhaps not before. For this reason, I construct the linear combinations as it was done in Figure 2 using the regression results from equation (3.1) but with a dummy of pre-electoral period for the first two months before elections, instead of for the first nine. These results, shown in Figure 3, change little compared with the ones in Figure 2.

2.5.2 Interludes' length

Public expenses might not always be reported in the month they were generated. For instance, some federal expenses may be reported, say, in June while they were generated in April. Given this, the budget deficit during interludes can be contaminated with expenses before elections, and expenses during interludes may be leaked out to the post interlude period. Under this scenario, my results above may contain biased estimates. I keep from the sample only those countries with interludes no shorter than two months in order to minimize this potential measurement error. Those countries are Brazil, Costa Rica, Dominican Republic, Mexico, Panama, Peru and Uruguay.

I regress Equation (3.1) using this subsample and report in Figure 4 the linear combinations as they were done in Figure 2. As it is shown, results in Figure 4 are not significantly different than results in Figure 2.

2.6 Conclusion

Theoretical studies on the strategic use of debt argue that governments issue more debt when facing a higher probability of electoral defeat (Alessina and Tabellini, 1990). Since incumbents may use the budget to help their reelection prospects (Rogoff and Sibert 1988; Rogoff, 1990; Shi and Svensson, 2006), proxies of the probability of reelection may be affected by deficits, leading to a reverse causality problem when trying to test the hypothesis of that theoretical prediction.

This paper improved on the identification strategy of earlier studies of strategic debt. I exploited the existence of extended interlude periods (i.e., time between election and government change date) from Latin American presidential democracies to identify a causal effect of a change in the probability of electoral defeat on a change in the budget deficit. I found that the higher the increase (decrease) in the probability of being defeated, the higher the increase (decrease) in the deficit between interlude and pre-electoral periods (e.g., the budget deficit as a percentage of GDP changes by around 2 percentage points when the probability of electoral defeat changes by 0.35 points –average change in the probability).

Studying the interlude periods is interesting in itself. In practice, different electoral systems work with widely different interlude periods. For example, Mexico has around four months, while Peru has only one. In the former country, there is an ongoing debate about whether shortening interludes is beneficial. It is argued in Mexico that extended interlude periods are harmful because the outgoing incumbent's policies do not currently match citizens' preferences. Also, disagreement between the outgoing and the incoming incumbents constitute risks that have to be reduced.[8]

In contributing to this debate, my finding suggests that the existence of interludes may generate significant welfare losses. I find significant changes in fiscal deficit, and consequently in public expenditure, due to electoral surprises. This implies that citizens may suffer a decrease in inter-temporal utility levels because the principle of consumption smoothing is violated by incumbent politicians. The implication is that an electoral system which includes extended interlude periods is socially costly.

[8] "Acortarán periodo de transición presidencial." El Economista, October 2nd 2012 – 13:08.

Appendix A. Validity of the instrument

I test and show in this appendix the validity of the instruments used in the regressions shown in Column 2 of Tables 2 and 3 since the lagged dependent variables are measured with error. The model to be estimated is

$$y = Y_{-1}\beta_1 + X\beta_2 + u, \tag{A1}$$

where Y_{-1} is a matrix with dimension T (amount of observations) by K (amount of distributed lags). X is a matrix with dimension T by J (amount of exogenous variables). y and Y_{-1} are not observable since the dependent variable is measured with error, but we observe $y^* = y + e$ and $Y^*_{-1} = Y_{-1} + \mathbf{e_{-1}}$.

Using a partitioned regression approach we end up with the following 2×2 system of equations if we attempt to estimate the model (1) using the variables measured with error:

$$\begin{array}{c} \hat{\beta}_2 = \left(X'X\right)^{-1} X'y^* - \left(X'X\right)^{-1} X'Y^*_{-1}\hat{\beta}_1 \\ \hat{\beta}_1 = \left(Y^{*\prime}_{-1}Y^*_{-1}\right)^{-1} Y^{*\prime}_{-1}y^* - \left(Y^{*\prime}_{-1}Y^*_{-1}\right)^{-1} Y^{*\prime}_{-1}X\hat{\beta}_2 \end{array} . \tag{A2}$$

Solving this system for $\hat{\beta}_2$ –assuming that the measurement errors e is not autocorrelated– give us the following,

$$\hat{\beta}_2 \xrightarrow{p} \left(\frac{X'M_{Y^*_{-1}}X}{N}\right)^{-1} \left(\frac{X'M_{Y^*_{-1}}Y_{-1}}{N}\right)\beta_1 + \beta_2, \tag{A3}$$

where $M_{Y^*_{-1}} = \left[I - Y^*_{-1}(Y^{*\prime}_{-1}Y^*_{-1})^{-1}Y^{*\prime}_{-1}\right]$is the residual maker matrix. The measurement error in the lagged dependent variable yields to inconsistent estimates of β_2 since the first terms of equation (A3) does not converge in probability to zero.

I used one instrument for each lagged dependent variable (twelve in total) to solve the inconsistency generated by the measurement error; in which case the system becomes just identified. Z is the matrix of instruments with dimension T×K, corresponding to a just identified system. Rewriting Equation (A1) taking into consideration that the variables used to estimate it are measured with error gives us,

$$y^* = Y^*_{-1}\beta_1 + X\beta_2 + (u - \mathbf{e_{-1}}\beta_1 + e). \tag{A4}$$

Using the matrix of instruments Z defined above yields the following,

Lemma 1. Under the assumptions (i) $\frac{Z_t' e_t}{T} \xrightarrow{p} E\left(Z_t' e_t\right) = 0$; (ii) $\frac{Z_{t-i}' e_{t-j}}{T} \xrightarrow{p} E\left(Z_{t-i}' e_{t-j}\right) = 0$ for all $i < j \in \mathbf{R}^+$; (iii) $\frac{Z_{t-i}' e_{t-j}}{T} \xrightarrow{p} E\left(Z_{t-i}' e_{t-j}\right) = 0$ for all $i > j \in \mathbf{R}^+$; and (iv) $\frac{X_{t-i}' e_{t-j}}{T} \xrightarrow{p} E\left(X_{t-i}' e_{t-j}\right) = 0$ for all $i, j \in \mathbf{R}^+$; the instrumental variable estimation of Equation (5) will yield to consistent estimates of β_2.

The proof for Lemma 1 directly follows estimating Equation (A4) using the matrix of instruments Z under assumptions (i)-(iv).

Even though assumptions (i) to (iv) cannot be tested since e_t is not observable at monthly series, e_t can be estimated with quarterly data since quarterly GDP is reported by IFS for many countries. I estimate the quarterly measurement error by constructing quarterly deficit as a percentage of disaggregated quarterly GDP following also Fernandez's (1981) method, and resting quarterly deficit as a percentage of the "true" GDP. Then, I estimate the correlations to test assumptions (i) to (iii) for quarterly data (since our exogenous variables are elections, by definition they won't be correlated with the error term, then assumption (iv) does not need to be tested). This will give us the answer if the instrument (deficit as % of Resources) is correlated with the error term or not under quarterly data. Results are exposed in Table A1. As shown, all the second sample moments assumed equal to zero in Lemma 1, are indeed very close to zero. The additional assumption to be made is that quarterly data behaves pretty similar to monthly data, which is at least more accurate than directly assuming that assumptions (i)-(iii) are true without any empirical test.

It has to be known that for monthly series, I use twelve distributed lags that correspond to four quarters. For that reason I test assumptions (i) to (iii) only for four lags.

Appendix B. Determination of the number of lags for deficit as a percentage of GDP

In this appendix I show how I selected the amount of lags of the dependent variable to be included as controls in the regression of Equation (3.1). This is a relevant question because omitting any of those lags could potentially generate an omitted variable bias because deficit could increase during pre-electoral periods and decrease afterwards. In such a case, the dummies that controls for the post-electoral periods will be systematically positively correlated with the past deficits since those were higher due to elections, leading to an omitted variable bias. I used the adjusted R2 and the Akaike information criterion to choose the optimal amount of distributed lags. Table B1 shows that the adjusted R2 increases until the 12th lag. However the Akaike information criterion keeps on decreasing even after 12th lags (not reported) but at a very small rate. Under this facts, I considered choosing 12 lags the optimal choice.

Appendix C. Studying the behavior of the time series

In all the study I used 12 lags as the adjusted R2 criterion suggested, shown in Appendix B. In order to examine if this approach misses any salient features of the data, this appendix presents the results of an explicit dynamic regression pretty closed to the one used during all the study:

$$y_{it} = \sum_{n=1}^{12} \beta_n y_{i,t-n} + \sum_{i} \sum_{t=1980}^{2005} year_t^{'} \times country_i + \sum_{i} \sum_{t=Jan}^{Dec} month_t^{'} \times country_i + e_{it}, \quad \text{(C1)}$$

in which I control for all possible annual trend and seasonal effect by country. I use this regression to ascertain that the results during the study are not driven by the way I use the data. After estimating Equation (C1), let's assume that there is a positive shock equal to one at month 0. In Figure (C1) we can observe the behavior of the fiscal deficit throughout the time after the initial shock.

This figure indicates that the process is stable because after 12 months the shock is totally reversed.

Table 1: Definition of variables and descriptive statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
variable	**Description**	**Mean**	**Standard error**	**Min**	**Max**	**Frequency**	**Source**
d	Fiscal deficit as percentage of GDP.	1.6	5.0	-21.4	51.9		IFS
$\triangle real\ gdp$	Real GDP growth rate.	0.3	2.8	-16.6	17.4		IFS
E	Dummy that takes value 1 for the last nine months before an election takes place.	0.2	0.4	0.0	1.0	550	SUNY & Others
$E \times P$	Sum of ex-post electoral vote share excluding president's party vote share as a proxy of probability of being defeated.	0.3	0.2	0.0	0.65		SUNY & Others
I	Dummy that takes value 1 if the month/s lie between election and government change date.	0.0	0.2	0.0	1.0	110	
$I \times \widetilde{P}$	Dummy that takes value 1 if the month lies within the inerlude and the successor is NOT from the incumbent's party.	0.0	0.1	0.0	1.0	73	DPI
$I \times reelection$	Dummy that takes value 1 if the month lies within the interlude and incubent was reelected.	0.0	0.1	0.0	1.0	9	SUNY & Others
G	Dummy that takes value 1 for the nine months after constitutional government change.	0.2	0.4	0.0	1.0	567	SUNY & Others

Notes: Note: IFS refers to IMF International Financial Statistics; DPI to Database of Political Institutions; and SUNY to the Center on Democratic Performance, Binghamton University, SUNY. For calculating all the descriptive statistics non-democratic periods were excluded based on Polity IV project.

Table 2: **Partisanship effect on fiscal deficit around interludes**

	(1)	**(2)**
Estimators\Method	OLS	IV
$\hat{\beta}_{E\times P}$	0.677	1.091
	[1.660]	[1.650]
$\hat{\beta}_{E\times P^2}$	1.736	1.439
	[2.367]	[2.294]
$\hat{\beta}_I$	-1.024	-1.547
	[1.142]	[1.172]
$\hat{\beta}_{I\times\widetilde{P}}$	4.340***	4.899***
	[1.297]	[1.294]
$\hat{\beta}_G$	-0.254	-0.275
	[0.343]	[0.357]
Observations	2,926	2,771
R2 within	0.553	
Number of Countries	13	13

Notes: This table shows that during interludes deficit increases when the incumbent's party is not re-elected. Results come from regressing Equation (3.1). Depened variable is monthly deficit as a % of GDP. Regressions include year-country and month-country fixed effect. Countries used in the regressions are: Argentina, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, Honduras, Mexico, Nicaragua, Panama, Peru, Uruguay and Venezuela. Non-democratic episodes were excluded from the sample based on Polity IV project. Robust standard errors are reported in brackets. *, **, *** indicates significance at 10, 5 and 1% respectively.

Table 3: Partisanship vs. reelection effect on fiscal deficit around interludes

	(1)	**(2)**
Estimators\Method	OLS	IV
$\hat{\beta}_{E\times P}$	0.678	1.092
	[1.661]	[1.651]
$\hat{\beta}_{E\times P^2}$	1.736	1.438
	[2.368]	[2.295]
$\hat{\beta}_{I}$	-1.077	-1.775
	[1.450]	[1.538]
$\hat{\beta}_{I\times\widetilde{P}}$	4.393***	5.128***
	[1.579]	[1.634]
$\hat{\beta}_{I\times reelected}$	0.208	0.801
	[1.781]	[1.793]
$\hat{\beta}_{G}$	-0.254	-0.274
	[0.343]	[0.357]
Observations	2,926	2,771
R2 within	0.553	
Number of Countries	13	13

Notes: This table shows the differential effect during the interludes between a reeleted incumbent and a reelected party. Results come from regressing Equation (3.1). Depened variable is monthly deficit as a % of GDP. Regressions include year-country and month-country fixed effect. Countries used in the regressions are: Argentina, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, Honduras, Mexico, Nicaragua, Panama, Peru, Uruguay and Venezuela. Non-democratic episodes were excluded from the sample based on Polity IV project. Robust standard errors are reported in brackets. *, **, *** indicates significance at 10, 5 and 1% respectively.

Table A1: Second sample moements between deficit as % of resources and the predicted error term

	deficit % resources (t)		predicted error of deficit % GDP (t)
predicted error of deficit % GDP (t)	0.0000414	deficit % resources (t)	0.0000414
predicted error of deficit % GDP (t-1)	0.0000058	deficit % resources (t-1)	0.0000421
predicted error of deficit % GDP (t-2)	0.000054	deficit % resources (t-2)	0.0000151
predicted error of deficit % GDP (t-3)	0.0000456	deficit % resources (t-3)	0.0000142
predicted error of deficit % GDP (t-4)	0.0000423	deficit % resources (t-4)	0.0000187

Notes: Correlations were conducted as defned in Lemma 1.

Table B1: Determining the optimal amount of lags for surplus/GDP

Number of lags	adj R2	AIC
0	0.239	28146.07
1	0.262	27641.58
2	0.275	27187.669
3	0.274	26737.02
4	0.280	26324.53
5	0.294	25897.002
6	0.298	25528.084
7	0.307	25105.73
8	0.327	24638.031
9	0.343	24207.316
10	0.371	23672.709
11	0.419	22958.773
12	**0.423**	22347.842
13	**0.423**	22080.205

Note: Note: the correspondent adj R2 and AIC comes from a regression of the type of Equation (3.1) but changing the amount of distributed lags as indicated in Column 1 of this table.

Appendix D: Dates of changes in term limit's rule

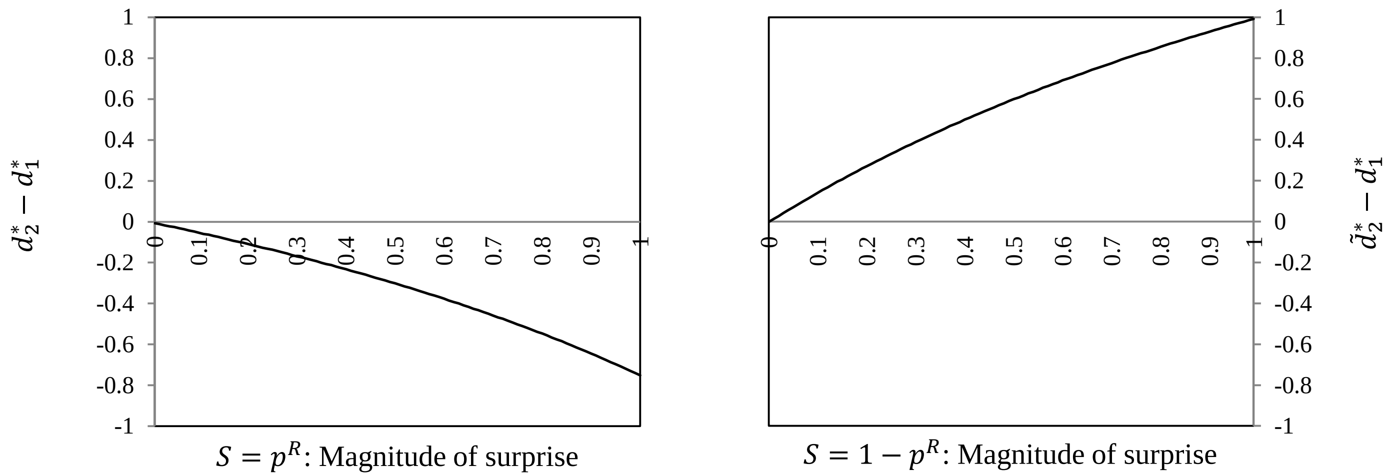
Table D1: Dates of changes in term limit's rule

Country	Immediate reelection was prohibited	Immediate reelection was allowed
Argentina	-	Nov-1993
Brazil	-	Dec-1996
Colombia	-	May-2000
Dom. Rep.	-	May-2002
Peru	-	Dec-1993
Peru	Nov-2000	
Venezuela	-	Dec-1999

Figure 1: Change in the budget deficit in response to electoral surprises

Panel A: Effect of a surprise win on budget deficit.

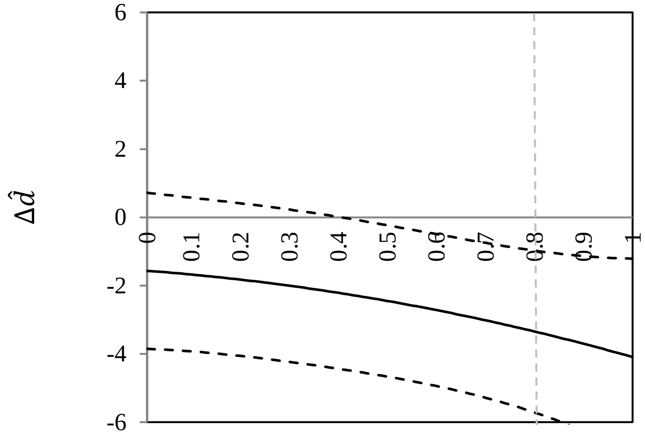
Panel B: Effect of a surprise defeat on budget deficit.



Note: Functions from panel A [and B] were constructed using equations (2.1), (2.2') and (2.3) by calculating d_2^*, $\tilde{d}_2^*$ and d_1 for $q = \delta = 1$ and using $u(x) = ln\, x$. The magnitude of a surprise win [defeat] is the probability of being defeated $p^R = 0$ [one minus the probability of being defeated $1 - p^R$].

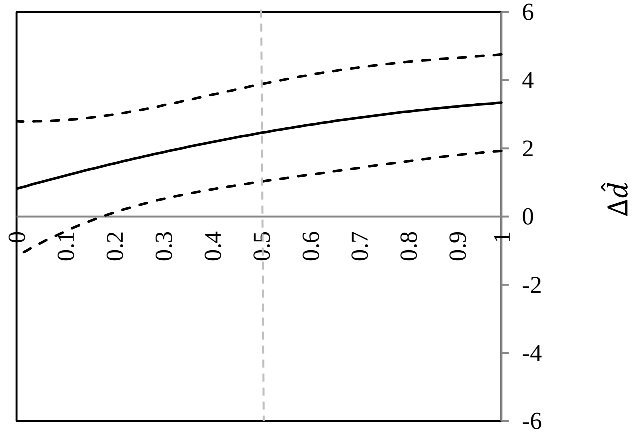
Figure 2: Change in the budget deficit in response to electoral surprises: Empirical test

Panel A: linear combination $\Delta\hat{d} = \hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$. Estimators used are from Table 2 -Column 2.



S: Magnitude of a surprise win. P

Panel B: linear combination $\Delta\hat{d} = \hat{\beta}_{E\times\tilde{P}} + \hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$. Estimators used are from Table 2 -Column 2.

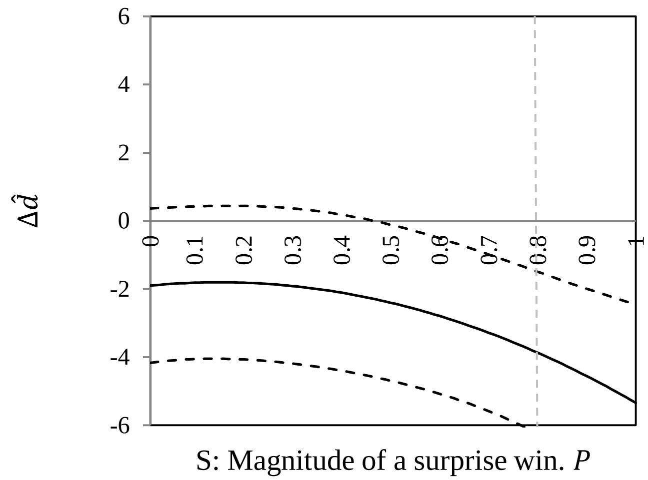


S: Magnitude of a surprise defeat. $1 - P$

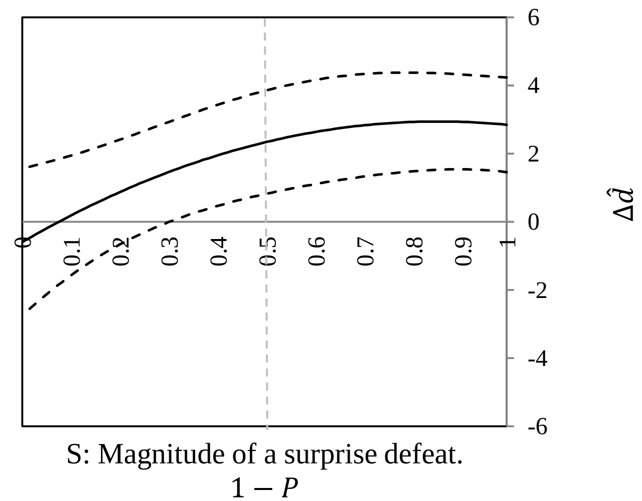
Notes: the solid line represents the linear combination $\hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$ in Panel A $[\hat{\beta}_{I\times\tilde{P}} + \hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$ in Panel B] of the coefficients reported in Column 2 from Table 2 for all $P = [0, 0.01, 0.02, \ldots, 1]$. The dotted lines represent the 95% confidence interval. The x-axis indicates the decrease [increase] in the probability of being defeated between the interlude (by definition $\tilde{P} = 0$ $[\tilde{P} = 1]$) and the pre-electoral period (by definition P is between 0 and 1) for the graph of Panel A [Panel B]. The y-axis indicates the change in the deficit between the interlude and pre-electoral periods as a consequence of the decrease [increase] in the probability of being defeated. The vertical dashed gray line represents the upper bound of the magnitude of the electoral surprise that comes from the data.

Figure 3: Change in the budget deficit in response to electoral surprises: Robustness check 1

Panel A: linear combination $\Delta\hat{d} = \hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$.



Panel B: linear combination $\Delta\hat{d} = \hat{\beta}_{E\times\tilde{P}} + \hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$.

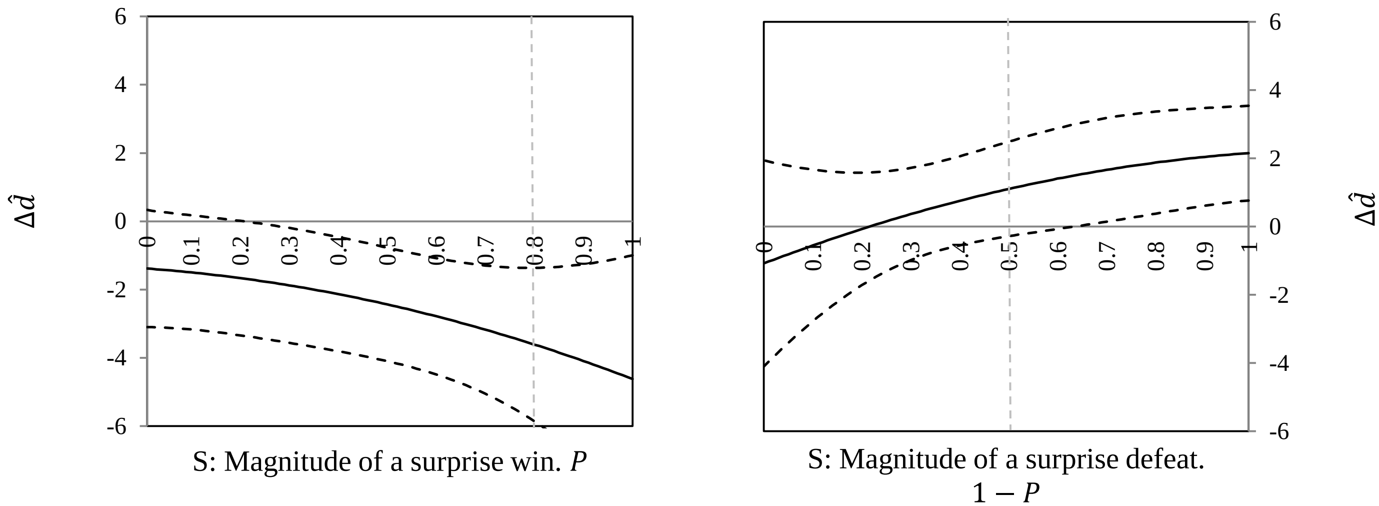


Notes: The linear combinations were conducted based on the estimation of Equation (3.1) but using a pre-electoral dummy of two months instead of nine, as in the previous linear combinations. The solid line represents the linear combination $\hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$ in Panel A [$\hat{\beta}_{I\times\tilde{P}} + \hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$ in Panel B] of the coefficients reported in Column 2 from Table 2 for all $P = [0, 0.01, 0.02, \ldots, 1]$. The dotted lines represent the 95% confidence interval. The x-axis indicates the decrease [increase] in the probability of being defeated between the interlude (by definition $\tilde{P} = 0$ [$\tilde{P} = 1$]) and the pre-electoral period (by definition P is between 0 and 1) for the graph of Panel A [Panel B]. The y-axis indicates the change in the deficit between the interlude and pre-electoral periods as a consequence of the decrease [increase] in the probability of being defeated. The vertical dashed gray line represents the upper bound of the magnitude of the electoral surprise that comes from the data.

Figure 4: Change in the budget deficit in response to electoral surprises: Robustness check 2

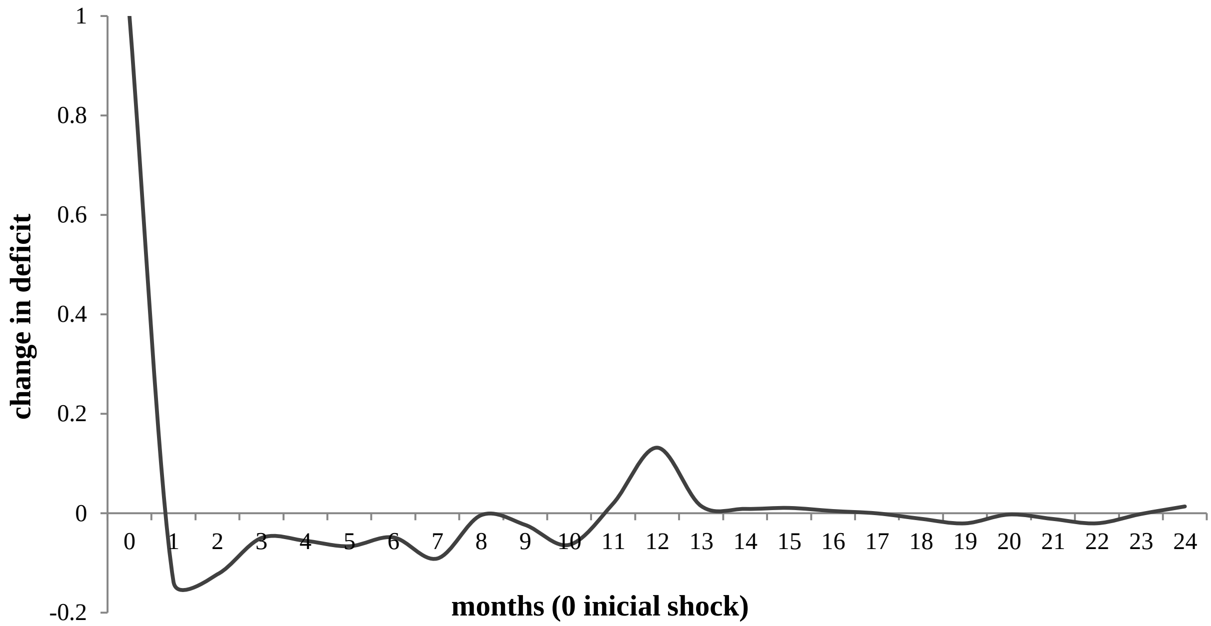
Panel A: linear combination $\Delta\hat{d} = \hat{\beta}_I - \left(\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P\right)$.

Panel B: linear combination $\Delta\hat{d} = \hat{\beta}_{E\times\tilde{P}} + \hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P\ )$.



Notes: The linear combinations were conducted based on the estimation of Equation (3.1) using a subset of countries with the longest interludes. Those countries are Brazil, Costa Rica, Dominican Republic, Mexico, Panama, Peru and Uruguay. The solid line represents the linear combination $\hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$ in Panel A $[\hat{\beta}_{I\times\tilde{P}} + \hat{\beta}_I - (\hat{\beta}_{E\times P^2}P^2 + \hat{\beta}_{E\times P}P)$ in Panel B] of the coefficients reported in Column 2 from Table 2 for all $P = [0, 0.01, 0.02, \ldots, 1]$. The dotted lines represent the 95% confidence interval. The x-axis indicates the decrease [increase] in the probability of being defeated between the interlude (by definition $\tilde{P} = 0$ [$\tilde{P} = 1$]) and the pre-electoral period (by definition P is between 0 and 1) for the graph of Panel A [Panel B]. The y-axis indicates the change in the deficit between the interlude and pre-electoral periods as a consequence of the decrease [increase] in the probability of being defeated. The vertical dashed gray line represents the upper bound of the magnitude of the electoral surprise that comes from the data.

Figure C1: Dynamic behavior of deficit/GDP after initial shock.



Notes: The solid line was constructed generating a shock equal to 1 in the dynamic Equation C1. Dependent variable is monthly deficit as a % of GDP. Regressions include year-country and month-country fixed effect. Countries used in the regressions are: Argentina, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, Honduras, Mexico, Nicaragua, Panama, Peru, Uruguay and Venezuela. Non-democratic episodes were excluded from the sample based on Polity IV project.

Chapter 3

Weak Identification in Dynamic Panel Data Models under Non-stationary Processes

3.1 Introduction

Many economic phenomena require the consideration of both the dynamic nature of a problem and unobserved heterogeneity within groups. For example, consider the dynamic model $y_{it} = \alpha y_{i,t-1} + \eta_i + u_{it}$, $i = 1, ..., N$ and $t = 1, ..., T$. Difference GMM estimation methods have been used, almost exclusively, in the estimation of this type of models because of their advantage that the validity of the moments are independent of the initial condition of the series. Anderson and Hsiao (1981) propose to estimate the model shown above by taking the first difference ($\triangle y_{it} = \alpha \triangle y_{i,t-1} + \triangle u_{it}$, $i = 1, ..., N$ and $t = 1, ..., T-1$) and using the dependent variable lagged for two periods as instrument ($y_{i,t-2}$). Moreover, the dependent variable lagged for more than two periods can also be used as valid instruments generating a more efficient estimation, as shown in Arellano and Bond (1991). Nevertheless, these methods can suffer from weak identification when the time dimension is very short (T).

Blundell and Bond (1998) propose to estimate the dynamic model shown above using instruments in first differences when difference GMM estimators suffer from weak identification.[1] However, stationarity of the process is required in order to have valid moments, which may not hold in many micro-econometric studies.[2]

This paper suggests for both stationary and non-stationary processes a likelihood-based estimator (MLE hereafter) - following Hsiao et al. (2002) - that is consistent

[1] These moment conditions are also shown in Arellano and Bover (1995) and Ahn and Schimidt (1995) but without making explicitly the point of weak identification.

[2] For example, the case of studying the performance of a new drug in clinical trials could follow a non-stationary process since the process starts with the treatment in the very near past.

and behaves much better than any IV-based estimator, especially when the latter fails to identify the parameter of the model due to weak identification issues. This MLE leads to a first difference OLS estimator with a bias corrector term that has the desirable property for identification that OLS estimators possess: a quadratic form in the denominator.

In the next section, I set up the model with a large number of groups (N) and a short time dimension (T), and analyze IV-based estimators (i.e., simple IV, difference GMM and Limited information maximum likelihood estimator) that lead to consistent estimates under the assumptions of the model. In section 3, I highlight explicitly the problem of weak identification under non-stationary processes that comes from using all those IV-based estimators. In section 4, I propose the maximum likelihood estimator based on Hsiao et al. (2002) and I show its properties for very short time periods. In section 5, I present the results of Monte Carlo experiments and in section 6 I conclude emphasizing strengths and weaknesses of the MLE estimator.

3.2 The model and some well known IV-based consistent estimators

I consider the simple AR(1) model with unobserved individual-specific fixed effect

$$y_{it} = \alpha y_{i,t-1} + \eta_i + u_{it}, \tag{2.1}$$

where $\alpha \in (0,1)$, $i = 1, ..., N$ and $t = 1, ..., T$.

Since the data starts being observed in time period 1, we can express the initial observation as a function of unobservables using recursive substitution:

$$y_{i1} = \alpha^m y_{i,1-m} + \frac{1-\alpha^m}{1-\alpha}\eta_i + \Sigma_{l=0}^{m-1}\alpha^l u_{i,1-l}. \qquad (2.2)$$

In which case, is stationary for $m \longrightarrow \infty$.

Furthermore, I will state the following set of assumptions for the data generating process:

$$n_i \sim N(0,\sigma_n^2),\ u_{it} \sim N(0,\sigma_u^2) \text{ for all } i = 1,...,N \text{ and } t = 1,...,T, \qquad (A.1)$$

$$E(n_i u_{it}) = 0 \text{ for all } i = 1,...,N \text{ and } t = 1,...,T, \qquad (A.2)$$

$$E(y_{i1} u_{it}) = 0 \text{ for all } i = 1,...,N \text{ and } t = 2,...,T, \qquad (A.3)$$

$$E(u_{it} u_{i,t+s}) = 0 \text{ for all } i = 1,...,N \text{ and } t = 1,...,T \text{ and } s > 0. \qquad (A.4)$$

Based on these set of assumptions, the following IV-based estimators are consistent:

3.2.1 The Instrumental Variable estimator (IV)

Using the first difference of equation (2.1) will eliminate the fixed effect η_i,

$$\triangle y_{it} = \alpha \triangle y_{i,t-1} + \triangle u_{it}, \qquad (2.3)$$

Anderson and Hsiao (1981) point out the consistency of the following IV estimator that comes from estimating the model (2.3) based on assumptions (A.1)-(A.4), using $y_{i,t-2}$ as an instrument,

$$\hat{\alpha}_{IV} = \left(\sum_{i=1}^{N} \mathbf{y}'_{i,-2} \triangle \mathbf{y}_{i,-1} \right)^{-1} \sum_{i=1}^{N} \mathbf{y}'_{i,-2} \triangle \mathbf{y}_{i}, \tag{2.4}$$

where $\triangle\mathbf{y}_i = (\triangle y_{i3}, ..., \triangle y_{iT})'$, $\triangle\mathbf{y}_{i,-1} = (\triangle y_{i2}, ..., \triangle y_{i,T-1})'$ and $\mathbf{y}_{i,-2} = (y_{i1}, ..., y_{i,T-2})'$ are vectors of dimension $(T-2)\times 1$.[3] Note that since $E(y_{i,t-2}\triangle u_{it}) = 0$, $\hat{\alpha}_{IV}$ is consistent and it does not depend on the initial condition y_{i1} in order to be valid.

3.2.2 Difference GMM estimator (GMM)

This estimator is developed in Arellano and Bond (1991) and exploits all the available lags as instruments, starting from the dependent variable lagged for two periods. Under assumptions (A.1)-(A.4), the following $M = (T-1)\times(T-2)/2$ linear moment restrictions are satisfied: $E[(y_{i,t-j})\triangle u_{it}] = 0$ for all $t = 3, ..., T$ and $j = 2, ..., t-1$.

These moment conditions can be written in vector form as $E[\mathbf{Z}_i'\triangle\mathbf{u}_i] = 0$ with sample moments $N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i'\triangle\mathbf{u}_i \xrightarrow{p} E[\mathbf{Z}_i'\triangle\mathbf{u}_i] = 0$ yielding to the following GMM estimator,

$$\hat{\alpha}_{GMM} = \mathbf{argmin}_{\alpha} \left\{ \sum_{i=1}^{N} \triangle\mathbf{u}'_i \mathbf{Z}_i \mathbf{A_N} \mathbf{Z}_i' \triangle\mathbf{u}_i \right\},$$

where $\Delta\mathbf{u}_i$ is a vector with dimension $(T-2)\times 1$ and $\mathbf{Z}_i = diag(y_{i1}, \ldots, y_{is})$ for $s = (1, \ldots, T-2)$, with dimension $(T-2)\times M$. $\hat{\alpha}_{GMM}$ can be re-expressed as,

[3] Anderson and Hsiao (1981) also proposed to use $\triangle\mathbf{y}_{i,-2}$ as instrument but for very short panels, as I analyze here, the estimator (2.4) is more efficient. See also Arellano (1989) for further discussions.

$$\hat{\alpha}_{GMM} = \left(\sum_{i=1}^{N} \triangle \mathbf{y}'_{i,-1} \mathbf{Z}_i \mathbf{A_N} \mathbf{Z}'_i \triangle \mathbf{y}_{i,-1} \right)^{-1} \sum_{i=1}^{N} \triangle \mathbf{y}'_{i,-1} \mathbf{Z}_i \mathbf{A_N} \mathbf{Z}'_i \triangle \mathbf{y}_i, \qquad (2.5)$$

$\mathbf{A_N}$ can be either $(N^{-1}(\sum_i \mathbf{Z}'_i \mathbf{H} \mathbf{Z}_i)^{-1}$ or $(N^{-1}(\sum_i \mathbf{Z}'_i \triangle \hat{\mathbf{u}}_i \triangle \hat{\mathbf{u}}'_i \mathbf{Z}_i)^{-1}$ generating the first or second step estimator, respectively. $\mathbf{H}$ is a $(T-2) \times (T-2)$ matrix with twos in the diagonal element and minus ones in the lower and upper first sub diagonal, and $\Delta \hat{\mathbf{u}}_i$ comes from estimating the error term using the first step estimator. Note that under the absent of heterokcedasticity and auto-correlation, based on assumptions (A.1) and (A.4), both first and second step estimators are asymptotically equivalent and there is no gains in efficiency of running the second step estimator.[4]

3.2.3 Limited information maximum likelihood estimator (LIML)

Now consider the LIML estimator used in Alonso-Borrego and Arellano (1999) and in Alvarez and Arellano (2003),

[4] An alternative to first differencing is the following forward orthogonal transformation, also known as Helmert's transformation: $\mathbf{u}_i^* = \mathbf{D}^* \mathbf{u}_i$, where $\mathbf{u}_i = (u_{i2}, u_{i3}, ..., u_{iT})'$, and$\mathbf{D}^*$ is a matrix of dimension $(T-2) \times (T-1)$ of the form, $\mathbf{D}^* = diag\left[\frac{T-1}{T}, ..., \frac{1}{2}\right]^{1/2} \times \begin{pmatrix} 1 & -(T-2)^{-1} & -(T-2)^{-1} & \cdots & -(T-2)^{-1} & -(T-2)^{-1} & -(T-2)^{-1} \\ 0 & 1 & -(T-3)^{-1} & \cdots & -(T-3)^{-1} & -(T-3)^{-1} & -(T-3)^{-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix}$. Under this transformation, the weighing matrix under the first step estimator becomes the identity matrix, i.e., $\mathbf{H} = \mathbf{I}$. The GMM estimator is invariant to any orthogonal linear transformation to eliminate the individual effect if all the available moment conditions are exploited. See for further details Arellano and Bover (1995), Alvarez and Arellano (2003) and Okui (2009) for further details.

$$\hat{\alpha}_{LIML} = \mathbf{argmin}_{\alpha} \left\{ \sum_{i=1}^{N} \triangle\mathbf{u}'_i \mathbf{Z}_i \left(\frac{\left(\sum_{i=1}^{N} \mathbf{Z}'_i \mathbf{Z}_i\right)^{-\mathbf{1}}}{\sum_{i=1}^{N} \left(\triangle\mathbf{u}'_i \triangle\mathbf{u}_i\right)} \right) \mathbf{Z}'_i \triangle\mathbf{u}_i \right\},$$

$\hat{\alpha}_{LIML}$ can be rewritten as,

$$\hat{\alpha}_{LIML} = \left(\sum_{i}^{N} \left(\triangle\mathbf{y}'_{i,-1} \mathbf{Z}_i \left(\sum_{i=1}^{N} \mathbf{Z}'_i \mathbf{Z}_i \right)^{-1} \mathbf{Z}'_i \triangle\mathbf{y}_{i,-1} - \hat{l} \left(\triangle\mathbf{y}'_{i,-1} \triangle\mathbf{y}_{i,-1} \right) \right) \right)^{-1} \times \left(\sum_{i}^{N} \left(\triangle\mathbf{y}'_{i,-1} \mathbf{Z}_i \left(\sum_{i=1}^{N} \mathbf{Z}'_i \mathbf{Z}_i \right)^{-1} \mathbf{Z}'_i \triangle\mathbf{y}_i - \hat{l} \left(\triangle\mathbf{y}'_{i,-1} \triangle\mathbf{y}_i \right) \right) \right), \quad (2.6)$$

where

$\hat{l} = minimun\, eigenvalue \; \sum_i^N \{ (\triangle\mathbf{y}_i\,,\, \triangle\mathbf{y}'_{i,-1})' \mathbf{Z}_i \left(\sum_{i=1}^{N} \mathbf{Z}'_i \mathbf{Z}_i \right)^{-1} \mathbf{Z}'_i (\triangle\mathbf{y}_i\,,\, \triangle\mathbf{y}'_{i,-1}) \times \left(\sum_i^N (\triangle\mathbf{y}_i\,,\, \triangle\mathbf{y}'_{i,-1})(\triangle\mathbf{y}_i\,,\, \triangle\mathbf{y}'_{i,-1})' \right)^{-1} \}$. $\underset{N \longrightarrow \infty}{lim} \hat{l} = 0$ indicates the asymptotic equivalence between LIML and IV, and also between LIML with GMM.[5]

3.3 The problem of weak instruments

When the time horizon is very short, $T = 3$, the previous IV-based estimators fail to identify the parameter α, when it has a specific value between zero and one. As

[5] Even though this estimator is still in spirit an IV-based estimator, it differs from IV and GMM in the sense that its asymptotic bias is of order of $1/(2N-T)$ when the process is stationary. Instead, for the same process, simple IV and GMM estimators has asymptotic bias of order $1/N$, as shown by Alvarez and Arellano (2003).

noted, when $T = 3$, all the estimators from section 2 becomes just identified,

$$\hat{\alpha}_{IV} \equiv \hat{\alpha}_{GMM} \equiv \hat{\alpha}_{LIML} = \frac{\sum_{i=1}^{N} y_{i1} \triangle y_{i3}}{\sum_{i=1}^{N} y_{i1} \triangle y_{i2}}. \tag{3.1}$$

The first stage or reduced form equation in this special case can be defined as $\triangle y_{i2} = y_{i1}\pi + \epsilon_i$ for $i = 1, ..., N$, in which case a simple OLS will give a consistent estimation of π,

$$\textbf{plim}\, \hat{\pi} = \left(\alpha^{2m} \sigma^2_{y_{1-m}} + \left(\frac{1-\alpha^m}{1-\alpha} \right)^2 \sigma^2_n + \left(\frac{1-\alpha^{2m}}{1-\alpha^2} \right) \sigma^2_u \right)^{-1} \times$$
$$\left(\alpha^{2m} (\alpha - 1) \sigma^2_{y_{1-m}} + \left(\frac{1-\alpha^m}{1-\alpha} \right) \alpha^m \sigma^2_n + (\alpha - 1) \left(\frac{1-\alpha^{2m}}{1-\alpha^2} \right) \sigma^2_u \right). \tag{3.2}$$

Blundell and Bond (1998) demonstrated the weak identification issue when the process is stationary. Namely, $\hat{\pi} \rightarrow 0$ as $\alpha \rightarrow 1$, when $m \rightarrow \infty$. They proposed moments in levels with instruments in difference in order to estimate the model. However, when the process is not stationary, the moments defined in Blundell and bond (1998) are not longer valid, generating inconsistent estimates. To explore the weakness of those IV-based estimators under non-stationary process, assume $m = 1$,

$$y_{i1} = \alpha y_{i0} + \eta_i + u_{i1} \text{ and } y_{i0} \sim N(0, \sigma^2_{y_0}) \text{ for } i = 1, ..., N\ . \tag{A.5}$$

The estimator of π converges in probability to,

$$\textbf{plim}\, \hat{\pi} = \left(\alpha^2 \sigma^2_{y_{1-m}} + \sigma^2_n + \sigma^2_u \right)^{-1} \left(\alpha^2 (\alpha - 1) \sigma^2_{y_{1-m}} + \alpha \sigma^2_n + (\alpha - 1) \sigma^2_u \right), \tag{3.3}$$

the weak identification still shows up but this time when α is in the neighborhood of 0.57, if we have unitary variances.[6]

The conventional wisdom is that the weak identification under GMM and IV occurs when $\alpha \longrightarrow 1$. However, this is not independent of the initial condition of the process, as shown when the process is not stationary.

3.4 A transformed maximum likelihood-based estimator

After first differencing equation (2.1) and following assumptions (A.1)-(A.5), we can construct a log likelihood function following Hsiao et al. (2002) that satisfies the usual regularity conditions,

$$ln\, L = -\frac{NT}{2} ln(2\pi) - \frac{N}{2} ln|\Omega| - \frac{1}{2}\sum_{i=1}^{N} \triangle\mathbf{u}_i^{\prime}\Omega^{-1}\triangle\mathbf{u}_i, \tag{4.1}$$

where $\triangle\mathbf{u}_i = (\triangle y_{i2} - b,\ \triangle u_{i3}, ...,\ \triangle u_{iT})'$, and $E(\triangle y_{i2}) = b = 0$ for $i = 1, .., N$ due to assumption (A.5). Also,

[6] This issue of weak identification is deeply studied in Nelson and Startz (1990), and Staiger and Stock (1997).

$$\Omega = \sigma_u^2 \begin{pmatrix} var(\triangle y_2)/\sigma_u^2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \iddots & \vdots \\ 0 & -1 & 2 & -1 & 0 \\ \vdots & \iddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix} = \sigma_u^2 \Omega^*.$$

As being the process non-stationary, the matrix Ω does not depend on α, then the principle of the maximum likelihood becomes easier to achieve. For example, if we want to use Newton-Raphtom procedure.

Alternatively, a grid search method can be used starting from small values of α (e.g., $\alpha_0 = 0.0001$ and $\sigma_{u_0}^2 = \frac{1}{2N(T-2)} \sum_{i=1}^{N} \sum_{t=3}^{T} (\triangle y_{it} - \alpha_0 \triangle y_{i,t-1})^2)$ and stopping when the log likelihood reaches its maximum value. This grid search method is easy to achieve due to the fact that $var(\triangle y_2)$ and b do not depend explicitly of α and they can be estimated consistently by averaging the initial differenced observations within groups, namely, $\frac{1}{N-1} \sum_{i=1}^{N} \left(\triangle y_{i2} - \frac{1}{N} \sum_{i=1}^{N} \triangle y_{i2} \right)^2 \equiv \widehat{var}(\triangle y_2) \xrightarrow{p} var(\triangle y_2))$ and $\frac{1}{N} \sum_{i=1}^{N} \triangle y_{i2} \equiv \hat{b} \xrightarrow{p} b$.

To study the behavior of this estimator when $T = 3$, as done for the IV-based estimator, we can estimate α and σ_u^2 by maximizing the likelihood function (4.1), yielding to the following system of two equations,

$$\frac{\partial ln\, L}{\partial \alpha} = 0,\ \alpha = \frac{\sum_{i=1}^{N} \triangle y_{i2} \triangle y_{i3}}{\sum_{i=1}^{N} \triangle y_{i2} \triangle y_{i2}} + \frac{\sigma_u^2}{v\hat{a}r(\triangle y_2)} - \frac{\hat{b}^2 \sigma_u^2}{v\hat{a}r(\triangle y_2) \frac{1}{N} \sum_{i=1}^{N} \triangle y_{i2}^2}, \qquad (4.2)$$

$$\frac{\partial ln\, L}{\partial \sigma_u^2} = -\frac{N}{2}\frac{1}{|\Omega|}\left(3\left(\sigma_u^2\right)^2\left(1+3\left(\frac{v\hat{a}r(\triangle y_2)}{\sigma_u^2}-1\right)\right)-3\left(\sigma_u^2\right)^3\frac{v\hat{a}r(\triangle y_2)}{\left(\sigma_u^2\right)^2}\right)+$$
$$+\frac{1}{4\left(\sigma_u^2\right)^2}\sum_{i=1}^{N}\triangle\mathbf{u}_i^{'}\Omega^{*-1}\triangle\mathbf{u}_i = 0. \quad (4.3)$$

The derivative of the log likelihood function (4.1) with respect to α can easily expressed as an explicit function of the form $\alpha = \nu\left(\sigma_u^2\right)$, as shown in (4.2). Plugging $\alpha = \nu\left(\sigma_u^2\right)$ into (4.3) give us a consistent and efficient estimation of the variance of the model $(\hat{\sigma}_u^2)$.[7] Under this setting, the estimator of α becomes

$$\hat{\alpha}_{MLE} = \frac{\sum_{i=1}^{N}\triangle y_{i2}\triangle y_{i3}}{\sum_{i=1}^{N}\left(\triangle y_{i2}\right)^2}+\frac{\hat{\sigma}_u^2}{v\hat{a}r(\triangle y_2)}-\frac{\hat{b}^2\hat{\sigma}_u^2}{v\hat{a}r(\triangle y_2)\frac{1}{N}\sum_{i=1}^{N}\left(\triangle y_{i2}\right)^2}. \quad (4.4)$$

It is easy to see that this estimator does not suffer from weak identification. Intuitively, because quadratic forms are in the denominator, obviously being always positive.[8] As said in the introduction, the first term of this estimator looks like a simple OLS that comes from the estimation equation (2.3). The second and third term are the bias correction.

[7] Unfortunately, since the process is not stationary, equation (4.3) does not look like a quasi-concave or concave function in which it would have been guaranteed the existence of a unique value of σ_u^2 that satisfies the equality to zero. Under this problem, a numerical procedure should be used and keep the value of $\hat{\sigma_u^2}$ that makes the likelihood function the highest possible. Since we have the presumption that $\alpha\ \epsilon\ (0,1)$, it is not a huge complication.

[8] See that $N^{-1}\sum_{i=1}^{N}\left(\triangle y_{i2}\right)^2 \xrightarrow{p} (\alpha-1)^2\left(\alpha^2\sigma_{y_0}^2+\sigma_n^2+\sigma_u^2\right)+(2\alpha-1)\,\sigma_n^2+\sigma_u^2$, which give us enough variation to identify the parameter strongly, avoiding having weak or closed to zero denominators like in the case of IV-based estimators.

3.5 Monte Carlo simulations

This section presents the results of a series of the Monte Carlo experiments.[9] I observe the finite sample performance of the estimators presented in sections 2 and 4 following assumptions (A.1)-(A.5). I construct the average bias of the estimators, the standard error and the root mean square error for 500 simulations.

In Table 1 can be seen the results for a just identified case in which all the IV-based estimators presented in section 2 are equivalent to the simple IV estimator showed in section 2.1. Column 1 reports the result of the transformed likelihood-based estimator of section 4, meanwhile Column 2 presents the result of the traditional IV estimator. As seen from Table 1 Column 2, when α is closed to the singularity point (i.e., $\alpha \approx 0.57$) described in Nelson and Startz (1990), the IV behaves poorly. For example, when $\alpha = 0.50$, the bias in IV is -0.97, something close to a bias of 194% in terms of the real value of the parameter α. Instead, under MLE, the bias is comparatively nil. Being away from the singularity point area, when α is equal to 0.1 or 0.9, both estimators behaves well. However, the root mean squared error is much smaller under MLE.

In Table 2 can be seen the results when the system is over identified in the case of GMM and LIML (panel A for $T = 4$ and panel B for $T = 5$) are presented. It is important to note that both GMM and LIML behaves poorly not only in a very closed neighborhood of the singularity point, but even when α is $\pm$ 0.2 away from it. Namely, when α is 0.4 in panel A, the bias as a percentage of α is 3% for MLE, 12% for IV, and 17, 16 and 20% for GMM1, GMM2 and LIML respectively. RMSE is pretty low in MLE, and 10 times greater for LIML and around 4.5 times for IV,

[9] The simulations were conducted on Gauss 6.0.

GMM1 and GMM2. This tests the efficiency and accuracy of the MLE against all the other IV-based estimators. The differences between MLE and the other IV-based estimators becomes enormous around the singularity point for both, panels A and B. Specifically, when α is 0.6, the bias is not a mayor concerned in IV, but it is under GMM or LIML. Instead, the size of the RMSE is a enormous in all IV-based estimators, compared with the MLE, specially in the case of IV and LIML. In panel B, when $T = 5$, things start to behave better in both GMM and LIML because both methods exploit all the available instruments (3). However, the bias and the RMSE is still huge compare with MLE. In part, because further lags as instruments would not be as informative as the first lag in the case of informative first lag instruments. For that reason, GMM and LIML still behaves pretty bad when T becomes larger compared with MLE.

3.6 Conclusion

This paper shows that IV-based estimators (i.e., difference GMM and LIML) may estimate the parameters of dynamic panel data model when the process is not stationary as weakly as when it is stationary. When the time horizon is the shortest necessary for identification (i.e., $T = 3$), the IV-based estimators become just identified and equivalent. I show in this case the singularity point defined by Nelson and Startz (1990). A transformed MLE is proposed to solve the weak identification problem generated under IV-based estimators. In the case of over-identified systems with short time series (i.e., $T = 4$, $T = 5$), difference GMM and LIML still shows problems identifying the parameter of the dynamic panel data model. The intuition is that further lags as instruments are not informative enough to offset the weak identification problem generated by the first instrument. The likelihood method definitely solves the weak identification issue and the RMSE is smaller than the ones under GMM and LIML, even when there is no weak identification under the latter two. The drawback of the MLE is that it is sensitive to misspecification of the initial condition for short time periods. For this reason, researchers should inspect carefully the initial condition.

Non-linear moment conditions for the GMM estimator might have solved the weak identification problem as well. I did not explicitly stated it in this paper because the non-linear moment conditions defined by Ahn and Schmidt (1995) require the time length to be no shorter than 4, for which case it would not have been feasible to estimate the dynamic model for $T = 3$. Also, when difference GMM and LIML estimate the parameter without weak identification problems, if $T = 4$, the amount of observations in the non-linear model condition would have decreased by N, generating inefficient estimates.

Table 1. Monte Carlo simulation. $N = 100$, $T = 3$.

		MLE	IV
	$\hat{\alpha} - \alpha$	0.021	0.014
$\alpha = 0.1$	$\widehat{\sigma}(\hat{\alpha})$	[0.107]	[0.270]
	$RMSE$	(0.109)	(0.270)
	$\hat{\alpha} - \alpha$	0.013	-0.675
$\alpha = 0.4$	$\widehat{\sigma}(\hat{\alpha})$	[0.139]	[16.787]
	$RMSE$	(0.140)	(16.801)
	$\hat{\alpha} - \alpha$	0.012	-0.974
$\alpha = 0.5$	$\widehat{\sigma}(\hat{\alpha})$	[0.135]	[15.575]
	$RMSE$	(0.135)	(15.606)
	$\hat{\alpha} - \alpha$	0.011	0.093
$\alpha = 0.6$	$\widehat{\sigma}(\hat{\alpha})$	[0.133]	[13.643]
	$RMSE$	(0.133)	(13.643)
	$\hat{\alpha} - \alpha$	0.009	0.106
$\alpha = 0.9$	$\widehat{\sigma}(\hat{\alpha})$	[0.130]	[0.466]
	$RMSE$	(0.130)	(0.478)

Notes: 500 simulations. $m = 1$, $\sigma^2_{y_{1-m}} = 1$, $\sigma^2_n = 1$, $\sigma^2_u = 1$. Values for α are given in column 1.

Table 2. Monte Carlo simulation. Over identified models.

		Panel A: $N = 100$, $T = 4$					Panel B: $N = 100$, $T = 5$				
		MLE	IV	GMM1	GMM2	LIML1	MLE	IV	GMM1	GMM2	LIML1
	$\hat{\alpha} - \alpha$	0.006	0.021	0.001	-0.003	0.027	-0.004	0.001	-0.017	-0.017	0.001
$\alpha = 0.1$	$\widehat{\sigma}(\hat{\alpha})$	[0.079]	[0.152]	[0.148]	[0.152]	[0.166]	[0.062]	[0.109]	[0.103]	[0.109]	[0.124]
	$RMSE$	(0.079)	(0.153)	(0.148)	(0.152)	(0.168)	(0.062)	(0.109)	(0.105)	(0.110)	(0.124)
	$\hat{\alpha} - \alpha$	-0.012	0.049	-0.068	-0.064	0.084	-0.012	0.020	-0.062	-0.059	0.024
$\alpha = 0.4$	$\widehat{\sigma}(\hat{\alpha})$	[0.088]	[0.465]	[0.339]	[0.352]	[0.845]	[0.070]	[0.231]	[0.198]	[0.214]	[0.266]
	$RMSE$	(0.089)	(0.468)	(0.345)	(0.358)	(0.849)	(0.071)	(0.232)	(0.207)	(0.222)	(0.267)
	$\hat{\alpha} - \alpha$	-0.012	0.090	-0.270	-0.282	-0.589	-0.020	0.072	-0.173	-0.181	0.161
$\alpha = 0.5$	$\widehat{\sigma}(\hat{\alpha})$	[0.090]	[4.035]	[0.551]	[0.587]	[12.996]	[0.071]	[2.290]	[0.290]	[0.318]	[3.984]
	$RMSE$	(0.091)	(4.036)	(0.614)	(0.651)	(13.009)	(0.074)	(2.291)	(0.337)	(0.366)	(3.987)
	$\hat{\alpha} - \alpha$	-0.019	0.076	-0.519	-0.526	-1.097	-0.018	-0.954	-0.412	-0.446	-0.487
$\alpha = 0.6$	$\widehat{\sigma}(\hat{\alpha})$	[0.082]	[10.660]	[0.901]	[0.938]	[12.228]	[0.070]	[21.316]	[0.386]	[0.420]	[8.123]
	$RMSE$	(0.084)	(10.661)	(1.040)	(1.076)	(12.278)	(0.072)	(21.337)	(0.565)	(0.613)	(8.137)
	$\hat{\alpha} - \alpha$	-0.004	0.002	-0.024	-0.028	0.004	-0.003	0.007	-0.021	-0.022	0.004
$\alpha = 0.9$	$\widehat{\sigma}(\hat{\alpha})$	[0.071]	[0.164]	[0.153]	[0.154]	[0.180]	[0.053]	[0.118]	[0.102]	[0.106]	[0.126]
	$RMSE$	(0.071)	(0.164)	(0.154)	(0.157)	(0.180)	(0.054)	(0.118)	(0.104)	(0.108)	(0.126)

Notes: 500 simulations. $m = 1$, $\sigma^2_{y_{1-m}} = 1$, $\sigma^2_n = 1$, $\sigma^2_u = 1$. Values for α are given in column 1.

References

[1] Ahn, Seung. C., and Peter Schmidt. 1995. "Efficient estimation of models for dynamic panel data". Journal of Econometrics, 68: 5-27.

[2] Alesina, Alberto, and Guido Tabellini. 1990. "A positive theory of fiscal deficits and government debt." Review of Economic Studies, 57: 403-414.

[3] Alonso-Borrego, Cesar, and Manuel Arellano. 1999. "Symmetrically normalized instrumental-variable estimation using panel data." Journal of Business and Economic Statistics, 17: 36-49.

[4] Javier Alvarez, and Manuel Arellano. 2003. "The time series and cross-section asymptotics of dynamic panel data estimators." Econometrica, 71: 1121-1159.

[5] Anderson, T.W., and Cheng Hsiao. 1981. "Estimation of dynamic models with error components." Journal of the American Statistical Association, 76: 598-606.

[6] Arellano, Manuel. 1989. "A note on the Anderson-Hsiao estimator for panel data." Economics Letters, 31: 337-341.

[7] Arellano, Manuel, and Olympia Bover. 1995. "Another look at the instrumental variable estimation of the error-components models." Journal of Econometrics, 68: 29-52.

[8] Arellano, Manuel, and Stephen R. Bond. 1991. "Some test of specification for panel data: Monte Carlo evidence and an application to employment equations." Review of Economic Studies, 58: 277-297.

[9] Akai, Nobuo, and Masayo Sakata. 2002. "Fiscal decentralization contributes to economic growth: evidence from state-level cross-sectional data for the United States." Journal of Urban Economics, 52: 93-108.

[10] Akhmedov, Akhmed, and Ekaterina Zhuravskaya. 2004. "Opportunisctic political cycles: Test in a young democracy setting." Quarterly Journal of Economics, 119:1301-1338.

[11] Ansolabehere, Stephen, Alan Gerber, and Jim Snyder. 2002. "Equal votes, equal money: court-ordered redistricting and public expenditures in the American states." American Political Science Review, 96: 767-777.

[12] Ansolabehere, Stephen, and James M. Snyder, Jr. 2006. "Party control of state government and the distribution of public expenditures." Scandinavian Journal of Economics, 108: 547-569.

[13] Albouy, David. 2013. "Partisan representation in congress and the geographic distribution of federal funds." Review of Economics and Statistics, 95: 127-141.

[14] Arulampalam, Wiji, Sugato Dasgupta, and Amrita Dhillon. 2009. "Electoral goals and center-state transfers: A theoretical model and empirical evidence from India." Journal of Development Economics, 88: 103-119.

[15] Baicker , Katherine, Jeffrey Clemens, and Monica Singhal. 2012. "The rise of the states: U.S. fiscal decentralization in the postwar period." Journal of Public Economics, 96:1079-1091.

[16] Barro, Robert. 1979. "On the determination of the public debt." Journal of Political Economy, 87: 940-947.

[17] Berry Christopher R., Barry C. Burden, and William G. Howell. 2010. "The president and the distribution of federal spending." American Political Science Review, 104: 783-799.

[18] Besley, Timothy, and Stephen Coate. 2003. "Centralized versus decentralized provision of local public goods: a political economy approach." Journal of Public Economics, 87: 2611-2637.

[19] Blundell, Richard, and Stephen Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models". Journal of Econometrics, 87: 115-143.

[20] Brender, Adi, and Allan Drazen. 2005. "Political budget cycles in new versus established democracies." Journal of Monetary Economics, 52: 1271-1295.

[21] Bun, Maurice, and Martin Carree. 2005. "Bias-corrected estimation in dynamic panel data models." Journal of Business and Economic Statistics, 23: 200-210.

[22] Case, Anne. 2001. "Election goals and income redistribution: recent evidence from Albania." European Economic Review, 45: 405-423.

[23] Crain, Mark, and Robert Tollison. 1993. "Time inconsistency and fiscal policy. Empirical analysis of US States, 1969-89." Journal of Public Economics, 51: 153-159.

[24] Coleman, John J. 1999. "Unified government, divided government, and party responsiveness." American Political Science Review, 93: 299-309.

[25] Cox, Gary, and Mathew McCubbins. 1986. "Electoral politics as a redistributive game". The Journal of Politics, 48: 370-389.

[26] Craig, Steven, and Robert Inman. 1982. "Aid and public education: an empirical look at the new fiscal federalism." Review of Economics and Statistics, 64: 541-552.

[27] Davoodi, Hamid, and Hens-fu Zou. 1998. "Fiscal decentralization and economic growth: a cross-country study." Journal of Urban Economics, 43: 244-257.

[28] Dixit, Avinash, and John Londregan. 1996. "The determinants of success of special interests in redistributive politics." The Journal of Politics, 58: 1132-1155.

[29] Dixit, Avinash, and John Londregan. 1998. "Fiscal federalism and redistributive politics." Journal of Public Economics, 68: 153-180.

[30] Drazen, Allan. 2000. "Political economy in macroeconomics." Princeton University Press.

[31] Fernández, Roque. 1981. "A methodological note on the estimation of time series." Review of Economics and Statistics, 63: 471-476.

[32] Frederickson, Gerorge, and Yong Cho. 1974. "Legislative reapportionment and fiscal policy in the American states." Western Political Quarterly, 27: 5-37.

[33] Grossman, Philip. 1994. "A political theory of intergovernmental grants." Public Choice, 78: 295-303.

[34] Hahn, Jinyong. 1999. "How informative is the initial condition in the dynamic panel model with fixed effect?" Journal of Econometrics, 93: 309-326.

[35] Howell, William, Scott Adler, Charles Cameron, and Charles Riemann. 2000. "Divided government and the legislative productivity of Congress, 1945-1994." Legislative Studies Quarterly, 25: 285-312.

[36] Howell, William and David Lewis. 2002. "Agencies by presidential design." The Journal of Politics, 64: 1095-1114.

[37] Hsiao, Cheng, Hashem Pesaran, and Kamil Tahmiscioglu. 2002. "Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods." Journal of Econometrics, 109: 107-150.

[38] Iimi, Atsushi. 2005. "Decentralization and economic growth revisited: an empirical note." Journal of Urban Economics, 57: 449-461.

[39] Knight, Brian. 2002. "Endogenous Federal grants and crowd-out of state government spending: Theory and evidence from the Federal Highway Aid Program." American Economic Review, 92: 71-91.

[40] Larcinese, Valentino, Leonzio Rizzo, and Cecilia Testa. 2006. "Allocating the U.S. Federal budget to the states: the impact of the President." The Journal of Politics, 68: 447-456.

[41] Lambertini, Luisa. 2003. "Are budget deficits used strategically?" Manuscript. Boston College.

[42] Levitt, Steven D., and James M. Snyder, Jr. 1995. "Political parties and the distribution of Federal outlays." The American Journal of Political Science, 39: 958-980.

[43] Levitt, Steven D., and James M. Snyder, Jr. 1997. "The impact of Federal spending on House election outcomes." Journal of Political Economy, 105: 30-52.

[44] Lindbeck, Assar, and Jörgen W. Weibull. 1987. "Balanced-budget redistribution as the outcome of political competition." Public choice, 52: 273-297.

[45] Lindbeck, Assar, and Jörgen W. Weibull. 1993. "A model of political equilibrium in a representative democracy." Journal of Public Economics, 51: 195–209.

[46] Lockwood, Ben. 2002. "Distributive politics and the costs of centralization." Review of Economic Studies, 69: 313-337.

[47] Lucas, Robert, and Nancy Stokey. 1983. "Optimal fiscal and monetary policy in an economy without capital." Journal of Monetary Economics, 17: 117-134.

[48] Nelson, Charles, and Richard Startz. 1990. "The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one." Journal of Business, 63: 125-140.

[49] McCarty, Nolan. 2000. "Presidential pork: executive veto power and distributive politics." American Political Science Review, 94: 117-129.

[50] Musgrave, Richard A. 1959. "The theory of public finance." New York: McGraw-Hill.

[51] Oates, Wallace E. 1972. "Fiscal federalism." New York: Harcourt Brace Jovanovich.

[52] Oates, Wallace E. 1999. "An essay on fiscal federalism." Journal of Economic Literature, 37: 1120-1149.

[53] Oates, Wallace E. 2005. "Toward a second-generation theory of fiscal federalism." International Tax and Public Finance, 12: 349-373.

[54] Okui, Ryo. 2009. "The optimal choice of moments in dynamic panel data models." Journal of Econometrics, 151: 1-16.

[55] Persson, Torsten, and Lars E. O. Svensson. 1989. "Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences." Quarterly Journal of Economics, 104: 325-345.

[56] Pettersson-Lidbom, Per. 2001. "An empirical investigation of the strategic use of debt." Journal of Political Economy, 109: 570-583.

[57] Rogoff, Kenneth. 1990. "Equilibrium political budget cycles." American Economic Review, 80: 21-36.

[58] Rogoff, Kenneth, and Anne Sibert. 1988. "Elections and macroeconomic policy cycles." Review of Economic Studies, 55: 1-16.

[59] Shi, Min, and Jacob Svensson. 2006. "Political budget cycles: Do they differ across countries and why?" Journal of Public Economics, 90: 1367-1389.

[60] Staiger, Douglas, and James Stock. 1997. "Instrumental variables regression with weak instruments." Econometrica, 65: 557-586.

[61] Stock, James, Jonathan Wright, and Yogo Motohiro. 2002. "A survey of weak instruments and weak identification in generalized method of moments." Journal of Business and Economic Statistics, 20: 518-529.

[62] Streb, Jorge M., Daniel Lema, and Pablo Garofalo. 2012. "Temporal aggregation in political budget cycles." Journal of LACEA Economia, 13: 39-78.

[63] Streb, Jorge M., Daniel Lema, and Gustavo Torrens. 2009. "Checks and balances on political budget cycles: Cross-Country Evidence." Kyklos, 62: 426-447.

[64] Strömberg, David. 2008. "How the electoral college influences campaigns and policy: The probability of being Florida." American Economic Review, 98: 769-807.

[65] Tabellini, Guido, and Alberto Alesina. 1990. "Voting on the budget deficit." American Economic Review, 80: 37-49.

[66] Ujhelyi, Gergely. 2013. "Civil service rules and policy choices: evidence from US state governments." American Economic Journal: Economic Policy, forthcoming.

[67] Volden, Craig. 2005. "Intergovernmental political competition in American Federalism." American Journal of Political Science, 49: 327-342.

[68] Wildasin, David E. "Intergovernmental transfers to local government." Institute for Federalism and Intergovernmental Relations. No. 2009-11, August 2009.

[69] World Bank, World Development Report 1999/2000, Oxford Univ. Press, New York, 2000.

[70] Xie, Danyang, and Heng-fu Zou. 1999. "Fiscal decentralization and economic growth in the United States." Journal of Urban Economics, 45: 228-239.